S5810-48TS-P Switch Configuration Guide

Model: S5810-48TS-P

Configuration Guide

Contents

Chapter1 System Configuration 1
1 Configuring CLI 2
2 Configuring Basic Management 14
3 Configuring Lines 44
4 Configuring Time Range 61
5 Configuring HTTP Service 65
6 Configuring Syslog 82
7 Configuring CWMP 124
8 Configuring Monitoring 144
9 Configuring Module Hot Swapping 146
10 Configuring Supervisor Module Redundancy 152
11 Configuring Package Management 161
12 Configuring PoE 172
Chapter2 Ethernet Switching Configuration 189
1 Configuring Interfaces 190
2 Configuring MAC Address 230
3 Configuring Aggregated Port 246
4 Configuring VLAN 264
5 Configuring Super VLAN 280
6 Configuring MAC VLAN 289
7 Configuring Protocol VLAN 297
8 Configuring Private VLAN 308
9 Configuring Voice VLAN 326
10 Configuring MSTP 348
11 Configuring GVRP 412
12 Configuring LLDP 428
13 Configuring QinQ 464
14 Configuring ERPS 501
Chapter3 IP Address & Application Configuration 536
1 Configuring IP Addresses and Services 537
2 Configuring ARP 556
3 Configuring IPv6 579
4 Configuring DHCP 612
5 Configuring DHCPv6 640
6 Configuring DNS 656
7 Configuring FTP Server 662
8 Configuring FTP Client 671
9 Configuring TFTP 681
10 Configuration Tunnel 687
11 Configuring Network Communication Test Tools 701
12 Configuring TCP 712
13 Configuring IPv4/IPv6 REF 723

Chapter4 IP Routing Configuration **727**
1 Configuring RIP 728
2 Configuring OSPFv2 773
3 Configuring OSPFv3 851
4 Configuring IS-IS 917
5 Configuring BGP 1000
6 Configuring PBR 1103
7 Configuring VRF 1144
8 Configuring RIPng 1166
9 Managing Routes 1187
10 Configuring Keys 1205
11 Configuring Routing Policies 1211
Chapter5 MPLS Configuration **1230**
1 Configuring MPLS 1231
2 Configuring BGP/MPLS IP VPN 1297
3 Configuring MPLS GR 1476
4 Configuring MPLS BFD 1497
Chapter6 Multicast Configuration **1528**
1 Configuring IP Multicast 1529
2 Configuring IPv6 Multicast 1537
3 Configuring IGMP 1560
4 Configuring MLD 1582
5 Configuring PIM-DM 1603
6 Configuring PIM-SM 1622
7 Configuring PIM-SMv6 1678
8 Configuring IGMP Snooping 1730
9 Configuring MLD Snooping 1776
10 Configuring MSDP 1790
Chapter7 Security Configuration **1825**
1 Configuring AAA 1826
2 Configuring RADIUS 1869
3 Configuring TACACS+ 1891
4 Configuring 802.1X 1902
5 Configuring Web Authentication 1952
6 Configuring SCC 2004
7 Configuring Global IP-MAC Binding 2019
8 Configuring Password Policy 2027
9 Configuring Port Security 2033
10 Configuring Storm Control 2045
11 Configuring SSH 2051
12 Configuring URPF 2090
13 Configuring CPP 2104
14 Configuring DHCP Snooping 2115
15 Configuring DHCPv6 Snooping 2131

16 Configuring ARP Check.....2149
17 Configuring Dynamic ARP Inspection.....2155
18 Configuring IP Source Guard.....2161
19 Configuring IPv6 Source Guard.....2167
20 Configuring Gateway-targeted ARP Spoofing Prevention.....2174
21 Configuring NFPP.....2179
22 Configuring DoS Protection.....2233
Chapter8 ACL & QoS Configuration.....2240
1 Configuring the ACL.....2241
2 Configuring QoS.....2291
3 Configuring MMU.....2323
Chapter9 Reliability Configuration.....2332
1 Configuring REUP.....2333
2 Configuring RLDP.....2358
3 Configuring DLDP.....2370
4 Configuring VRRP.....2378
5 Configuring VRRP Plus.....2424
6 Configuring BFD.....2435
7 Configuring IP Event Dampening.....2456
8 Configuring Stacking.....2461
9 Configuring RNS.....2498
Chapter10 Network Management & Monitoring Configuration.....2525
1 Configuring SNMP.....2526
2 Configuring RMON.....2552
3 Configuring NTP.....2567
4 Configuring SNTP.....2581
5 Configuring SPAN-RSPAN.....2587
6 Configuring ERSPAN.....2604
7 Configuring sFlow.....2613

Chapter1 System Configuration

1. Configuring CLI
2. Configuring Basic Management
3. Configuring Lines
4. Configuring Time Range
5. Configuring HTTP Service
6. Configuring Syslog
7. Configuring CWMP
8. Configuring Monitoring
9. Configuring Module Hot Swapping
10. Configuring Supervisor Module Redundancy
11. Configuring Package Management
12. Configuration PoE

1 Configuring CLI

1.1 Overview

The command line interface (CLI) is a window used for text command interaction between users and network devices. You can enter commands in the CLI window to configure and manage network devices.

Protocols and Standards

N/A

1.2 Applications

Application	Description
Configuring and Managing Network Devices Through CLI	You can enter commands in the CLI window to configure and manage network devices

1.2.1 Configuring and Managing Network Devices Through CLI

Scenario

As shown in Figure 1-1, a user accesses network device A using a PC, and enter commands in the CLI window to configure and manage the network device.

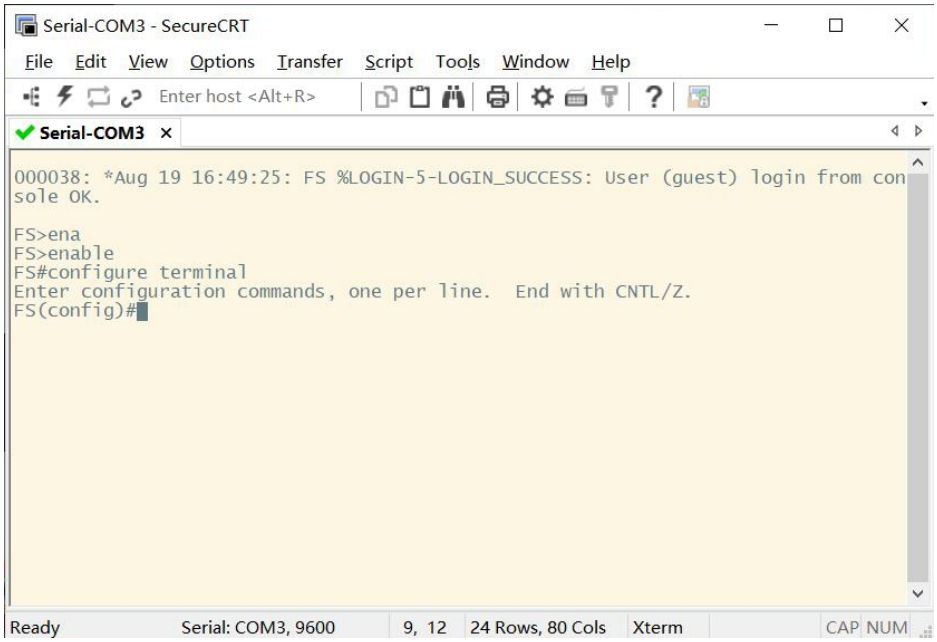
Figure 1-1

Remarks	A is the network device to be managed. PC is a terminal.

Deployment

As shown in Figure 1-2, the user uses the Secure CRT installed on a PC to set up a connection with network device A, and opens the CLI window to enter configuration commands.

Figure 1- 2



1.3 Features

Overview

Feature	Description
Accessing CLI	You can log in to a network device for configuration and management.
Command Modes	The CLI provides several command modes. Commands that can be used vary according to command modes.
System Help	You can obtain the help information of the system during CLI configuration.
Abbreviated Commands	If the entered string is sufficient to identify a unique command, you do not need to enter the full string of the command.
No and Default Options of Commands	You can use the **no** option of a command to disable a function or perform the operation opposite to the command, or use the **default** option of the command to restore default settings.
Prompts Indicating Incorrect Commands	An error prompt will be displayed if an incorrect command is entered.
History Commands	You can use short-cut keys to display or call history commands.
Featured Editing	The system provides short-cut keys for editing commands.
Searching and Filtering of the Show Command Output	You can run the **show** command to search or filter specified commands.
Command Alias	You can configure alias of a command to replace the command.

1.3.1 Accessing CLI

Before using the CLI, you need to connect a terminal or PC to a network device. You can use the CLI after starting the network device and finishing hardware and software initialization. When used for the first time, the network device can be connected only through the console port, which is called out band management. After performing relevant configuration, you can connect and manage the network device through Telnet.

1.3.2 Command Modes

Due to the large number of commands, these commands are classified by function to facilitate the use of commands. The CLI provides several commands modes, and all commands are registered in one or several command modes. You must first enter the command mode of a command before using this command. Different command modes are related with each other while distinguished from each other.

As soon as a new session is set up with the network device management interface, you enter User EXEC mode. In this mode, you can use only a small number of commands and the command functions are limited, such as the **show** commands. Execution results of commands in User EXEC mode are not saved.

To use more commands, you must first enter Privileged EXEC mode. Generally, you must enter a password to enter Privileged EXEC mode. In Privileged EXEC mode, you can use all commands registered in this command mode, and further enter global configuration mode.

Using commands of a certain configuration mode (such as global configuration mode and interface configuration mode) will affect configuration in use. If you save the configuration, these commands will be saved and executed next time the system is restarted. You must enter global configuration mode before entering another configuration mode, such as interface configuration mode.

The following table summarizes the command modes by assuming that the name of the network device is "FS".

Command Mode	Access Method	Prompt	Exit or Entering Another Mode	About
User EXEC (User EXEC mode)	Enter User EXEC mode by default when accessing a network device.	FS>	Run the **exit** command to exit User EXEC mode. Run the **enable** command to enter Privileged EXEC mode.	Use this command mode to conduct basic tests or display system information.
Privileged EXEC (Privileged EXEC mode)	In User EXEC mode, run the **enable** command to enter Privileged EXEC mode.	FS#	Run the **disable** command to return to User EXEC mode. Run the **configure** command to enter global configuration mode.	Use this command mode to check whether the configuration takes effect. This mode is password protected.

Command Mode	Access Method	Prompt	Exit or Entering Another Mode	About
Global configuration (Global configuration mode)	In Privileged EXEC mode, run the **configure** command to enter global configuration mode.	FS(config)#	Run the **exit** or **end** command, or press **Ctrl+C** to return to Privileged EXEC mode. Run the **interface** command to enter interface configuration mode. When using the **interface** command, you must specify the interface. Run the **vlan** *vlan_id* command to enter VLAN configuration mode.	Using commands in this mode will affect the global parameters of the network device.
Interface configuration (Interface configuration mode)	In global configuration mode, run the **interface** command to enter interface configuration mode.	FS(config-if)#	Run the **end** command, or press **Ctrl+C** to return to Privileged EXEC mode. Run the **exit** command to return to global configuration mode. When using the **interface** command, you must specify the interface.	Use this configuration mode to configure various interfaces of the network device.
Config-vlan (VLAN configuration mode)	In global configuration mode, run the **vlan** *vlan_id* command to enter VLAN configuration mode.	FS(config-vlan)#	Run the **end** command, or press **Ctrl+C** to return to the Privileged EXEC mode. Run the **exit** command to return to global configuration mode.	Use this configuration mode to configure VLAN parameters.

1.3.3 System Help

When entering commands in the CLI window, you can obtain the help information using the following methods:

1. At the command prompt in any mode, enter a question mark (?) to list the commands supported by the current command mode and related command description.

For example

```
FS>?
Exec commands:
  <1-99>      Session number to resume
  disable     Turn off privileged commands
  disconnect  Disconnect an existing network connection
  enable      Turn on privileged commands
```

```
  exit          Exit from the EXEC
  help          Description of the interactive help system
  lock          Lock the terminal
  ping          Send echo messages
  show          Show running system information
  telnet        Open a telnet connection
  traceroute    Trace route to destination
```

2. Enter a space and a question mark (?) after a keyword of a command to list the next keyword or variable associated with the keyword.

For example

```
FS(config)#interface ?
  Aggregateport       Aggregate port interface
  Dialer              Dialer interface
  GigabitEthernet     Gigabit Ethernet interface
  Loopback            Loopback interface
  Multilink           Multilink-group interface
  Null                Null interface
  Tunnel              Tunnel interface
  Virtual-ppp         Virtual PPP interface
  Virtual-template    Virtual Template interface
  Vlan                Vlan interface
  range               Interface range command
```

If the keyword is followed by a parameter value, the value range and description of this parameter are displayed as follows:

```
FS(config)#interface vlan ?
  <1-4094>    Vlan port number
```

3. Enter a question mark (?) after an incomplete string of a command keyword to list all command keywords starting with the string.

For example

```
FS#d?
debug   delete   diagnostic   dir   disable   disconnect
```

4. After an incomplete command keyword is entered, if the suffix of this keyword is unique, press the **Tab** key to display the complete keyword.

For example

```
FS# show conf<Tab>
FS# show configuration
```

5. In any command mode, run the **help** command to obtain brief description about the help system.

For example

```
FS(config)#help
Help may be requested at any point in a command by entering
a question mark '?'.    If nothing matches, the help list will
be empty and you must backup until entering a '?' shows the
available options.
Two styles of help are provided:
1. Full help is available when you are ready to enter a
    command argument (e.g. 'show ?') and describes each possible
    argument.
2. Partial help is provided when an abbreviated argument is entered
    and you want to know what arguments match the input
    (e.g. 'show pr?'.)
```

1.3.4 Abbreviated Commands

If a command is long, you can enter a part of the command that is sufficient to identify the command keyword.

For example, to run the **interface** *gigabitEthernet 0/1* command in GigabitEthernet 0/1 interface configuration mode, enter the abbreviated command as follows:

```
FS(config)#int g0/1
FS(config-if-GigabitEthernet 0/1)#
```

1.3.5 No and Default Options of Commands

Most commands have the **no** option. Generally, the **no** option is used to disable a feature or function, or perform the operation opposite to the command. For example, run the **no shutdown** command to perform the operation opposite to the **shutdown** command, that is, enabling the interface. The keyword without the **no** option is used to enable a disabled feature or a feature that is disabled by default.

Most configuration commands have the **default** option. The **default** option is used to restore default settings of the command. Default values of most commands are used to disable related functions. Therefore, the function of the **default** option is the same as that of the **no** option in most cases. For some commands, however, the default values are used to enable related functions. In this case, the function of the **default** option is opposite to that of the **no** option. At this time, the **default** option is used to enable the related function and set the variables to default values.

For specific function of the **no** or **default** option of each command, see the command reference.

1.3.6 Prompts Indicating Incorrect Commands

When you enter an incorrect command, an error prompt is displayed.

The following table lists the common CLI error messages.

Error Message	Meaning	How to Obtain Help
% Ambiguous command: "show c"	The characters entered are insufficient for identifying a unique command.	Re-enter the command, and enter a question mark after the word that is ambiguous. All the possible keywords will be displayed.
% Incomplete command.	The mandatory keyword or variable is not entered in the command.	Re-enter the command, and enter a space and a question mark. All the possible keywords or variables will be displayed.
% Invalid input detected at '^' marker.	An incorrect command is entered. The sign (^) indicates the position of the word that causes the error.	At the current command mode prompt, enter a question mark. All the command keywords allowed in this command mode will be displayed.

1.3.7 History Commands

The system automatically saves commands that are entered recently. You can use short-cut keys to display or call history commands.

The methods are described in the following table.

Operation	Result
Ctrl+P or the UP key	Display the previous command in the history command list. Starting from the latest record, you can repeatedly perform this operation to query earlier records.
Ctrl+N or the DOWN key	After pressing Ctrl+N or the DOWN key, you can return to a command that is recently executed in the history command list. You can repeatedly perform this operation to query recently executed commands.

1.3.8 Featured Editing

When editing the command line, you can use the keys or short-cut keys listed in the following table:

Function	Key or Short-Cut Key	Description
Move the cursor on the editing line.	Left key or Ctrl+B	Move the cursor to the previous character.
	Right key or Ctrl+B	Move the cursor to the next character.
	Ctrl+A	Move the cursor to the head of the command line.
	Ctrl+E	Move the cursor to the end of the command line.
Delete an entered character.	Backspace key	Delete one character to the left of the cursor.
	Delete key	Delete one character to the right of the cursor.
Move the output by one line or one page.	Return key	When displaying contents, press the **Return** key to move the output one line upward and display the next line. This operation is performed when the output does not end yet.
	Space key	When displaying contents, press the **Space** key to page down and display the next page. This operation is performed when the output does not end yet.

When the editing cursor is close to the right boundary, the entire command line will move to the left by 20 characters, and the hidden

front part is replaced by the dollar ($) signs. You can use the related keys or short-cut keys to move the cursor to the characters in the front or return to the head of the command line.

For example, the whole **access-list** may exceed the screen width. When the cursor is close to the end of the command line for the first time, the entire command line moves to the left by 20 characters, and the hidden front part is replaced by the dollar signs ($). Each time the cursor is close to the right boundary, the entire command line moves to the left by 20 characters.

```
access-list 199 permit ip host 192.168.180.220 host
$ost 192.168.180.220 host 202.101.99.12
$0.220 host 202.101.99.12 time-range tr
```

Press **Ctrl+A** to return to the head of the command line. At this time, the hidden tail part of the command line is replaced by the dollar signs ($).

```
access-list 199 permit ip host 192.168.180.220 host 202.101.99.$
```

The default screen width is 80 characters.

1.3.9 Searching and Filtering of the Show Command Output

To search specified contents from the output of the **show** command, run the following command:

Command	Description
show *any-command* | **begin** *regular-expression*	Searches specified contents from the output of the **show** command. The first line containing the contents and all information that follows this line will be output.

The **show** command can be executed in any mode.

Searched contents are case sensitive.

To filter specified contents from the output of the **show** command, run the following commands:

Command	Description
show *any-command* | **exclude** *regular-expression*	Filters the output of the **show** command. Except those containing the specified contents, all lines will be output.
show *any-command* | **include** *regular-expression*	Filters the output of the **show** command. Only the lines containing the specified contents will be output.

To search or filter the output of the **show** command, you must enter a vertical line (|). After the vertical line, select the searching or filtering rules and contents (character or string). Searched and filtered contents are case sensitive.

```
FS#show running-config | include interface
interface GigabitEthernet 0/0
interface GigabitEthernet 0/1
interface GigabitEthernet 0/2
interface GigabitEthernet 0/3
interface GigabitEthernet 0/4
```

```
interface GigabitEthernet 0/5
interface GigabitEthernet 0/6
interface GigabitEthernet 0/7
interface Mgmt 0
```

1.3.10 Command Alias

You can configure any word as the alias of a command to simply the command input.

Configuration Effect

1. Replace a command with a word.

For example, configure "mygateway" as the alias of the **ip route** *0.0.0.0 0.0.0.0192.1.1.1* command. To run this command, you only need to enter "mygateway".

2. Replace the front part of a command with a word, and enter the later part.

For example, configure "ia" as the alias of the **ip address** command. To run this command, you need to enter "ia" and then the specified IP address and subnet mask.

Configuration Steps

Displaying Default Alias

In User EXEC or Privileged EXEC mode, default alias are available for some commands. You can run the **show aliases** command to display these default aliases.

```
FS(config)#show aliases
Exec mode alias:
  h                    help
  p                    ping
  s                    show
  u                    undebug
  un                   undebug
```

These default aliases cannot be deleted.

Configuring a Command Alias

Command	**alias** *mode command-alias original-command*
Parameter Description	*mode*: indicates the command mode of the command represented by the alias. *command-alias*: indicates the command alias. *original-command*: indicates the command represented by the alias.
Command Mode	Global configuration mode
Usage Guide	In global configuration mode, run the **alias** ? command to list all command modes that can be configured with

	aliases.

Displaying Settings of Command Aliases

Run the **show aliases** command to display alias settings in the system.

Notes

- The command replaced by an alias must start from the first character of the command line.
- The command replaced by an alias must be complete.
- The entire alias must be entered when the alias is used; otherwise, the alias cannot be identified.

Configuration Example

Defining an Alias to Replace the Entire Command

Configuration Steps	In global configuration mode, configure the alias "ir" to represent the default route configuration command **ip route** *0.0.0.0 0.0.0.0 192.168.1.1*.
	FS#configure terminal FS(config)#alias config ir ip route 0.0.0.0 0.0.0.0 192.168.1.1
Verification	● Run the **show alias** command to check whether the alias is configured successfully. FS(config)#show alias Exec mode alias: h help p ping s show u undebug un undebug Global configuration mode alias: ir ip route 0.0.0.0 0.0.0.0 192.168.1.1
	● Use the configured alias to run the command, and run the **show running-config** command to check whether the alias is configured successfully.
	FS(config)#ir FS(config)#show running-config Building configuration… ! alias config ir ip route 0.0.0.0 0.0.0.0 192.168.1.1 //Configuring an alias …

	ip route 0.0.0.0 0.0.0.0 192.168.1.1 //Configuration result after the alias "ir" is entered !

Defining an Alias to Replace the Front Part of a Command

Configuration Steps	In global configuration mode, configure the alias "ir" to represent the front part "**ip route**" of the default route configuration command.
	FS#configure terminal FS(config)#alias config ir ip route
Verification	● Run the **show alias** command to check whether the alias is configured successfully. FS(config)#show alias Exec mode alias: h help p ping s show u undebug un undebug Global configuration mode alias: ir ip route
	● Enter the alias "ir" and then the later part of the command "0.0.0.0 0.0.0.0 192.168.1.1". ● Run the **show ap-config running** command to check whether the configuration is successful.
	FS(config)#ir 0.0.0.0 0.0.0.0 192.168.1.1 FS(config)#show running Building configuration… ! alias config ir ip route //Configuring an alias ! ip route 0.0.0.0 0.0.0.0 192.168.1.1 //Configuration result after the alias "ir" and the later part of the command are entered !

System Help

1. The system provides help information for command alias. An asterisk (*) will be displayed in front of an alias. The format is as

follows:

```
*command-alias=original-command
```

For example, in Privileged EXEC mode, the default command alias "s" represents the **show** keyword. If you enter "s?", the keywords starting by "s" and alias information are displayed.

```
FS#s?
*s=show    show    start-chat    start-terminal-service
```

2. If the command represented by an alias contains more than one word, the command is displayed in a pair of quotation marks.

For example, in Privileged EXEC mode, configure the alias "sv" to replace the **show version** command. If you enter "s?", the keywords starting by "s" and alias information are displayed.

```
FS#s?
*s=show    *sv="show version" show    start-chat
start-terminal-service
```

3. You can use the alias to obtain help information about the command represented by the alias.

For example, configure the alias "ia" to represent the **ip address** command in interface configuration mode. If you enter "ia?" in interface configuration mode, the help information on "ip address?" is displayed, and the alias is replaced by the command.

```
FS(config-if)#ia ?
  A.B.C.D    IP address
  dhcp       IP Address via DHCP
FS(config-if)#ip address
```

If you enter a space in front of a command, the command represented by this alias will not be displayed.

2 Configuring Basic Management

2.1 Overview

This document is a getting started guide to network device management. It describes how to manage, monitor, and maintain network devices.

2.2 Applications

Application	Description
Network Device Management	A user logs in to a network device from a terminal and runs commands on a command line interface (CLI) to manage device configurations.

2.2.1 Network Device Management

Scenario

Network device management described in this document is performed through the CLI. A user logs in to Network Device A from a terminal and runs commands on the CLI to manage device configurations. See Figure 2- 1.

Figure 2- 1

2.3 Features

Basic Concepts

TFTP

Trivial File Transfer Protocol (TFTP) is a TCP/IP protocol which allows a client to transfer a file to a server or get a file from a server.

AAA

AAA is short for Authentication, Authorization and Accounting.

Authentication refers to the verification of user identities and the related network services.

Authorization refers to the granting of network services to users according to authentication results.

Accounting refers to the tracking of network service consumption by users. A billing system charges users based on consumption records.

AAA provides effective means of network management and security protection.

RADIUS

Remote Authentication Dial In User Service (RADIUS) is the most widely used AAA protocol at present.

Telnet

Telnet is a terminal emulation protocol in the TCP/IP protocol stack which provides access to a remote host through a virtual terminal connection. It is a standard protocol located at Layer 7 (application layer) of the Open System Interconnection (OSI) model and used on the internet for remote login. Telnet sets up a connection between the local PC and a remote host.

System Information

System information includes the system description, power-on time, hardware and software versions, control-layer software version, and boot-layer software version.

Hardware Information

Hardware information includes the physical device information as well as slot and module information. The device information includes the device description and slot quantity. The slot information includes the slot ID, module description (which is empty if a slot does not have a module), and actual and maximum number of physical ports.

Overview

Feature	Description
User Access Control	Controls the terminal access to network devices on the internet based on passwords and privileges.
Login Authentication Control	Performs username-password authentication to grant access to network devices when AAA is enabled. (Authentication is performed by a dedicated server.)
Basic System Parameters	Refer to the parameters of a system, such as the clock, banner, and Console baud rate.
Displaying Configurations	Displays the system configurations, including the configurations that the system is currently running and the device configurations stored in the nonvolatile random access memory (NVRAM).
Multiple-configuration Booting	Allows users to modify the path for saving startup configurations of the device and the corresponding file name.
Telnet	Telnet is an application-layer protocol in the TCP/IP protocol stack. It provides the standard governing remote login and virtual terminal communication on the internet.
Restart	Introduces system restart.
Running Batch File Commands	Runs the commands in batches.

2.3.1 User Access Control

User access control refers to the control of terminal access to network devices on the internet based on passwords and privileges.

Working Principle

Privilege Level

16 privilege levels are defined ranging from 0 to 15 for CLI on network devices to grant users access to different commands. Level 0 is the lowest level granting access to just a few commands, whereas level 15 is the highest level granting access to all commands. Levels 0 and

1 are common user levels without the device configuration permission (users are not allowed to enter global configuration mode by default). Levels 2–15 are privileged user levels with the device configuration permission.

Password Classification

Passwords are classified into two types: password and security. The first type refers to simple encrypted passwords at level 15. The second type refers to secure encrypted passwords at levels 0–15. If a level is configured with both simple and secure encrypted passwords, the simple encrypted password will not take effect. If you configure a non-15 level simple encrypted password, a warning is displayed and the password is automatically converted into a secure encrypted password. If you configure the same simple encrypted password and secure encrypted password at level 15, a warning is displayed.

Password Protection

Each privilege level on a network device has a password. An increase in privilege level requires the input of the target level password, whereas a reduction in privilege level does not require password input.

By default, only two privilege levels are password-protected, namely, level 1 (common user level) and level 15 (privileged user level). Sixteen privilege levels with password protection can be assigned to the commands in each mode to grant access to different commands.

If no password is configured for a privileged user level, access to this level does not require password input. It is recommended that a password be configured for security purposes.

Command Authorization

Each command has its lowest execution level. A user with a privilege level lower than this level is not allowed to run the command. After the command is assigned a privilege level, users at this level and higher have access to the command.

Related Configuration

Configuring a Simple Encrypted Password

- Run the **enable password** command.

Configuring a Secure Encrypted Password

- Run the **enable secret** command.
- A secure encrypted password is used to control the switching between user levels. It has the same function as a simple encrypted password but uses an enhanced password encryption algorithm. Therefore, secure encrypted passwords are recommended out of security consideration.

Configuring Command Privilege Levels

- Run the **privilege** command to assign a privilege level to a command.
- A command at a lower level is accessible by more users than a command at a higher level.

Raising/Lowering a User Privilege Level

- Run the **enable** command or the **disable** command to raise or lower a user privilege level respectively.
- After logging in to a network device, the user can change his/her level to obtain access to commands at different privilege levels.
- To enable level increase logging, run the **login privilege log** command.

Enabling Line Password Protection

- Line password protection is required for remote login (such as login through Telnet).
- Run the **password**[**0** | **7**] *line* command to configure a line password, and then run the **login** command to enable password protection.
- By default, terminals do not support the **lock** command.

2.3.2 Login Authentication Control

In login authentication with AAA disabled, the password entered by a user is checked against the configured line password. If they are consistent, the user can access the network device. In local authentication, the username and password entered by a user are checked against those stored in the local user database. If they are matched, the user can access the network device with proper management permissions.

In AAA, the username and password entered by a user are authenticated by a server. If authentication is successful, the user can access the network device and enjoy certain management permissions.

For example, a RADIUS server can be used to authenticate usernames and passwords and control users' management permissions on network devices. Network devices no longer store users' passwords, but send encrypted user information to the RADIUS server, including usernames, passwords, shared passwords, and access policies. This provides a convenient way to manage and control user access and improve user information security.

Working Principle

Line Password

If AAA is disabled, you can configure a line password used to verify user identities during login. After AAA is enabled, line password verification does not take effect.

Local Authentication

If AAA is disabled, you can configure local authentication to verify user identities and control management permissions by using the local user database. After AAA is enabled, local authentication does not take effect.

AAA

AAA provides three independent security functions, namely, Authentication, Authorization and Accounting. A server (or the local user database) is used to perform authentication based on the configured login authentication method list and control users' management permissions. For details about AAA, see *Configuring AAA*.

Related Configuration

Configuring Local User Information

- Run the **username** command to configure the account used for local identity authentication and authorization, including usernames, passwords, and optional authorization information.

Configuring Local Authentication for Line-Based Login

- Run the **login local** command (in the case that AAA is disabled).
- Perform this configuration on every device.

Configuring AAA Authentication for Line-Based Login

- The default authentication method is used after AAA is enabled.
- Run the **login authentication** command to configure a login authentication method list for a line.
- Perform this configuration when the local AAA authentication is required.

Configuring Non-AAA Authentication for Line-Based Login When AAA Is Enabled

- Run the **login access non-aaa** command in global configuration mode.
- Perform this configuration on every device.

Configuring the Connection Timeout Time

- The default connection timeout time is 10 minutes.
- Run the **exec-timeout** command to change the default connection timeout time. An established connection will be closed if no output is detected during the timeout time.
- Perform this configuration when you need to increase or reduce the connection timeout time.

Configuring the Session Timeout Time

- The default session timeout time is 0 minutes, indicating no timeout.
- Run the **session-timeout** command to change the default session timeout time.
- The session established to a remote host through a line will be disconnected if no output is detected during the timeout time. Then the remote host is restored to Idle. Perform this configuration when you need to increase or reduce the session timeout time.

Locking a Session

- By default, terminals do not support the **lock** command.
- Run the **lockable** command to lock the terminals connected to the current line.
- To lock a session, first enable terminal lock in line configuration mode, and then run the **lock** command in terminal EXEC mode to lock the terminal.

2.3.3 Basic System Parameters

System Time

The network device system clock records the time of events on the device. For example, the time shown in system logs is obtained from the system clock. Time is recorded in the format of *year-month-day, hour:minute:second, day of the week*.

When you use a network device for the first time, set its system clock to the current date and time manually.

Configuring a System Name and Command Prompt

You can configure a system name to identify a network device. The default system name is **FS**. A name with more than 32 characters will be truncated to keep only the first 32 characters. The command prompt keeps consistent with the system name.

Banner

A banner is used to display login prompt information. There are two types of banner: Daily notification and login banner.

- Daily notification is displayed on all terminals connected to network devices soon after login. Urgent messages (such as immediate system shutdown) can be delivered to users through daily notification.
- A login banner appears after daily notification to display login information.

Configuring the Console Baud Rate

You can manage network device through a Console port The first configuration on the network device must be performed through the Console port. The serial port baud rate can be changed based on actual requirements. Note that the management terminal must have consistent baud rate setting with the device console.

Configuring the Connection Timeout Time

The connection timeout time is used to control device connections (including established connections and sessions established to remote hosts). A connection will be closed when no input is detected during the timeout time.

Related Configuration

Configuring the System Date and Clock

- Run the **clock set** command to configure the system time of a network device manually. The device clock starts from the configured time and keeps running even when the device is powered off.

Updating the Hardware Clock

- If the hardware clock and software clock are not synchronized, run the **clock update-calendar** command to copy the date and time of the software clock to the hardware clock.

Configuring a System Name

- Run the **hostname** command to change the default system name.
- The default host name is **FS**.

Configuring a Command Prompt

- Run the **prompt** command.

Configuring Daily Notification

- By default, no daily notification is configured.
- Run the **banner motd** command to configure daily notification.
- Daily notification is displayed on all terminals connected to network devices soon after login. Urgent messages (such as immediate system shutdown) can be delivered to users through daily notification.

Configuring a Login Banner

- By default, no login banner is configured.
- Run the **banner login** command to configure a login banner to display login information.

Configuring the Console Baud Rate

- Run the **speed** command.
- The default baud rate is 9,600 bps.

2.3.4 Displaying Configurations

Displays the system configurations, including the configurations that the system is currently running and the device configurations stored in the NVRAM.

Working Principle

Running Configurations

Running configurations, namely, running-config, are the configurations that individual component modules run in real time. A request can be made to all running components to collect configurations, which will be orchestrated before being displayed to users. Only running components may provide real-time configurations, whereas unloaded components do not display configurations. In the case that the system is started, a component process is restarted, the configurations collected during this period may be inaccurate due to the component unstable state. For example, the configurations of a component may not be missing initially but can be displayed later.

Startup Configurations

The configurations stored in the NVRAM, namely, startup-config, are the configurations executed during device startup. When the system is restarted, startup-config is loaded to become new running-config. To display permanent configurations, the system needs to read the **startup-config** file in the NVRAM.

The **startup-config** file copied to the device only supports the UTF-8 (no BOM) format.

Related Configuration

Displaying Running Configurations

Run the **show running-config** [**interface** *interface*] command to display the configurations that the system is currently running or the configurations on an interface.

Displaying Startup Configurations

Run the **show startup-config** command.

Storing Startup Configurations

Run the **write** or **copy running-config startup-config** command to store the current running configurations as new startup configurations.

2.3.5 Multiple-configuration Booting

Multiple-configuration booting allows users to modify the path for saving startup configurations of the device and the corresponding file name. At present, configurations can be saved to an extended flash memory and an extended USB flash drive of a device. To save configurations in an extended USB flash drive, the device must support at least one USB interface. If the device supports two or more USB interfaces, startup configurations are saved in **/mnt/usb0**.

Working Principle

- By default, the startup configuration file of a device is saved in **Flash:/config.text** and named **config.text**. Use this command to modify the path for saving startup configurations of the device and the corresponding file name.

The startup configuration file name follows a slash "/", for example, **Flash:/fs.text** and **Usb0:/fs.text**.

The startup configuration file name consists of a path and a file name. The path is mandatory. Otherwise, configurations cannot be saved by using the **write** command. Take **Flash:/fs/fs.text** and **Usb0:/fs/fs.text** as examples, where the **Flash:/fs** and **Usb0:/fs** folders must exist. In master-slave mode, all device paths are required.

To save the startup configuration file to a USB flash drive, the device must provide a USB interface with a USB flash drive inserted. Otherwise, configurations cannot be saved by using the **write** command. In master-slave mode, all devices must have USB flash drives connected.

Related Configuration

Modifying the Path for Saving Startup Configurations and the Corresponding File Name

Run the **boot config** { **flash:***filename* | **usb0:***filename* } command to modify the path for saving startup configurations and the corresponding file name.

Displaying the Path for Saving Startup Configurations and the Corresponding File Name

Run the **show boot config** command to display the path for saving startup configurations and the corresponding file name.

2.3.6 Telnet

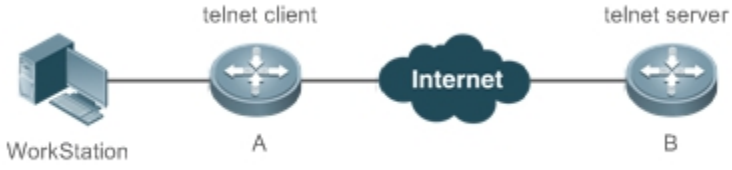
Working Principle

Telnet is an application-layer protocol in the TCP/IP protocol stack. It provides the standard governing remote login and virtual terminal communication on the internet.

The Telnet Client service allows a local or remote user who has logged in to a network device to use its Telnet Client program to access other remote system resources on the internet. In Figure 2-2, a user with a PC connects to Network Device A by using the terminal emulation or Telnet program and then logs in to Network Device B by using the **telnet** command to perform configuration management.

FS Telnet program supports the use of IPv4 and IPv6 addresses. A Telnet server accepts Telnet connection requests that carry IPv4 and IPv6 addresses. A Telnet client can send connection requests to hosts identified by IPv4 and IPv6 addresses.

Figure 2-2



Related Configuration

Enabling the Telnet Client Service

- Run the **telnet** command to log in to a remote device.

Restoring a Telnet Client Session

- Run the **<1-99>** command.

Disconnecting a Suspended Telnet Client Session

- Run the **disconnect** *session-id* command.

Enabling the Telnet Server Service

- Run the **enable service telnet-server** command.
- Perform this configuration when you need to enable Telnet login.

2.3.7 Restart

The timed restart feature makes user operation easier in some scenarios (such as tests).

- If you configure a time interval, the system will restart after the interval. The interval is in the format of *mmm* or *hhh:mm*, in the unit of minutes. You can specify the interval name to reflect the restart purpose.
- If you define a future time, the system will restart when the time is reached.

The clock feature must be supported by the system if you want to use the **at** option. It is recommended that you configure the system clock in advance. A new restart plan will overwrite the existing one. A restart plan will be invalid if the system is restarted before the plan takes effect.

The span between the restart time and current time must not exceed 31 days, and the restart time must be later than the current system time. After you configure a restart plan, do not to change the system clock; otherwise, the plan may fail (for example, the system time is changed to a time after the restart time.)

Related Configuration

Configuring Restart

- Run the **reload** command to configure a restart policy.
- Perform this configuration when you need to restart a device at a specific time.

2.3.8 Running Batch File Commands

In system management, sometimes it takes a long time to enter many commands on the CLI to manage a function. This process is prone to errors and omissions. You can put the commands to a batch file according to configuration steps and execute the file to complete related configuration.

You can specify the name and content of the batch file on your PC and transfer the file to the device flash memory through TFTP. The batch processing content simulates user input. Therefore, you need to edit the batch file content according to the CLI command configuration sequence. In addition, you need to write the responses to interactive commands to the batch file to ensure normal command execution.

The batch file size must not exceed 128 KB; otherwise, it will fail to be executed. You can divide a large batch file into multiple parts not larger than 128 KB each.

Related Configuration

Batch-Running Commands

- Run **execute** to run the commands in batches.
- This command provides a convenient way to run multiple commands at a time.

2.4 Configuration

Configuring Passwords and Privileges	(Optional) It is used to configure passwords and command privilege levels.	
	enable password	Configures a simple encrypted password.
	enable secret	Configures a secure encrypted password.
	enable	Raises a user privilege level.
	login privilege log	Outputs log information of user privilege level increase.
	disable	Lowers a user privilege level.
	privilege	Configures command privilege levels.
	password	Specifies a line password.
	login	Enables line password protection.
Configuring Login and Authentication	(Optional) It is used to configure different login modes and authentication methods.	
	username	Configures local user account information and optional authorization information.
	login local	Configures local authentication for line-based login.
	login access non-aaa	Configures non-AAA authentication for line-based login when AAA is enabled.
	login authentication	Configures AAA authentication for line-based login.
	telnet	Enables the Telnet Client service.
	enable service telnet-server	Enables the Telnet Server service.
	exec-timeout	Configures the connection timeout time.
	session-timeout	Configures the session timeout time.
	lockable	Enables line-based terminal lock.
	lock	Locks a terminal connected to the current line.
Configuring Basic System Parameters	(Optional) It is used to configure basic system parameters.	
	clock set	Configures the system date and clock.
	clock update-calendar	Updates the hardware clock.
	hostname	Configures a system name.
	prompt	Configures a command prompt.
	banner motd	Configures daily notification.

	bannerlogin	Configures a login banner.
	speed	Configures the Console baud rate.
Enabling and Disabling a Specific Service	(Optional) It is used to enable and disable a specific service.	
	enable service	Enables a service.
Configuring Multiple-configuration Booting	(Optional) It is used to modify the startup configuration file.	
	boot config { **flash:***filename* | **usb0:***filename* }	Modifies the path for saving startup configurations and the corresponding file name.
Configuring a Restart Policy	(Optional) It is used to configure a system restart policy.	
	reload	Restarts a device.
Running Batch File Commands	(Optional) It is used to run the commands in batches.	
	execute { [**flash:**] *filename* }	Runs the commands in batches.

2.4.1 Configuring Passwords and Privileges

Configuration Effect

- Configure passwords to control users' access to network devices.
- Assign a privilege level to a command to grant the command access to only the users at or higher than the level.
- Lower the command privilege level to grant more users access to the command.
- Raise the command privilege level to limit the command access to a few users.

Notes

- You can use the password configuration command with the **level** option to configure a password for a specific privilege level. After you specify the level and the password, the password works for the users who need to access this level.
- By default, no password is configured for any level. The default level is 15.
- If you configure a simple encrypted password with a non-15 level, a warning is displayed and the password is automatically converted into a secure encrypted password.
- The system chooses the secure encrypted password over the simple encrypted password if both of them are configured.

Configuration Steps

Configuring a Simple Encrypted Password

- (Optional) Perform this configuration when you need to establish simple encrypted password verification when users switch between different privilege levels.
- Run the **enable password** command to configure a simple encrypted password.

Configuring a Secure Encrypted Password

- (Optional) Perform this configuration when you need to establish secure encrypted password verification when users switch between different privilege levels.
- Run the **enable secret** command to configure a secure encrypted password.
- A secure encrypted password has the same function as a simple encrypted password but uses an enhanced password encryption algorithm. Therefore, secure encrypted passwords are recommended out of security consideration.

Configuring Command Privilege Levels

- Optional.
- A command at a lower level is accessible by more users than a command at a higher level.

Raising/Lowering a User Privilege Level

- After logging in to a network device, the user can change his/her level to obtain access to commands at different privilege levels.
- Run the **enable** command or the **disable** command to raise or lower a user privilege level respectively.
- To enable level increase logging, run the **login privilege log** command.

Enabling Line Password Protection

- (Optional) Line password protection is required for remote login (such as login through Telnet).
- Run the **password** [**0** | **7**] *line* command to configure a line password, and then run the **login** command to enable login authentication.

If a line password is configured but login authentication is not configured, the system does not display password prompt.

Verification

- Run the **show privilege** command to display the current user level.
- Run the **show running-config** command to display the configuration.

Related Commands

Configuring a Simple Encrypted Password

Command	**enable password** [**level** *level*] { *password* | [**0** | **7**] *encrypted-password* }
Parameter Description	*level*: Indicates a specific user level. *password:* Indicates the password used to enter privileged EXEC mode. *0:* Indicates that the password is entered in plaintext. *7:* Indicates that the password is entered in cyphertext. *encrypted-password:* Indicates the password text, which must contain case-sensitive English letters and digits. Leading spaces are allowed, but will be ignored. However, intermediate and trailing spaces are recognized.
Command Mode	Global configuration mode
Usage Guide	Currently, simple encrypted passwords can be configured with only level 15 and take effect only when no secure encrypted password is configured.

	If you configure a simple encrypted password with a non-15 level, a warning is displayed and the password is automatically converted into a secure encrypted password. If the level 15 simple encrypted password and secure encrypted password are configured the same, a warning is displayed. If you specify an encryption type and enter a password in plaintext, you cannot re-enter privileged EXEC mode. An encrypted password cannot be retrieved once lost. You have to configure a new password.

Configuring a Secure Encrypted Password

Command	**enable secret** [**level** *level*] {*secret* | [**0 | 5**] *encrypted-secret* }
Parameter Description	*level*: Indicates a specific user level. secret: Indicates the password used to enter privileged EXEC mode. **0 | 5**: Indicates the password encryption type. 0 indicates no encryption, and 5 indicates secure encryption. *encrypted-password*: Indicates the password text.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure passwords for different privilege levels.

Raising a User Privilege Level

Command	**enable** [*privilege-level*]
Parameter Description	*privilege-level*: Indicates a specific privilege level.
Command Mode	Privileged EXEC mode
Usage Guide	An increase in privilege level requires the input of the target level password.

Lowering a User Privilege Level

Command	**disable** [*privilege-level*]
Parameter Description	*privilege-level*: Indicates a specific privilege level.
Command Mode	Privileged EXEC mode
Usage Guide	A reduction in privilege level does not require password input. Use this command to exit Privileged EXEC mode and return to user EXEC mode. If *privilege-level* is specified, the current privilege level is reduced to the specified level. *privilege-level* must be lower than the current level.

Enabling Level Increase Logging

Command	**login privilege log**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command to enable logging of privilege level increase. The configuration takes effect for all terminals.

Configuring Command Privilege Levels

Command	**privilege** *mode* [**all**] { **level** *level* | **reset** } *command-string*
Parameter Description	*mode*: Indicates the CLI mode of the command. For example, **config** indicates the global configuration mode, **EXEC** indicates the privileged command mode, and **interface** indicates the interface configuration mode. **all**: Changes the subcommand privilege levels of a specific command to the same level. **level** *level*: Indicates a privilege level, ranging from 0 to 15. **reset**: Restores the command privilege level to the default. *command-string*: Indicates the command to be assigned a privilege level.
Command Mode	Global configuration mode
Usage Guide	To restore a command privilege level, run the **no privilege** *mode* [**all**] **level** *level command* command in global configuration mode.

Specifying a Line Password

Command	**Password** [**0** | **7**] *line*
Parameter Description	**0**: Indicates to configure a password in plaintext. **7**: Indicates to configure a password in cyphertext. line: Indicates the password string.
Command Mode	Line configuration mode
Usage Guide	N/A

Enabling Line Password Protection

Command	**login**
Parameter Description	N/A
Command Mode	Line configuration mode
Usage Guide	N/A

Configuration Example

Configuring Command Authorization

Scenario	Assign privilege level 1 to the **reload** command and its subcommands and configure level 1 as the valid level (by configuring the **test** password).
Configuration Steps	● Assign privilege level 1 to the **reload** command and its subcommands.
	FS# configure terminal FS(config)# privilege exec all level 1 reload FS(config)# enable secret level 1 0 test FS(config)# end
Verification	● Check whether the **reload** command and its subcommands are accessible at level 1.
	FS# disable 1 FS> reload ? at reload at<cr>

2.4.2 Configuring Login and Authentication

Configuration Effect

- Establish line-based login identity authentication.
- Run the **telnet** command on a network device to log in to a remote device.
- Close an established connection if no output is detected during the timeout time.
- Disconnect an established session connecting to a remote host and restore the host to Idle if no output is detected during the timeout time.
- Lock a terminal to deny access. When a user enters any character on the locked terminal, the password prompt is displayed. The terminal will be automatically unlocked if the entered password is correct.

Configuration Steps

Configuring Local User Information

- Mandatory.
- Run the **username** command to configure the account used for local identity authentication and authorization, including usernames, passwords, and optional authorization information.
- Perform this configuration on every device.

Configuring Local Authentication for Line-Based Login

- Mandatory.
- Configure local authentication for line-based login in the case that AAA is disabled.
- Perform this configuration on every device.

Configuring AAA Authentication for Line-Based Login

- (Optional) Perform this configuration to configure AAA authentication for line-based login.
- Configure AAA authentication for line-based login in the case that AAA is enabled.
- Perform this configuration on every device.

Configuring Non-AAA Authentication for Line-Based Login When AAA Is Enabled

- Optional.
- Run the **login access non-aaa** command in global configuration mode to authenticate line-based login in non-AAA mode in the case that AAA is enabled.
- Perform this configuration on every device.

Enabling the Telnet Client Service

- Run the **telnet** command to log in to a remote device.

Restoring a Telnet Client Connection

- (Optional) Perform this configuration to restore the connection on a Telnet client.

Closing a Suspended Telnet Client Connection

- (Optional) Perform this configuration to close the suspended connection on a Telnet client.

Enabling the Telnet Server Service

- Optional.
- Enable the Telnet Server service when you need to enable Telnet login.

Configuring the Connection Timeout Time

- Optional.
- An established connection will be closed if no output is detected during the timeout time.
- Perform this configuration when you need to increase or reduce the connection timeout time.

Configuring the Session Timeout Time

- Optional.
- The session connecting to a remote host will be disconnected and the host be restored to Idle if no output is detected during the timeout time.
- Perform this configuration when you need to increase or reduce the session timeout time.

Locking a Session

- (Optional) Perform this configuration when you need to temporarily exit a session on a device.

- To lock a session, first enable terminal lock in line configuration mode, and then run the **lock** command to lock the terminal.

Verification

- Run the **show running-config** command to display the configuration.
- In the case that AAA is disabled, after local user information and line-based local authentication are configured, check whether users are prompted for username and password input for access to the CLI.
- In the case that AAA is enabled, after local user information and local AAA authentication are configured, check whether users are prompted for username and password input for access to the CLI.
- Run the **show user** command to display the information about the users who have logged in to the CLI.
- Telnet clients can connect to devices enabled with the Telnet Server service.
- When a user presses **Enter** on a locked CLI, the user is prompted for password input. The session is unlocked only when the entered password is the same as the configured one.
- Run the **show sessions** command to display every established Telnet client instance.

Related Commands

Configuring Local User Information

Command	**username** *name* [**login mode** { **aux** | **console** | **ssh** | **telnet** }] [**online amount** *number*] [**permission** *oper-mode path*] [**privilege** *privilege-level*] [**reject remote-login**] [**web-auth**] [**pwd-modify**] [**nopassword** | **password** [**0** | **7**] *text-string* | **secret** [**0** | **5**] *text-string*
Parameter Description	*name*: Indicates a user name. **login mode**: Indicates the login mode. **aux**: Sets the login mode to AUX. **console**: Sets the login mode to Console. **ssh**: Sets the login mode to SSH. **telnet**: Sets the login mode to Telnet. **online amount** *number*: Indicates the maximum number of online accounts. **permission** *oper-mode path*: Configures the file operation permission. *op-mode* indicates the operation mode, and *path* indicates the directory or path of a specific file. **privilege** *privilege-level*: Indicates the account privilege level, ranging from 0 to 15. **reject remote-login**: Rejects remote login by using the account. **web-auth**: Allows only Web authentication for the account. **pwd-modify**: Allows the account owner to change the password. This option is available only when **web-auth** is configured. **nopassword**: Indicates that no password is configured for the account. **password** [**0** | **7**] *text-string*: Indicates the password configured for the account. 0 indicates that the password is input in plaintext, and 7 indicates that the password is input in cyphertext. The default is plaintext. **secret** [**0** | **5**] *text-string*: Indicates the password configured for the account. 0 indicates that the password is input in plaintext, and 5 indicates that the password is input in cyphertext. The default is plaintext.
Command	Global configuration mode

Mode	
Usage Guide	Use this command to create a local user database to be used by authentication. If the value 7 is selected for the encryption type, the entered cyphertext string must consist of an even number of characters. This setting is applicable to the scenario where encrypted passwords may be copied and pasted. In other cases, the value 7 is not selected.

Configuring Local Authentication for Line-Based Login

Command	**login local**
Parameter Description	N/A
Command Mode	Line configuration mode
Usage Guide	Use this command to configure local authentication for line-based login in the case that AAA is disabled. Local user information is configured by using the **username** command.

Configuring AAA Authentication for Line-Based Login

Command	**login authentication** { **default** | *list-name* }
Parameter Description	**default:** Indicates the default authentication method list name. *list-name*: Indicates the optional method list name.
Command Mode	Line configuration mode
Usage Guide	Use this command to configure AAA authentication for line-based login in the case that AAA is enabled. The AAA authentication methods, including RADIUS authentication, local authentication, and no authentication, are used during the authentication process.

Configuring Non-AAA Authentication for Line-Based Login When AAA Is Enabled

Command	**login access non-aaa**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command when you need to perform non-AAA authentication on line-based login in the case that AAA is enabled. The configuration takes effect for all terminals.

Enabling the Telnet Client Service

Command	**telnet** [**oob**] *host* [*port*] [**/source** { **ip** *A.B.C.D* **| ipv6** *X:X:X:X::X* **| interface** *interface-name* }] [**/vrf** *vrf-name*]
Parameter	**oob:** Remotely connects to a Telnet server through out-of-band communication (by using a management port). This

Description	option is available only when the device has a management port. *host*: Indicates the IPv4 address, IPv6 address, or host name of the Telnet server. *port*: Indicates the TCP port number of the Telnet server. The default value is 23. **/source:** Indicates the source IP address or source port used by a Telnet client. **ip** *A.B.C.D:* Indicates the source IPv4 address used by the Telnet client. **ipv6** *X:X:X:X::X:* Indicates the source IPv6 address used by the Telnet client. **interface** *interface-name*: Indicates the source port used by the Telnet client. **/vrf** *vrf-name*: Indicates the name of the virtual routing and forwarding (VRF) table to be queried.
Command Mode	Privileged EXEC mode
Usage Guide	A user can telnet to a remote device identified by an IPv4 host name, IPv6 host name, IPv4 address, or IPv6 address.

Restoring a Telnet Client Session

Command	<1-99>
Parameter Description	N/A
Command Mode	User EXEC mode
Usage Guide	Use this command to restore a Telnet client session. A user can press the shortcut key **Ctrl**+**Shift**+**6 X** to temporarily exit the Telnet client session that is established using the **telnet** command, run the **<1-99>** command to restore the session, and run the **show sessions** command to display the session information.

Closing a Suspended Telnet Client Connection

Command	**disconnect** *session-id*
Parameter Description	*session-id*: Indicates the suspended Telnet client session ID.
Command Mode	User EXEC mode
Usage Guide	Use this command to close a specific Telnet client session by entering the session ID.

Enabling the Telnet Server Service

Command	**enable service telnet-server**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command to enable the Telnet Server service. The IPv4 and IPv6 services are also enabled after the command is executed.

Configuring the Connection Timeout Time

Command	**exec-timeout** *minutes* [*seconds*]
Parameter Description	*minutes*: Indicates the connection timeout time in the unit of minutes. *seconds*: Indicates the connection timeout time in the unit of seconds.
Command Mode	Line configuration mode
Usage Guide	Use this command to configure the timeout time for the established connections on a line. A connection will be closed when no input is detected during the timeout time. To remove the connection timeout configuration, run the **no exec-timeout** command in line configuration mode.

Configuring the Session Timeout Time

Command	**session-timeout** *minutes*[**output**]
Parameter Description	*minutes*: Indicates the session timeout time in the unit of minutes. **output**: Indicates whether to add data output as a timeout criterion.
Command Mode	Line configuration mode
Usage Guide	Use this command to configure the timeout time for the remote host sessions on a line. A session will be disconnected when no input is detected during the timeout time. To cancel the session timeout time, run the **no session-timeout** command in line configuration mode.

Enabling Line-Based Terminal Lock

Command	**lockable**
Parameter Description	N/A
Command Mode	Line configuration mode
Usage Guide	N/A

Locking a Terminal Connected to the Current Line

Command	**lock**
Parameter Description	N/A
Command Mode	Line configuration mode
Usage Guide	N/A

Configuration Example

Establishing a Telnet Session to a Remote Network Device

Configuration Steps	● Establish a Telnet session to a remote network device with the IP address 192.168.65.119. ● Establish a Telnet session to a remote network device with the IPv6 address 2AAA:BBBB::CCCC. ● Run the **telnet** command in privileged EXEC mode, and run the **do telnet** command in privileged EXEC mode/configuration mode/interface configuration mode.
	FS# telnet 192.168.65.119 Trying 192.168.65.119 ... Open User Access Verification Password:
	FS# telnet 2AAA:BBBB::CCCC Trying 2AAA:BBBB::CCCC ... Open User Access Verification Password:
Verification	● Check whether the Telnet sessions are established to the remote network devices.

Configuring the Connection Timeout Time

Configuration Steps	● Set the connection timeout time to 20 minutes.
	FS# configure terminal//Enter global configuration mode. FS# line vty 0 //Enter line configuration mode. FS(config-line)#exec-timeout 20 //Set the connection timeout time to 20 minutes.
Verification	● Check whether the connection between a terminal and the local device is closed when no input is detected during the timeout time.

Configuring the Session Timeout Time

Configuration Steps	● Set the session timeout time to 20 minutes.

	FS# configure terminal//Enter global configuration mode.
	FS(config)# line vty 0 //Enter line configuration mode.
	FS(config-line)#session-timeout 20//Set the session timeout time to 20 minutes.

Verification	● Check whether the session between a terminal and the local device is disconnected when no input is detected during the timeout time.

2.4.3 Configuring Basic System Parameters

Configuration Effect

- Configure basic system parameters.

Configuration Steps

Configuring the System Date and Clock

- Mandatory.
- Configure the system time of a network device manually. The device clock starts from the configured time and keeps running even when the device is powered off.

⚠ The time configuration is applied only to the software clock if the network device does not provide a hardware clock. The configuration will be invalid when the device is powered off.

Updating the Hardware Clock

- Optional.
- Perform this configuration when you need to copy the date and time of the software clock to the hardware clock so that the hardware clock is synchronized with the software clock.

Configuring a System Name

- (Optional) Perform this configuration to change the default system name.

Configuring a Command Prompt

- (Optional) Perform this configuration to change the default command prompt.

Configuring Daily Notification

- (Optional) Perform this configuration when you need to display important prompts or warnings to users.
- You can configure notification in one or multiple lines, which will be displayed to users after login.

Configuring a Login Banner

- (Optional) Perform this configuration when you need to display important messages to users upon login or logout.

Configuring the Console Baud Rate

- (Optional) Perform this configuration to change the default Console baud rate.

Verification

- Run the **show clock** command to display the system time.
- Check whether a login banner is displayed after login.

- Run the **show version** command to display the system information and version.

Related Commands

Configuring the System Date and Clock

Command	**clock set** *hh:mm:ss month day year*
Parameter Description	*hh*:*mm*:*ss*: Indicates the current time, in the format of *hour* (24-hour format):*minute*:*second*. *day*: Indicates a day (1–31) of the month. *month*: Indicates a month (from January to December) of the year. *year*: Indicates a year, ranging from 1993 to 2035. Abbreviation is not supported.
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to configure the system time. If the device does not provide a hardware clock, the time configuration will be invalid when the device is powered off.

Updating the Hardware Clock

Command	**clock update-calendar**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	After the configuration, the time of the software clock will overwrite that of the hardware clock.

Configuring a System Name

Command	**hostname** *name*
Parameter Description	*name*: Indicates the system name, which must consist of printable characters and must not exceed 63 bytes.
Command Mode	Global configuration mode
Usage Guide	To restore the system name to the default, run the **no hostname** command in global configuration mode.

Configuring a Command Prompt

Command	**prompt** *string*
Parameter Description	*string*: Indicates the command prompt name. A name with more than 32 characters will be truncated to keep only the first 32 characters.
Command Mode	Privileged EXEC mode
Usage Guide	To restore the command prompt to the default settings, run the **no prompt** command in global configuration mode.

Configuring Daily Notification

Command	**banner motd** *c message c*
Parameter Description	*c*: Indicates a delimiter, which can be any character, such as "&".
Command Mode	Global configuration mode
Usage Guide	A message must start and end with delimiter+carriage return respectively. Any characters following the ending delimiter will be dropped. Any letter contained in the message must not be used as the delimiter. The message must not exceed 255 bytes.

Configuring a Login Banner

Command	**banner login** *c message c*
Parameter Description	*c*: Indicates a delimiter, which can be any character, such as "&".
Command Mode	Global configuration mode
Usage Guide	A message must start and end with delimiter+carriage return respectively. Any characters following the ending delimiter will be dropped. Any letter contained in the message must not be used as the delimiter. The message must not exceed 255 bytes. To remove the login banner configuration, run the **no banner login** command in global configuration mode.

Configuring the Console Baud Rate

Command	**speed** *speed*
Parameter Description	*speed*: Indicates the console baud rate, in the unit of bps. The serial port baud rate can be set to 9,600 bps, 19,200 bps, 38,400 bps, 57,600 bps, or 115,200 bps. The default is 9,600 bps.
Command Mode	Line configuration mode
Usage Guide	You can configure the asynchronous line baud rate based on requirements. The **speed** command is used to configure receive and transmit rates for the asynchronous line.

Configuration Example

Configuring the System Time

Configuration Steps	● Change the system time to 2003-6-20, 10:10:12.
	FS# clock set 10:10:12 6 20 2003 //Configure the system time and date.
Verification	● Run the **show clock** command in privileged EXEC mode to display the system time.

Configuration Steps	● Change the system time to 2003-6-20, 10:10:12.
	FS# clock set 10:10:12 6 20 2003 //Configure the system time and date.
Verification	● Run the **show clock** command in privileged EXEC mode to display the system time.
	FS# show clock //Confirm that the changed system time takes effect. clock: 2003-6-20 10:10:54

Configuring Daily Notification

Configuration Steps	● Configure the daily notification message "Notice: system will shutdown on July 6th." with the pound key (#) as the delimiter.
	FS(config)# banner motd #//Starting delimiter Enter TEXT message. End with the character '#'. Notice: system will shutdown on July 6th.# //Ending delimiter FS(config)#
Verification	● Run the **show running-config** command to display the configuration. ● Connect to the local device through the Console, Telnet or SSH, and check whether daily notification is displayed before the CLI appears.
	C:\>telnet 192.168.65.236 Notice: system will shutdown on July 6th. Access for authorized users only. Please enter your password. User Access Verification Password:

Configuring a Login Banner

Configuration Steps	● Configure the login banner message "Access for authorized users only. Please enter your password." with the pound key (#) as the delimiter.
	FS(config)# banner login #//Starting delimiter Enter TEXT message. End with the character '#'. Access for authorized users only. Please enter your password. # //Ending delimiter FS(config)#

Verification	● Run the **show running-config** command to display the configuration. ● Connect to the local device through the Console, Telnet or SSH, and check whether the login banner is displayed before the CLI appears.
	C:\>telnet 192.168.65.236 Notice: system will shutdown on July 6th. Access for authorized users only. Please enter your password. User Access Verification Password:

Configuring the Serial Port Baud Rate

Configuration Steps	● Set the serial port baud rate to 57,600 bps.

	FS# configure terminal //Enter global configuration mode.
	FS(config)# line console 0 //Enter console line configuration mode.
	FS(config-line)# speed 57600 //Set the console baud rate to 57,600 bps.
	FS(config-line)# end //Returns to privileged mode.

Verification	● Run the **show** command to display the configuration.
	FS# show line console 0 //Displays the console configuration. CON Type speed Overruns * 0 CON 57600 0 Line 0, Location: "", Type: "vt100" Length: 25 lines, Width: 80 columns Special Chars: Escape Disconnect Activation ^^x none ^M Timeouts: Idle EXEC Idle Session never never History is enabled, history size is 10. Total input: 22 bytes Total output: 115 bytes Data overflow: 0 bytes stop rx interrupt: 0 times Modem: READY

2.4.4 Enabling and Disabling a Specific Service

Configuration Effect

- Dynamically adjust system services when the system is running, and enable and disable specific services (SNMP Agent, SSH Server, and Telnet Server).

Configuration Steps

Enabling the SNMP Agent, SSH Server, and Telnet Server Services

- (Optional) Perform this configuration when you need to use these services.

Verification

- Run the **show running-config** command to display the configuration.
- Run the **show services** command to display the service Enabled/Disable state.

Related Commands

Enabling the SSH Server, Telnet Server, and SNMP Agent Services

Command	**enable service** { **ssh-server** | **telnet-server** | **snmp-agent** }
Parameter Description	**ssh-server:** Enables or disables the SSH Server service. The IPv4 and IPv6 services are also enabled together with this service. **telnet-server:** Enables or disables the Telnet Server service. The IPv4 and IPv6 services are also enabled together with this service. **snmp-agent:** Enables or disables the SNMP Agent service. The IPv4 and IPv6 services are also enabled together with this service.
Command Mode	Global configuration mode
Usage Guide	Use this command to enable and disable specific services.

Configuration Example

Enabling the SSH Server Service

Configuration Steps	• Enable the SSH Server service.
	FS# configure terminal //Enter global configuration mode. FS(config)#enable service ssh-server //Enable the SSH Server service.
Verification	• Run the **show running-config** command to display the configuration. • Run the **show ip ssh** command to display the configuration and running state of the SSH Server service.

2.4.5 Configuring Multiple-configuration Booting

Configuration Effect

- Modify the path for saving startup configurations and the corresponding file name.

Notes

- The startup configuration file name consists of a path and a file name. The path is mandatory. Otherwise, configurations cannot be saved by using the **write** command. Take **Flash:/fs/fs.text** and **Usb0:/fs/fs.text** as examples, where the **Flash:/fs** and **Usb0:/fs** folders must exist. In master-slave mode, all device paths are required.
- To save the startup configuration file to a USB flash drive, the device must provide a USB interface with a USB flash drive inserted. Otherwise, configurations cannot be saved by using the **write** command. In master-slave mode, all devices must have USB flash drives connected.

Configuration Steps

Modifying the Path for Saving Startup Configurations and the Corresponding File Name

- (Optional) Perform this configuration when you need to modify the startup configuration file.

Verification

- Run the **show boot config** command to display the path for saving startup configurations and the corresponding file name.

Related Commands

Modifying the Path for Saving Startup Configurations and the Corresponding File Name

Command	**boot config** { **flash:***filename* | **usb0:***filename* }
Parameter Description	**flash:** Saves the startup configuration file in the extensible Flash. **usb0:** Saves the startup configuration file in USB0 device. The device must have a USB interface into which a USB flash drive is inserted.
Command Mode	Global configuration mode
Usage Guide	Use this command to modify the path for saving startup configurations and the corresponding file name.

Configuration Example

Changing the Path of the Startup Configuration File to Flash:/fs.text

Configuration Steps	• Change the startup configuration file path into Flash:/fs.text.
	FS# configure terminal //Enter global configuration mode. FS(config)# boot config flash:/fs.text//Change the path and file name into flash:/fs.text.

Verification	● Run the **show boot config** command to display the path for saving startup configurations and the corresponding file name.

2.4.6 Configuring a Restart Policy

Configuration Effect

Configure a restart policy to restart a device as scheduled.

Configuration Steps

Configuring Direct Restart

Run the **reload** command in privileged EXEC mode to restart the system immediately.

Configuring Timed Restart

```
reload    at    hh:mm:ss month day year
```

If you configure a specific time, the system will restart at the time. The time must be a time in the future. The **month day year** parameter is optional. If it is not specified, the system clock time is used by default.

The clock feature must be supported by the system if you want to use the **at** option. It is recommended that you configure the system clock in advance. A new restart plan will overwrite the existing one. A restart plan will be invalid if the system is restarted before the plan takes effect.

The restart time must be later than the current system time. After you configure a restart plan, do not change the system clock; otherwise, the plan may fail (for example, the system time is changed to a time after the restart time.)

Related Commands

Restarting a Device

Command	**reload** [**at** { *hh* [:*mm* [:*ss*]] } [*month* [*day* [*year*]]]]
Parameter Description	**at** *hh:mm:ss*: Indicates the time when the system will restart. *month*: Indicates a month of the year, ranging from 1 to 12. *day*: Indicates a date, ranging from 1 to 31. *year*: Indicates a year, ranging from 1993 to 2035. Abbreviation is not supported.
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to enable a device to restart at a specific time.

2.4.7 Running Batch File Commands

Configuration Effect

Run the commands in batches.

Configuration Steps

Running the execute Command

Run the **execute** command, with the path set to the batch file to be executed.

You can specify the name and content of the batch file on your PC and transfer the file to the device flash memory through TFTP. The batch processing content simulates user input. Therefore, you need to edit the batch file content according to the CLI command configuration sequence. In addition, you need to write the responses to interactive commands to the batch file to ensure normal command execution.

The batch file size must not exceed 128 KB; otherwise, it will fail to be executed. You can divide a large batch file into multiple parts not larger than 128 KB each.

Related Commands

Command	**execute** { [**flash:**] *filename* }
Parameter Description	*filename*: Indicates the path for the batch file to be executed.
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to run the commands related to a function in batches.

2.5 Monitoring

Displaying

Description	Command
show boot config	Displays the save path and file name.
show clock	Displays the current system time.
show line { **console** *line-num* | **vty** *line-num* | *line-num* }	Displays line configurations.
show reload	Displays system restart settings.
show running-config [**interface** *interface*]	Displays the current running configurations of the device or the configurations on an interface.
show startup-config	Displays the device configurations stored in the NVRAM.
show this	Displays the current system configurations.
show version [**devices** | **module** | **slots**]	Displays system information.
show sessions	Displays the information of each established Telnet client instance.

3 Configuring Lines

3.1 Overview

There are various types of terminal lines on network devices. You can manage terminal lines in groups based on their types. Configurations on these terminal lines are called line configurations. On network devices, terminal lines are classified into multiple types such as CTY, and VTY.

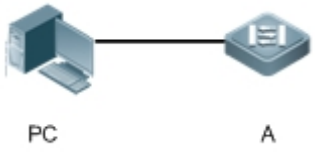
3.2 Applications

Application	Description
Accessing a Device Through Console	Enter the command-line interface (CLI) of a network device through the Console.
Accessing a Device Through VTY	Enter the CLI of a network device through Telnet or SSH.

3.2.1 Accessing a Device Through Console

Scenario

Figure 3- 1



Remarks	A is a network device to be managed. PC is a network management station.

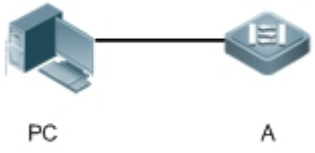
Deployment

The network management station connects to the Console port of a network device through a serial cable. Using the Console software (Hyper Terminal or other terminal simulation software) on the network management station, you can access the Console of the network device and enter the CLI to configure and manage the network device.

3.2.2 Accessing a Device Through VTY

Scenario

Figure 3- 2



Remarks	A is a network device to be managed. PC is a network management station.

Deployment

The network management station connects to a network device through the network. Using a VTY client (such as Putty) on the network management station, you can access the network device through Telnet or SSH and enter the CLI to configure and manage the network device.

3.3 Features

Basic Concepts

CTY

The CTY line refers to the line connected to the Console port. Most network devices have a Console port. You can access the local system through the Console port.

VTY

The VTY line is a virtual terminal line that does not correspond to any hardware. It is used for Telnet or SSH connection.

Overview

Feature	Description
Basic Features	Configures a terminal, displays and clears terminal connection information.

3.3.1 Basic Features

Related Configuration

Configuring Terminal Lines

Run the **line** command in global configuration mode to enter the configuration mode of a specified line.

Configure the line attributes.

Clearing Terminal Connections

When a terminal connects to the network device, the corresponding terminal line is occupied. Run the **show user** command to display the connection status of these terminal lines. If you want to disconnect the terminal from the network device, run the **clear line** command to clear the terminal line. After the terminal lines are cleared, the related connections (such as Telnet and SSH) are interrupted, the CLI exits, and the terminal lines restore to the unoccupied status. Users can re-establish connections.

Specifying the Number of VTY Terminals

Run the **line vty** command to enter the VTY line configuration mode and specify the number of VTY terminals.

By default, there are 5 VTY terminals, numbered from 0 to 4. You can increase the number of VTY terminals to 36, with new ones numbered from 5 to 35. Only new terminals can be removed.

3.4 Configuration

Configuration	Description and Command	
Entering Line Configuration Mode	(Mandatory) It is used to enter the line configuration mode.	
	line [**console** | **vty**] *first-line* [*last-line*]	Enters the specified line configuration mode.
	line vty *line-number*	Increases or reduces the number of available VTY lines.

3.4.1 Entering Line Configuration Mode

Configuration Effect

Enter line configuration mode to configure other functions.

Configuration Steps

Entering Line Configuration Mode

- Mandatory.
- Unless otherwise specified, enter line configuration mode on each device to configure line attributes.

Increasing/Reducing the Number of VTY Lines

- Optional.
- Run the (**no**) **line vty** *line-number* command to increase or reduce the number of VTY lines.

Verification

Run the **show line** command to display line configuration.

Related Commands

Entering Line Configuration Mode

Command	**line** [**console** | **vty**] *first-line* [*last-line*]
Parameter Description	**console**: Indicates the Console port. **vty**: Indicates a virtual terminal line, which supports Telnet or SSH. *first-line*: Indicates the number of the first line. *last-line*: Indicates the number of the last line.
Command Mode	Global configuration mode
Usage Guide	N/A

Increasing/Reducing the Number of VTY Lines

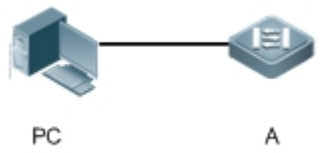
Command	**line vty** *line-number*

Parameter Description	*line-number*: Indicates the number of VTY lines. The value ranges from 0 to 35.
Command Mode	Global configuration mode
Usage Guide	Run the **no line vty** *line-number* command to reduce the number of available VTY lines.

Displaying Line Configuration

Command	**show line** { **console** *line-num* | **vty** *line-num* | *line-num* }
Parameter Description	**console**: Indicates the Console port. **vty**: Indicates a virtual terminal line, which supports Telnet or SSH. *line-num*: Indicates the line to be displayed.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuration Example

Scenario **Figure 3- 3**	
Configuration Steps	● Connect the PC to network device A through the Console line and enter the CLI on the PC. ● Run the **show user** command to display the connection status of the terminal line. ● Run the **show line console** 0 command to display the status of the Console line. ● Enter global configuration mode and run the **line vty** command to increase the number of VTY terminals to 36.
A	FS#show user Line User Host(s) Idle Location ---------- ---------- ---------- ---------- ---------- * 0 con 0 --- idle 00:00:00 --- FS#show line console 0 CON Type speed Overruns * 0 CON 9600 0

	Line 0, Location: "", Type: "vt100" Length: 24 lines, Width: 79 columns Special Chars: Escape Disconnect Activation ^^x ^D ^M Timeouts: Idle EXEC Idle Session 00:10:00 never History is enabled, history size is 10. Total input: 490 bytes Total output: 59366 bytes Data overflow: 0 bytes stop rx interrupt: 0 times FS#show line vty ? <0-5> Line number FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#line vty 35 FS(config-line)# *Oct 31 18:56:43: %SYS-5-CONFIG_I: Configured from console by console
Verification	● After running the **show line** command, you can find that the number of terminals increases. ● Run the **show running-config** command to display the configuration.
A	FS#show line vty ? <0-35> Line number FS#show running-config Building configuration... Current configuration : 761 bytes version 11.0(1C2B1)(10/16/13 04:23:54 CST -ngcf78) ip tcp not-send-rst vlan 1

```
!
interface GigabitEthernet 0/0
!
interface GigabitEthernet 0/1
 ip address 192.168.23.164 255.255.255.0
!
interface GigabitEthernet 0/2
!
interface GigabitEthernet 0/3
!
interface GigabitEthernet 0/4
!
interface GigabitEthernet 0/5
!
interface GigabitEthernet 0/6
!
interface GigabitEthernet 0/7
!
interface Mgmt 0
!
line con 0
line vty 0 35
 login
!
end
```

3.4.2 Configuring Line Attributes

Configuration Effect

Configure line attributes in line configuration mode.

Configuration Steps

Configuring the Absolute Timeout for Line Disconnection

- Optional.
- Run the **absolute-timeout** command to ensure that a line is disconnected after the specified time.

Configuring the Character You Enter at a Vacant Terminal to Begin a Terminal Session

- Optional.
- Run the **activation-character** command in line configuration mode to configure a character to activate a terminal.

Enabling Automatic Command Execution

- Optional.
- Run the **autocommand** command in line configuration mode to enable automatic command execution on terminals with asynchronous ports.

Configuring the Number of Data Bits per Character for Physical Terminal Connections

- Optional.
- Run the **databits** command in line configuration mode.

Configuring the EXEC Character Width for Physical Terminal Connections

- Optional.
- Run the **exec-character-bits** command in line configuration mode.

Configuring Flow Control Mode for Physical Terminal Connections

- Optional.
- Run the **flowcontrol** command in line configuration mode.

Configuring the Parity Bit for Physical Terminal Connections

- Optional.
- Run the **parity** command in line configuration mode.

Configuring the Start Character of Software Flow Control for Physical Terminal Connections

- Optional.
- Run the **start-character** command in line configuration mode.

Configuring the Stop Character of Software Flow Control for Physical Terminal Connections

- Optional.
- Run the **stop-character** command in line configuration mode.

Configuring the Number of Stop Bits per Byte for Physical Terminal Connections

- Optional.
- Run the **stopbits** command in line configuration mode.

Configuring the Type of Terminal Connected to a Line

- Optional.
- Run the **terminal-type** command in line configuration mode.

Verification

Run the **show line** command to display line configuration.

Related Commands

Configuring the Absolute Timeout for Line Disconnection

Command	**absolute-timeout** *minutes*
Parameter Description	*minutes*: Indicates the absolute timeout of the current line in minutes. The value ranges from 0 to 60.
Command Mode	Line configuration mode
Usage Guide	Configure the absolute timeout for line disconnection. As long as the specified time expires, the line is disconnected no matter whether you are on the operating terminal or not. Before the line is disconnected, the system displays the remaining time after which the terminal will exit: Terminal will be login out after 20 second

Configuring the Character You Enter at a Vacant Terminal to Begin a Terminal Session

Command	**activation-character** *ascii-value*
Parameter Description	*ascii-value*: Indicates the ASCII value of the hotkey character for beginning a terminal session. The value ranges from 0 to 127.
Command Mode	Line configuration mode
Usage Guide	If auto-selection is enabled for the current line, the hotkey character for beginning a terminal session must be set to the default value.

Enabling Automatic Command Execution

Command	**autocommand** *autocommand-string*
Parameter Description	*autocommand-string*: Indicates the command line to be automatically executed.
Command Mode	Line configuration mode
Usage Guide	In most cases, after a user acts as a dumb terminal to connect to a router through an asynchronous serial port, the user can remotely log in to the specified host through Telnet or obtain the specified application-based terminal service with the **autocommand** command.

Configuring the Number of Data Bits per Character for Physical Terminal Connections

Command	**databits** *bit*
Parameter	*bit*: Indicates the number of data bits per character. The value ranges from 5 to 8.

Description	
Command Mode	Line configuration mode
Usage Guide	The asynchronous hardware (such as an asynchronous serial port and AUX port) of a router generates seven data bits with parity in flow communication mode. If parity is being generated, specify 7 data bits per character. If no parity is being generated, specify 8 data bits per character. Only early devices support 5 or 6 data bits, which are seldom used.

Configuring the EXEC Character Width for Physical Terminal Connections

Command	**exec-character-bits { 7 | 8 }**
Parameter Description	**7**: Selects the 7-bit ASCII character set. **8**: Selects the 8-bit ASCII character set.
Command Mode	Line configuration mode
Usage Guide	If you need to enter Chinese characters or display Chinese characters, images, or other international characters in the command line, run the **exec-character-bits 8** command.

Configuring Flow Control Mode for Physical Terminal Connections

Command	**flowcontrol { hardware | none | software }**
Parameter Description	**hardware**: Configures hardware flow control. **none**: Configures no flow control. **software**: Configures software flow control.
Command Mode	Line configuration mode
Usage Guide	By running this command, you can specify the flow control mode to keep the Tx rate of one end the same as the Rx rate of the peer end. Since terminals cannot receive data while sending data, flow control serves to prevent data loss. When high-data-rate devices communicate with low-rate-data devices (e.g., a printer communicates with a network port), you also need to enable flow control to prevent data loss. FS general operating system (FSOS) provides two flow control modes: software flow control (controlled with control keys) and hardware flow control (controlled by hardware). The default stop character and start character for software flow control are respectively Ctrl+S (XOFF, with the ASCII value 19) and Ctrl+Q (XON, with the ASCII value 17). You can also run the **stop-character** and **start-character** commands to configure them.

Configuring the Parity Bit for Physical Terminal Connections

Command	**parity { even | none | odd }**
Parameter Description	**even**: Indicates the even parity check. **none**: Indicates no parity check. **odd**: Indicates the odd parity check.
Command	Line configuration mode

Mode	
Usage Guide	When using certain hardware (such as an asynchronous serial port and Console port) for communication, you are usually required to configure a parity bit.

Configuring the Start Character of Software Flow Control for Physical Terminal Connections

Command	**start-character** *ascii-value*
Parameter Description	*ascii-value*: Indicates the ASCII value of the start character of software flow control for physical terminal connections. The value ranges from 0 to 255.
Command Mode	Line configuration mode
Usage Guide	After software flow control is enabled, the start character for software flow control indicates the start of data transmission.

Configuring the Stop Character of Software Flow Control for Physical Terminal Connections

Command	**stop-character** *ascii-value*
Parameter Description	*ascii-value*: Indicates the ASCII value of the stop character of software flow control for physical terminal connections. The value ranges from 0 to 255.
Command Mode	Line configuration mode
Usage Guide	After software flow control is enabled, the stop character for software flow control indicates the end of data transmission.

Configuring the Number of Stop Bits per Byte for Physical Terminal Connections

Command	**stopbits { 1 | 2 }**
Parameter Description	**1**: Indicates one stop bit. **2**: Indicates two stop bits.
Command Mode	Line configuration mode
Usage Guide	You should configure the stop bits for communication between the asynchronous line and the connected network device (such as a conventional numb terminal and modem).

Configuring the Type of Terminal Connected to a Line

Command	**terminal-type** *terminal-type-string*
Parameter Description	*terminal-type-string*: Indicates the description of the terminal type, such as vt100 and ansi.
Command Mode	Line configuration mode

Usage Guide	You can run the **terminal-type vt100** command to restore the default terminal type or run the **terminal-type** command to configure the type of terminal connected to a line as required. Upon Telnet connection, one end negotiates with the other end about the terminal type based on its terminal type configuration (Telnet ID: 0x18). For details, see RFC 854.

Configuration Example

Configuring the Baud Rate, Data Bits, Parity Bits, and Stop Bits

Scenario **Figure 3- 4**	
Configuration Steps	● Connect the PC to network device A through the Console line and enter the CLI on the PC. ● Configure the baud rate, data bits, parity bit, and stop bits in global configuration mode. ● Run the **show line console** 0 command to display the status of the Console line.
A	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#line console 0 FS(config-line)#speed 115200 FS(config-line)#databits 8 FS(config-line)#parity even FS(config-line)#stopbits 1 FS#show line console 0 CON Type speed Overruns * 0 CON 115200 0 Line 0, Location: "", Type: "vt100" Length: 24 lines, Width: 79 columns Special Chars: Escape Disconnect Activation ^^x none Timeouts: Idle EXEC Idle Session 00:10:00 never History is enabled, history size is 10. Total input: 636 bytes Total output: 30498 bytes

	Data overflow: 0 bytes stop rx interrupt: 0 times
Verification	● Run the **show running-config** command to display the configuration.
A	FS#show line vty ? <0-35> Line number FS#show running-config Building configuration... Current configuration : 761 bytes version 11.0(1C2B1)(10/16/13 04:23:54 CST -ngcf78) ip tcp not-send-rst vlan 1 ! interface GigabitEthernet 0/0 ! interface GigabitEthernet 0/1 ip address 192.168.23.164 255.255.255.0 ! interface GigabitEthernet 0/2 ! interface GigabitEthernet 0/3 ! interface GigabitEthernet 0/4 ! interface GigabitEthernet 0/5 ! interface GigabitEthernet 0/6 ! interface GigabitEthernet 0/7 ! interface Mgmt 0

```
!
line con 0
 parity even
 stopbits 1
 speed 115200
line vty 0 35
 login
!
end
```

3.4.3 Configuring Terminal Attributes

Configuration Effect

Configure terminal attributes in privileged EXEC mode of a terminal.

Configuration Steps

Configuring the Number of Data Bits per Character for the Current Session

- Optional.
- Run the **terminal databits** command on the terminal.

Configuring the EXEC Character Width for the Current Session

- Optional.
- Run the **terminal exec-character-bits** command on the terminal.

Configuring Flow Control Mode for the Current Session

- Optional.
- Run the **terminal flowcontrol** command on the terminal.

Configuring the Parity Bits for the Current Session

- Optional.
- Run the **terminal parity** command on the terminal.

Configuring the Start Character of Software Flow Control for the Current Session

- Optional.
- Run the **terminal start-character** command on the terminal.

Configuring the Stop Character of Software Flow Control for the Current Session

- Optional.
- Run the **terminal stop-character** command on the terminal.

Configuring the Number of Stop Bits in Each Byte for the Current Session

- Optional.
- Run the **terminal stopbits** command on the terminal.

Configuring the Type of Terminal Connected to the Current Line for the Current Session

- Optional.
- Run the **terminal terminal-type** command on the terminal.

Verification

Run the **show line** command to display line configuration.

Related Commands

Configuring the Number of Data Bits per Character for the Current Session

Command	**terminal databits** *bit*
Parameter Description	*bit*: Indicates the number of data bits per character, ranging from 5 to 8.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuring the EXEC Character Width for the Current Session

Command	**terminal exec-character-bits { 7 | 8 }**
Parameter Description	**7**: Selects the 7-bit ASCII character set. **8**: Selects the full 8-bit ASCII character set.
Command Mode	Privileged EXEC mode
Usage Guide	If you need to enter Chinese characters or display Chinese characters, images, or other international characters in the command line, run the **terminal exec-character-bits 8** command.

Configuring Flow Control Mode for the Current Session

Command	**terminal flowcontrol { hardware | none | software }**
Parameter Description	**hardware**: Configures hardware flow control. **none**: Configures no flow control. **software**: Configures software flow control.
Command Mode	Privileged EXEC mode

Usage Guide	N/A

Configuring the Parity Bit of the Asynchronous Line for the Current Session

Command	**terminal parity { even | none | odd }**
Parameter Description	**even**: Indicates the even parity check. **none**: Indicates no parity check. **odd**: Indicates the odd parity check.
Command Mode	Line configuration mode
Usage Guide	When using certain hardware (such as an asynchronous serial port and Console port) for communication, you are usually required to configure a parity bit.

Configuring the Start Character of Software Flow Control for the Current Session

Command	**terminal start-character** *ascii-value*
Parameter Description	*ascii-value*: Indicates the ASCII value of the start character of software flow control for the current session. The value ranges from 0 to 255.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuring the Stop Character of Software Flow Control for the Current Session

Command	**terminal stop-character** *ascii-value*
Parameter Description	*ascii-value*: Indicates the ASCII value of the stop character of for the current session. The value ranges from 0 to 255.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuring the Number of Stop Bits for the Current Session

Command	**terminal stopbits { 1 | 2 }**
Parameter Description	**1**: Indicates one stop bit. **2**: Indicates two stop bits.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuring the Type of Terminal Connected to the Current Line for the Current Session

Command	**terminal terminal-type** *terminal-type-string*
Parameter Description	*terminal-type-string*: Indicates the description of the terminal type, such as vt100 and ansi.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuration Example

Configuring the Terminal Type and Baud Rate of a Terminal

Scenario **Figure 3- 5**	
Configuration Steps	● Connect the PC to network device A through the Console line and enter the CLI on the PC. ● Configure the terminal type and baud rate of the terminal in privileged EXEC mode.
A	FS#terminal terminal-type ansi FS#terminal speed 115200
Verification	● Run the **show line console** 0 command to display the status of the Console line.
A	FS#show line console 0 CON Type speed Overruns * 0 CON **115200** 0 Line 0, Location: "", Type: "**ansi**" Length: 24 lines, Width: 79 columns Special Chars: Escape Disconnect Activation ^^x none Timeouts: Idle EXEC Idle Session 00:10:00 never History is enabled, history size is 10. Total input: 858 bytes Total output: 57371 bytes Data overflow: 0 bytes stop rx interrupt: 0 times

3.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the line connection status.	**clear line** { **console** *line-num* | **vty** *line-num* | *line-num* }

Displaying

Description	Command
Displays the line configuration.	**show line** { **console** *line-num* | **vty** *line-num* | *line-num* }
Displays historical records of a line.	**show history**
Displays the privilege level of a line.	**show privilege**
Displays users on a line.	**show user** [**all**]

4 Configuring Time Range

4.1 Overview

Time Range is a time-based control service that provides some applications with time control. For example, you can configure a time range and associate it with an access control list (ACL) so that the ACL takes effect within certain time periods of a week.

4.2 Typical Application

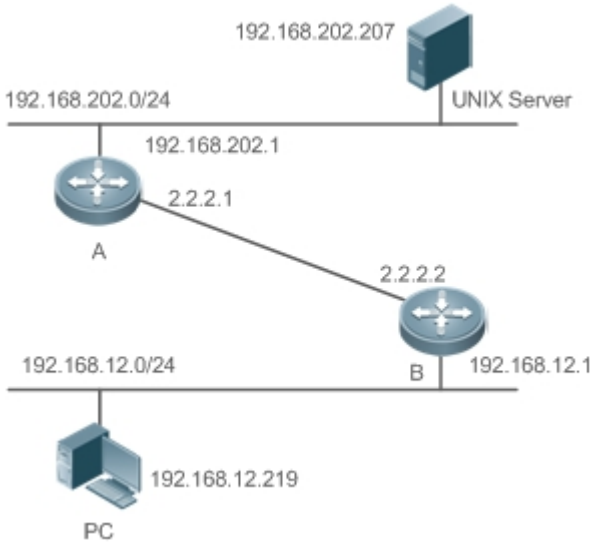
Typical Application	Scenario
Applying Time Range to an ACL	Apply a time range to an ACL module so that the time-based ACL takes effect

4.2.1 Applying Time Range to an ACL

Application Scenario

An organization allows users to access the Telnet service on a remote Unix host during working hours only, as shown in Figure 4- 1.

Figure 4- 1



Note	Configure an ACL on device B to implement the following security function: Hosts in network segment 192.168.12.0/24 can access the Telnet service on a remote Unix host during normal working hours only.

Functional Deployment

- On device B, apply an ACL to control Telnet service access of users in network segment 192.168.12.0/24. Associate the ACL with a time range, so that the users' access to the Unix host is allowed only during working hours.

4.3 Function Details

Basic Concepts

Absolute Time Range

The absolute time range is a time period between a start time and an end time. For example, [12:00 January 1 2000, 12:00 January 1 2001] is a typical absolute time range. When an application based on a time range is associated with the time range, a certain function can be effective within this time range.

Periodic Time

Periodic time refers to a periodical interval in the time range. For example, "from 8:00 every Monday to 17:00 every Friday" is a typical periodic time interval. When a time-based application is associated with the time range, a certain function can be effective periodically from every Monday to Friday.

Features

Feature	Function
Using Absolute Time Range	Sets an absolute time range for a time-based application, so that a certain function takes effect within the absolute time range.
Using Periodic Time	Sets periodic time or a time-based application, so that a certain function takes effect within the periodic time.

4.3.1 Using Absolute Time Range

Working Principle

When a time-based application enables a certain function, it determines whether current time is within the absolute time range. If yes, the function is effective or ineffective at the current time depending on specific configuration.

4.3.2 Using Periodic Time

Working Principle

When a time-based application enables a certain function, it determines whether current time is within the period time. If yes, the function is effective or ineffective at the current time depending on specific configuration.

4.4 Configuration Details

Configuration Item	Suggestions and Related Commands	
Configuring Time Range	Mandatory configuration. Time range configuration is required so as to use the time range function.	
	time-range *time-range-name*	Configures a time range.
	Optional configuration. You can configure various parameters as necessary.	
	absolute { [**start** *time date*] | [**end** *time date*] }	Configures an absolute time range.
	periodic *day-of-the-week time* **to** [*day-of-the-week*] *time*	Configures periodic time.

4.4.1 Configuring Time Range

Configuration Effect

- Configure a time range, which may be an absolute time range or a periodic time interval, so that a time-range-based application can enable a certain function within the time range.

Configuration Method

Configuring Time Range

- Mandatory configuration.
- Perform the configuration on a device to which a time range applies.

Configuring Absolute Time Range

- Optional configuration.

Configuring Periodic Time

- Optional configuration.

Verification

- Use the **show time-range** [*time-range-name*] command to check time range configuration information.

Related Commands

Configuring Time Range

Command Syntax	**time-range** *time-range-name*
Parameter Description	*time-range-name*: name of the time range to be created.
Command Mode	Global configuration mode

Usage Guide	Some applications (such as ACL) may run based on time. For example, an ACL can be effective within certain time ranges of a week. To this end, first you must configure a time range, then you can configure relevant time control in time range configuration mode.

Configuring Absolute Time Range

Command Syntax	**absolute** { [**start** *time date*] | [**end** *time date*] }
Parameter Description	**start** *time date*: start time of the range. **end** *time date*: end time of the range.
Command Mode	Time range configuration mode
Usage Guide	Use the **absolute** command to configure a time absolute time range between a start time and an end time to allow a certain function to take effect within the absolute time range.

Configuring Periodic Time

Command Syntax	**periodic** *day-of-the-week time* **to** [*day-of-the-week*] *time*
Parameter Description	*day-of-the-week*: the week day when the periodic time starts or ends *time*: the exact time when the periodic time starts or ends
Command Mode	Time range configuration mode
Usage Guide	Use the **periodic** command to configure a periodic time interval to allow a certain function to take effect within the periodic time.

4.5 Monitoring and Maintaining Time Range

Displaying the Running Status

Function	Command
Displays time range configuration.	**show time-range** [*time-range-name*]

5 Configuring HTTP Service

5.1 Overview

Hypertext Transfer Protocol (HTTP) is used to transmit Web page information on the Internet. It is at the application layer of the TCP/IP protocol stack. The transport layer adopts connection-oriented Transmission Control Protocol (TCP).

Hypertext Transfer Protocol Secure (HTTPS) is an HTTP supporting the Secure Sockets Layer (SSL) protocol. HTTPS is mainly used to create a secure channel on an insecure network, ensure that information can hardly be intercepted, and provide certain reasonable protection against main-in-the-middle attacks. At present, HTTPS is widely used for secure and sensitive communication on the Internet, for example, electronic transactions.

Protocols and Standards

- RFC1945: Hypertext Transfer Protocol -- HTTP/1.0
- RFC2616: Hypertext Transfer Protocol -- HTTP/1.1
- RFC2818: Hypertext Transfer Protocol Over TLS -- HTTPS

5.2 Applications

Application	Description
HTTP Application Service	Users manage devices based on Web.

5.2.1 HTTP Application Service

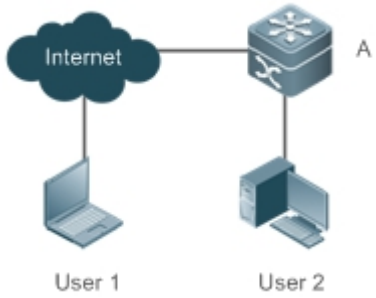
Scenario

After the HTTP service is enabled, users can access the Web management page after passing authentication by only entering **http://IP address of a device** in the browser of a PC. On the Web page, users you can monitor the device status, configure devices, upload and download files.

Take the following figure as an example to describe Web management.

- Users can remotely access devices on the Internet or configure and manage devices on the Local Area Network (LAN) by logging in to the Web server.
- According to actual conditions, users can choose to enable the HTTPS or HTTP service or enable the HTTPS and HTTP services at the same time.
- Users can also access the HTTP service of devices by setting and using HTTP/1.0 or HTTP/1.1 in the browser.

Figure 5- 1



Remarks	A is a FS device. User 1 accesses the device through the Internet. User 2 accesses the device through a LAN.

Deployment

- When a device runs HTTP, users can access the device by entering **http://IP address of the device** in the browser of a PC.
- When a device runs HTTPS, users can access the device by entering **https://IP address of the device** in the browser of a PC.

5.3 Features

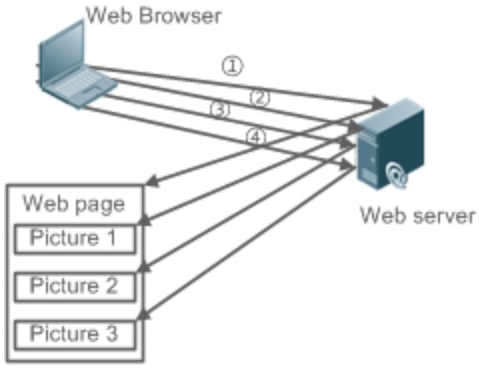
Basic Concepts

↘ HTTP Service

The HTTP service refers to transmission of Web page information on the Internet by using HTTP. HTTP/1.0 is currently an HTTP version that is the most widely used. As one Web server may receive thousands or even millions of access requests, HTTP/1.0 adopts the short connection mode to facilitate connection management. One TCP connection is established for each request. After a request is completed, the TCP connection is released. The server does not need to record or trace previous requests. Although HTTP/1.0 simplifies connection management, HTTP/1.0 introduces performance defects.

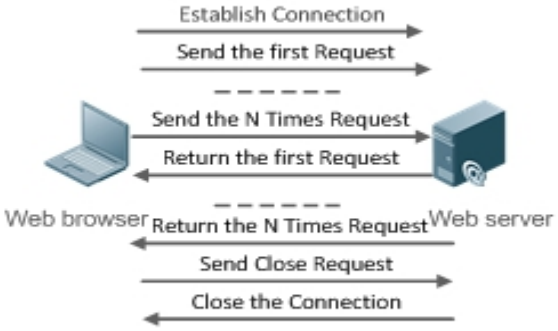
For example, a web page my need lots of pictures. However, the web page contains not real picture contents but URL connection addresses of the pictures. In this case, the browser sends multiple requests during access. Each request requires establishing an independent connection and each connection is completely isolated. Establishing and releasing connections is a relatively troublesome process, which severely affects the performance of the client and server, as shown in the following figure:

Figure 5- 2



HTTP/1.1 overcomes the defect. It supports persistent connection, that is, one connection can be used to transmit multiple requests and response messages. In this way, a client can send a second request without waiting for completion of the previous request. This reduces network delay and improves performance. See the following figure:

Figure 5- 3



At present, FS devices support both HTTP/1.0 and HTTP/1.1.

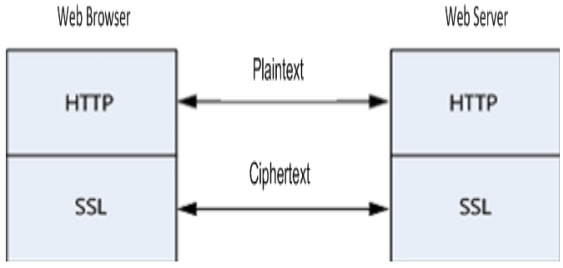
Which HTTP version will be used by a device is decided by the Web browser.

HTTPS Service

The HTTPS service adds the SSL based on the HTTP service. Its security basis is the SSL. To run HTTPS properly, a server must have a Public Key Infrastructure (PKI) certificate while a client may not necessarily need one. The SSL protocol provides the following services:

- Authenticating users and servers and ensuring that data is sent to the correct client and server.
- Encrypting data to prevent data from being stolen midway.
- Maintaining data integrity and ensuring that data is not changed during transmission.

Figure 5- 4



- During a local upgrade, a device serves as an HTTP server. Users can log in to the device through a Web browser and upload upgrade files to the device to realize file upgrade on the device.

Features

Feature	Description
HTTP Service	Users log in to devices through Web pages to configure and manage devices.
Local HTTP Upgrade Service	Upgrade files are uploaded to a device to realize file upgrade on the device.

5.3.1 HTTP Service

HTTP is a service provided for Web management. Users log in to devices through Web pages to configure and manage devices.

Working Principle

Web management covers Web clients and Web servers. Similarly, the HTTP service also adopts the client/server mode. The HTTP client is embedded in the Web browser of the Web management client. It can send HTTP packets and receive HTTP response packets. The Web server (namely HTTP server) is embedded in devices. The information exchange between the client and the server is as follows:

- A TCP connection is established between the client and the server. The default port ID of the HTTP service is 80 and the default port ID of the HTTPS service is 443.
- The client sends a request message to the server.
- The server resolves the request message sent by the client. The request content includes obtaining a Web page, executing a CLI command, and uploading a file.
- After executing the request content, the server sends a response message to the client.

Related Configuration

Enabling the HTTP Service

By default, the HTTP service is disabled.

The **enable service web-server** command can be used to enable HTTP service functions, including the HTTP service and HTTPS service.

The HTTP service must be enabled so that users can log in to devices through Web pages to configure and manage devices.

Configuring HTTP Authentication Information

By default, the system creates the **admin** account. The account cannot be deleted and only the password of the account can be changed. The administrator account is the admin account, which corresponds to the level 0 permission. The administrator account owns all permissions on the Web client and can edit other management accounts and authorize the accounts to access pages. The new accounts that are added correspond to the level 1 permission.

The **webmaster level** command can be used to configure an authenticated user name and a password.

After this command is run, you need to enter the configured user name and password to log in to the Web page.

Configuring an HTTP Service Port

By default, the HTTP service port ID is 80.

The **http port** command can be used to configure an HTTP service port ID. The value range of the port ID is 80 and 1025 to 65535.

By configuring an HTTP service port ID, you can reduce the number of attacks initiated by illegal users on the HTTP service.

Configuring an HTTPS Service Port

By default, the HTTPS service port ID is 443.

The **http secure-port** command can be used to configure an HTTPS service port ID. The value range of the port ID is 443 and 1025 to 65535.

By configuring an HTTPS service port ID, you can reduce the number of attacks initiated by illegal users on the HTTPS service.

5.3.2 Local HTTP Upgrade Service

When a device serves as the HTTP server, users can log in to the device through a Web browser and upload upgrade files (including component package and Web package) to the device or directly upload files to the device through Trivial File Transfer Protocol (TFTP).

Working Principle

- A component package or Web package is uploaded through the local upgrade function provided by Web.
- After successfully receiving a file, the device checks the version for its validity.
- After the file check is successful, if the file is a Web package, perform the upgrade directly; if the file is a component package, decide whether to perform the upgrade in the browser by restarting the device.

Related Configuration

Updating a Web Package

Run the **upgrade web download** command to download a Web package from the TFTP server.

After the command is run, download a Web package from the TFTP server. After the package passes the validity check, directly use the Web package for upgrade without restarting the device.

You can also run the **upgrade web** command to directly upgrade a Web package stored locally.

Updating a Subsystem Component

By default, a device does not upgrade subsystem components uploaded through a browser or TFTP.

To upgrade a subsystem component, you must restart the device.

5.4 Configuration

Configuration	Description and Command	
Configuring the HTTP Service	(Mandatory) It is used to enable the HTTP service.	
	enable service web-server	Enables the HTTP service.
	webmaster level	Configures HTTP authentication information.
	http port	Configures an HTTP service port.
	http secure-port	Configures an HTTPS service port.
Configuring HTTP redirection to HTTPS	(Optional) It is used to configure HTTP redirection to HTTPS.	
	web-server http redirect-to-https	Configures HTTP redirection to HTTPS.

Configuration	Description and Command	
Configuring HTTPS Certificate	(Optional) It is used to configure HTTPS certificate.	
	web-server https generate self-signed-certificate	Re-generates HTTP self-signed certificate.
	web-server https certificate	Installs HTTPS certificate.
Configuring a Local HTTP Upgrade	(Mandatory) It is used to realize a local HTTP upgrade.	
	upgrade web	Upgrades a Web package stored on a device.
	upgrade web download	Automatically downloads a Web package from a server and automatically upgrades the package.

5.4.1 Configuring the HTTP Service

Configuration Effect

After the HTTP service is enabled on a device, users can log in to the Web management page after passing authentication and monitor the device status, configure devices, upload and download files.

Configuration Steps

Enabling the HTTP Service

- Mandatory
- If there is no special requirement, enable the HTTP service on FS devices. Otherwise, the Web service is inaccessible.

Configuring HTTP Authentication Information

- By default, the user name **admin** and the password **admin** are configured.
- If there is no special requirement, you can log in to the Web page by using the default user name and directly update authentication information through the Web browser. If you always use the default account, security risks may exist because unauthorized personnel can obtain device configuration information once the IP address is disclosed.

Configuring an HTTP Service Port

- If an HTTP service port needs to be changed, the HTTP service port must be configured.
- If there is no special requirement, the default HTTP service port 80 can be used for access.

Configuring an HTTPS Service Port

- If an HTTPS service port needs to be changed, the HTTPS service port must be configured.
- If there is no special requirement, the default HTTPS service port 443 can be used for access.

Verification

- Enter **http://IP address of the device: service port** to check whether the browser skips to the authentication page.
- Enter **https://IP address of the device: service port** to check whether the browser skips to the authentication page.

Related Commands

Enabling the HTTP Service

Command	**enable service web-server** [**http** | **https** | **all**]
Parameter Description	**http** | **https** | **all**: Enables the corresponding service. **http** indicates enabling the HTTP service, **https** indicates enabling the HTTPS service, and **all** indicates enabling the HTTP and HTTPS services at the same time. By default, the HTTP and HTTPS services are enabled at the same time.
Command Mode	Global configuration mode.
Usage Guide	If no key word or **all** is put at the end of the command when the command is run, the HTTP and HTTPS services are enabled at the same time. If the key word **http** is put at the end of the command, only the HTTP service is enabled; if the key word **https** is put at the end of the command, only the HTTPS service is enabled. The **no enable service web-server** or **default enable service web-server** command is used to disable the corresponding HTTP service. If no key word is put at the end of the **no enable service web-server** or **default enable service web-server** command, the HTTP and HTTPS services are disabled.

Configuring HTTP Authentication Information.

Command	**webmaster level** *privilege-level* **username** *name* **password** { *password* | [**0 | 7**] *encrypted-password* }
Parameter Description	*privilege-level*: Permission level bound to a user. *name*: User name. *password*: User password. **0 | 7**: Password encryption type. 0: no encryption; 7: simple encryption. The default value is **0**. *encrypted-password*: Password text.
Command Mode	Global configuration mode.
Usage Guide	When the HTTP server is used, you need to be authenticated before logging in to the Web page. The **webmaster level** command is used to configure a user name and a password for logging in to the Web page. Run the **no webmaster level** *privilege-level* command to delete all user names and passwords of the specified permission level. Run the **no webmaster level** *privilege-level* **username** *name* command to delete the specified user name and password. User names and passwords involve three permission levels: Up to 10 user names and passwords can be configured for each permission level. By default, the system creates the **admin** account. The account cannot be deleted and only the password of the account can be changed. The administrator account is the **admin** account, which corresponds to the level 0 permission. The administrator account owns all permissions on the Web client and can edit other management accounts and authorize the accounts to access pages. The new accounts that are added correspond to the level 1 permission.

Configuring an HTTP Service Port

Command	**http port** *port-number*

Parameter Description	*port-number*: Configures an HTTP service port. The value range is 80 and 1025 to 65535.
Command Mode	Global configuration mode.
Usage Guide	Run the command to set an HTTP service port.

Configuring an HTTPS Service Port

Command	**http secure-port** *port-number*
Parameter Description	*port-number*: Configures an HTTPS service port. The value range is 443 and 1025 to 65535.
Command Mode	Global configuration mode.
Usage Guide	Run the command to set an HTTPS service port.

Configuration Example

Managing one FS Device by Using Web and Logging in to the Device through a Web Browser to Configure Related Functions

- Log in to the device by using the **admin** account configured by default.
- To improve security, the Web browser is required to support both HTTP and HTTPS for access.
- The user is required to configure an HTTP service port to reduce the number of attacks initiated by illegal users on HTTP.

Scenario **Figure 5- 5**	
Configuration Steps	• Enable the HTTP and HTTPS services at the same time. • Set the HTTP service port ID to 8080 and the HTTPS service port ID to 4430.
A	A#configure terminal A(config)# enable service web-server A(config)# http port 8080 A(config)# http secure-port 4430
Verification	Check HTTP configurations.
A	A# show web-server status http server status: enabled http server port: 8080

	https server status:enabled https server port: 4430

Common Errors

- If the HTTP service port is not the default port 80 or 443, you must enter a specific configured service port in the browser. Otherwise, you cannot access devices on the Web client.

5.4.2 Configuring HTTP Redirection to HTTPS

Configuration Effect

After enabling HTTP and HTTPS, users can configure HTTP redirection to HTTPS to improve security.

Configuration Steps

HTTP Redirection to HTTPS

- HTTP redirection to HTTPS is disabled by default.
- Run the **web-server http redirect-to-https** command to enable HTTP redirection to HTTPS.
- Configure HTTP redirection to HTTPS to improve security.

Command	**web-server http redirect-to-https**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Run the **no web-server http redirect-to-https** or **default web-server http redirect-to-https** command to configure HTTP redirection to HTTPS. HTTP and HTTPS must be enabled first. If the target IP address is a NAPT address, HTTP redirection to HTTPS may fail. In this case, please disable HTTP first and use HTTPS to access this IP address.

Verification

- Enter **http://Device IP:HTTP port** into the address bar of the browser to verify whether the browser will redirect to **http://Device IP:HTTP port**.
- Run the **show web-server status** command to configure HTTP redirection to HTTPS.

Configuration Example

Using Browser to Access Web

- Configure HTTP redirection to HTTPS to improve security.

Scenario

Figure 5- 6

Configuration Steps

- Enable both HTTP and HTTPS.
- Configure HTTP redirection to HTTPS.

A

```
A#configure terminal
A(config)# enable service web-server
A(config)# web-server http redirect-to-https
```

Verification

- Check Web status.
- Enter **http://Device IP:HTTP port** into the address bar of the browser to verify whether the browser will redirect to **http://Device IP:HTTP port**.

A

```
A(config)#show web-server status
http server status: enabled
http server port: 80
https server status:enabled
https server port: 443
http redirect to https: true
```

5.4.3 Configuring HTTPS Certificate

Configuration Effect

Configure HTTPS certificate to re-generate the self-signed certificate or the certificate assigned by Certificate Authority.

Configuration Steps

Re-generating HTTPS Self-signed Certificate

- HTTPS self-signed certificate is used by default.
- Run the **web-server https generate self-signed-certificate** command to re-generate the HTTPS certificate.

Command	**web-server https generate self-signed-certificate**
Parameter Description	N/A
Command	Global configuration mode

Mode

Usage Guide This command is an interactive command. After running this command, please enter the information required or press Ctrl+C to abort the task.

If the HTTPS certificate is installed, the HTTPS certificate will be used preferentially. The re-generated self-signed certificate will not replace the HTTPS certificate.

This command is not displayed in running-config.

It is recommended to open the Web management page again after closing the browser.

Installing HTTPS Certificate

- HTTPS self-signed certificate is used by default.
- Run the **web-server https certificate** command to install the HTTPS certificate.
- Installing the HTTPS certificate assigned by the Certificate Authority will prevent distrust prompt popping up during HTTPS access.

Command **web-server https certificate** { **pem** *cert-filename* **private-key** *key-filename* } | { **pfx** *cert-filename* } [**password** *password-text*]

Parameter Description **pem**: Imports the certificate and private key file in pem format.

pfx: Imports the certificate file in pfx format.

cert-filename: Specifies the name of the certificate file under the flash: directory.

key-filename: Specifies the name of the private key file under the flash: directory.

password *password-text*: Configures the decryption password.

Command Mode Global configuration mode

Usage Guide Run the **copy** command to copy the certificate/private key file to the **flash:** partition. After installation finishes, the certificate/private key file can be deleted.

You can run the **no web-server https certificate** command to delete the HTTPS certificate. Afterwards, the auto-signed HTTPS certificate will be used.

This command is not displayed in running-config.

It is recommended to open the Web management page again after closing the browser.

Verification

- Run the **show web-server https certificate information** command to display HTTPS certificate.

Configuration Example

Re-generating HTTPS Self-signed Certificate

Scenario

Figure 5- 7

Configuration Steps

- Re-generate the self-signed certificate.

A

```
A#configure terminal
A(config)# web-server https generate self-signed-certificate
RSA key modulus bits (1024~4096) [2048]:
Common Name (e.g. server IP) [Self-Signed-600B16C2]:
% Generate self-signed certificate successfully
```

Verification

- Run the **show web-server https certificate information** command to display certificate.

A

```
A#show web-server https certificate information
Source: Default
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number: 1 (0x1)
    Signature Algorithm: sha256WithRSAEncryption
        Issuer: CN=Self-Signed-600B16C2
        Validity
            Not Before: Feb 28 05:49:39 2019 GMT
            Not After : Feb 25 05:49:39 2029 GMT
        Subject: CN=Self-Signed-600B16C2
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (2048 bit)
                Modulus:
……
A#
```

Installing Third-party HTTPS Certificate

- The certificate file name is usercert.pfx and the key is 123456. Enable the TFTP server and place the certificate file under the TFTP server directory.

Scenario

Figure 5- 8

Configuration Steps

- Run the **copy** command to copy the certificate/private key file to the **flash:** partition.
- Run the **web-server https certificate** command to install HTTPS certificate.

A

```
A# copy tftp://192.168.1.1/usercert.pfx flash:usercert.pfx
Press Ctrl+C to quit
!
Copy success.
A#configure terminal
A(config)# web-server https certificate pfx usercert.pfx password 123456
*Feb 28 14:38:37: %HTTPD-4-CERT_CHANGE: HTTPS certificate changed.
% The certificate was successfully installed.
```

Verification

- Run the **show web-server https certificate information** command to display certificate.

A

```
A#show web-server https certificate information
Source: Installed
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number: 4 (0x4)
    Signature Algorithm: sha1WithRSAEncryption
        Issuer: C=CN, CN=mytestCA
        Validity
            Not Before: Jan 23 08:36:21 2019 GMT
            Not After : Jan 23 08:36:21 2020 GMT
        Subject: C=CN, CN=test-cert-2
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (2048 bit)
                Modulus:
……
 A#
```

5.4.4 Configuring a Local HTTP Upgrade

Configuration Effect

Perform an HTTP upgrade through the browser or the **upgrade web** command.

Notes

- So long as a Web package is uploaded successfully and passes the version check, the device directly performs an upgrade based on the latest Web package.
- The **upgrade web download** command is used to automatically download files from the TFTP server and automatically perform an upgrade.
- The **upgrade web** command is used to automatically upgrade the Web package in the local file system.

Configuration Steps

N/A

Verification

- Access and view the latest Web page through the browser.

Related Commands

Downloading a Web Package from the TFTP Server

Command	**upgrade web download { oob_tftp:** */path* | **tftp:** */path* **}**
Parameter Description	*oob_tftp:* Connects to the TFTP server through the management port and downloads a Web package. Only the devices that support the management port support the parameter. *tftp:* Connects the FTFP server through a common data port and downloads a Web package. *path*: Path of a Web package on the TFTP server.
Command Mode	Privileged mode
Usage Guide	This command is used to download a Web package from the TFTP server and automatically perform an upgrade.

Upgrading a Web Package Stored on a Local Device

Command	**upgrade web** *uri*
Parameter Description	*uri*: Local path for storing a Web package.
Command Mode	Privileged mode
Usage Guide	This command is used to upgrade a Web package stored on a device and automatically perform an upgrade.

Configuration Example

Obtaining the Latest Web Package from the Official Website and Running the Web Package

Scenario **Figure 5-6**	
Configuration Steps	● Connect to a local PC whose IP address is 10.10.10.13 and assign an IP address 10.10.10.131 in the same network segment to the device. ● Log in to the device through Web and upload the latest Web package to the device.
A	A#configure terminal A(config)# vlan 1 A(config-vlan)# exit A(config)# interface vlan 1 A(config-VLAN 1)# ip address 10.10.10.131 255.255.255.0 A(config-VLAN 1)# exit A(config)# enable service web-server
	On a PC, use the local upgrade function on the Web page to upload a Web package for upgrade.
Verification	On the PC, log in to the device through Web again and check whether the latest Web page is displayed.

Upgrading a Web Package by Running the upgrade web download Command

Scenario **Figure 5-7**	
Configuration Steps	● Connect to a local PC whose IP address is 10.10.10.13 and assign an IP address 10.10.10.131 in the same network segment to the device. ● Start the TFTP server.
A	A#configure terminal A(config)# vlan 1 A(config-vlan)# exit A(config)# interface vlan 1 A(config-VLAN 1)# ip address 10.10.10.131 255.255.255.0 A(config-VLAN 1)# end A#upgrade web download tftp:// 10.10.10.13/web.upd Press Ctrl+C to quit !!!!!!!! download 3896704 bytes

	Begin to upgrade the web package... Web package upgrade successfully.
Verification	On the PC, log in to the device through Web again and check whether the latest Web page is displayed.

Upgrading a Web Package by Running the upgrade web Command

Scenario **Figure 5- 8**	
Configuration Steps	● Connect to a local PC whose IP address is 10.10.10.13 and assign an IP address 10.10.10.131 in the same network segment to the device. ● Start the TFTP server.
A	A#configure terminal A(config)# vlan 1 A(config-vlan)# exit A(config)# interface vlan 1 A(config-VLAN 1)# ip address 10.10.10.131 255.255.255.0 A(config-VLAN 1)# end A#copy tftp://10.10.10.13/web.upd flash:/web.upd Press Ctrl+C to quit !!!!!!!! Accessing tftp:// 10.10.10.13/web.upd finished, 3896704 bytes prepared Flushing data to flash:/web.upd... Flush data done A #upgrade web flash:/web.upd Web package upgrade successfully. A #
Verification	On the PC, log in to the device through Web again and check whether the latest Web page is displayed.

Common Errors

● Access to the web page through the browser shows that the web page is not updated based on the latest Web package. This is possibly because the local browser has a cache. Clear the cache of the local browser and access the Web page again.

5.5 Monitoring

Displaying

Description	Command
Displays the configuration and status of the Web service.	**show web-server status**

6 Configuring Syslog

6.1 Overview

Status changes (such as link up and down) or abnormal events may occur anytime. FS products provide the syslog mechanism to automatically generate messages (log packets) in fixed format upon status changes or occurrence of events. These messages are displayed on the related windows such as the Console or monitoring terminal, recorded on media such as the memory buffer or log files, or sent to a group of log servers on the network so that the administrator can analyze network performance and identify faults based on these log packets. Log packets can be added with the timestamps and sequence numbers and classified by severity level so that the administrator can conveniently read and manage log packets.

Protocols and Standards

- RFC3164: The BSD syslog Protocol
- RFC5424: The_Syslog_Protocol

6.2 Applications

Application	Description
Sending Syslogs to the Console	Monitor syslogs through the Console.
Sending Syslogs to the Log Server	Monitor syslogs through the server.

6.2.1 Sending Syslogs to the Console

Scenario

Send syslogs to the Console to facilitate the administrator to monitor the performance of the system. The requirements are as follows:

1. Send logs of Level 6 or higher to the Console.
2. Send logs of only the ARP and IP modules to the Console.

Figure 6-1 shows the network topology.

Figure 6- 1 Network topology

Deployment

Configure the device as follows:

1. Set the level of logs that can be sent to the Console to informational (Level 6).
2. Set the filtering direction of logs to terminal.
3. Set log filtering mode of logs to contains-only.
4. Set the filtering rule of logs to single-match. The module name contains only ARP or IP.

6.2.2 Sending Syslogs to the Log Server

Scenario

Send syslogs to the log server to facilitate the administrator to monitor the logs of devices on the server. The requirements are as follows:

1. Send syslogs to the log server 10.1.1.1.
2. Send logs of Level 7 or higher to the log server.
3. Send syslogs from the source interface Loopback 0 to the log server.

Figure 6- 2 shows the network topology.

Figure 6- 2 Network topology

Deployment

Configure the device as follows:

1. Set the IPv4 address of the server to 10.1.1.1.
2. Set the level of logs that can be sent to the log server to debugging (Level 7).
3. Set the source interface of logs sent to the log server to Loopback 0.

6.3 Features

Basic Concepts

Classification of Syslogs

Syslogs can be classified into two types:

- Log type
- Debug type

Levels of Syslogs

Eight severity levels of syslogs are defined in descending order, including emergency, alert, critical, error, warning, notification, informational, and debugging. These levels correspond to eight numerical values from 0 to 7. A smaller value indicates a higher level.

Only logs with a level equaling to or higher than the specified level can be output. For example, if the level of logs is set to informational (Level 6), logs of Level 6 or higher will be output.

The following table describes the log levels.

Level	Numerical Value	Description
emergencies	0	Indicates that the system cannot run normally.
alerts	1	Indicates that the measures must be taken immediately.

critical	2	Indicates a critical condition.
errors	3	Indicates an error.
warnings	4	Indicates a warning.
notifications	5	Indicates a notification message that requires attention.
informational	6	Indicates an informational message.
debugging	7	Indicates a debugging message.

Output Direction of Syslogs

Output directions of syslogs include Console, monitor, server, buffer, and file. The default level and type of logs vary with the output direction. You can customize filtering rules for different output directions.

The following table describes output directions of syslogs.

Output Direction	Description	Default Output Level	Description
Console	Console	Debugging (Level 7)	Logs and debugging information are output.
monitor	Monitoring terminal	Debugging (Level 7)	Logs and debugging information are output.
server	Log server	Informational (Level 6)	Logs and debugging information are output.
buffer	Log buffer	Debugging (Level 7)	Logs and debugging information are output. The log buffer is used to store syslogs.
file	Log file	Informational (Level 6)	Logs and debugging information are output. Logs in the log buffer are periodically written into files.

RFC3164 Log Format

Formats of syslogs may vary with the syslog output direction.

- If the output direction is the Console, monitor, buffer, or file, the syslog format is as follows:

```
seq no: *timestamp: sysname %module-level-mnemonic: content
```

For example, if you exit configuration mode, the following log is displayed on the Console:

```
001233: *May 22 09:44:36: FS %SYS-5-CONFIG_I: Configured from console by console
```

- If the output direction is the log server, the syslog format is as follows:

```
<priority>seq no: *timestamp: sysname %module-level-mnemonic: content
```

For example, if you exit configuration mode, the following log is displayed on the log server:

```
<189>001233: *May 22 09:44:36: FS %SYS-5-CONFIG_I: Configured from console by console
```

The following describes each field in the log in details:

4. Priority

This field is valid only when logs are sent to the log server.

The priority is calculated using the following formula: Facility x 8 + Level Level indicates the numerical code of the log level and Facility indicates the numerical code of the facility. The default facility value is local7 (23). The following table lists the value range of the facility.

Numerical Code	Facility Keyword	Facility Description
0	kern	kernel messages
1	user	user-level messages
2	mail	mail system
3	daemon	system daemons
4	auth1	security/authorization messages
5	syslog	messages generated internally by syslogs
6	lpr	line printer subsystem
7	news	network news subsystem
8	uucp	UUCP subsystem
9	clock1	clock daemon
10	auth2	security/authorization messages
11	ftp	FTP daemon
12	ntp	NTP subsystem
13	logaudit	log audit
14	logalert	log alert
15	clock2	clock daemon
16	local0	local use 0 (local0)
17	local1	local use 1 (local1)
18	local2	local use 2 (local2)
19	local3	local use 3 (local3)
20	local4	local use 4 (local4)
21	local5	local use 5 (local5)
22	local6	local use 6 (local6)
23	local7	local use 7 (local7)

5. Sequence Number

The sequence number of a syslog is a 6-digit integer, and increases sequentially. By default, the sequence number is not displayed. You can run a command to display or hide this field.

6. Timestamp

The timestamp records the time when a syslog is generated so that you can display and check the system event conveniently. FS devices support two syslog timestamp formats: datetime and uptime.

If the device does not have the real time clock (RTC), which is used to record the system absolute time, the device uses its startup time (uptime) as the syslog timestamp by default. If the device has the RTC, the device uses its absolute time (datetime) as the syslog timestamp by default.

The two timestamp formats are described as follows:

- Datetime format

The datetime format is as follows:

```
Mmm dd yyyy hh:mm:ss.msec
```

The following table describes each parameter of the datetime.

Timestamp Parameter	Parameter Name	Description
Mmm	Month	Mmm refers to abbreviation of the current month. The 12 months in a year are written as Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, and Dec.
dd	Day	dd indicates the current date.
yyyy	Year	yyyy indicates the current year, and is not displayed by default.
hh	Hour	hh indicates the current hour.
mm	Minute	mm indicates the current minute.
ss	Second	ss indicates the current second.
msec	Millisecond	msec indicates the current millisecond.

By default, the datetime timestamp displayed in the syslog does not contain the year and millisecond. You can run a command to display or hide the year and millisecond of the datetime timestamp.

- Uptime format

The uptime format is as follows:

```
dd:hh:mm:ss
```

The timestamp string indicates the accumulated days, hours, minutes, and seconds since the system is started.

7. Sysname

This field indicates the name of the device that generates the log so that the log server can identify the host that sends the log. By default, this field is not displayed. You can run a command to display or hide this field.

8. Module

This field indicates the name of the module that generates the log. The module name is an upper-case string of 2 to 20 characters, which contain upper-case letters, digits, or underscores. The module field is mandatory in the log-type information, and optional in the debug-type information.

9. Level

Eight syslog levels from 0 to 7 are defined. The level of syslogs generated by each module is fixed and cannot be modified.

10. Mnemonic

This field indicates the brief information about the log. The mnemonic is an upper-case string of 4 to 32 characters, which may include upper-case letters, digits, or underscore. The mnemonic field is mandatory in the log-type information, and optional in the debug-type information.

11. Content

This field indicates the detailed content of the syslog.

RFC5424 Log Format

The syslog format in the output direction is as follows:

```
<priority>version timestamp sysname MODULE LEVEL MNEMONIC [structured-data] description
```

For example, if you exit configuration mode, the following log is displayed on the Console:

```
<133>1 2013-07-24T12:19:33.130290Z fs SYS 5 CONFIG - Configured from console by console
```

The following describes each field in the log in details:

12. Priority

The priority is calculated using the following formula: Facility x 8 + Level. Level indicates the numerical code of the log level and Facility indicates the numerical code of the facility. When the RFC5424 format is enabled, the default value of the facility field is local0 (16).

13. Version

According to RFC5424, the version is always 1.

14. Timestamp

The timestamp records the time when a syslog is generated so that you can display and check the system event conveniently. FS devices use the following uniformed timestamp format when the RFC5424 logging function is enabled:

```
YYYY-MM-DDTHH:MM:SS.SECFRACZ
```

The following table describes each parameter of the timestamp.

Timestamp Parameter	Description	Remark
YYYY	Year	YYYY indicates the current year.
MM	Month	MM indicates the current month.
DD	Day	DD indicates the current date.
T	Separator	The date must end with "T".
HH	Hour	HH indicates the current hour.
MM	Minute	MM indicates the current minute.
SS	Second	SS indicates the current second.
SECFRAC	Millisecond	SECFRAC indicates the current millisecond (1–6 digits).
Z	End mark	The time must end with "Z".

15. Sysname

This field indicates the name of the device that generates the log so that the log server can identify the host that sends the log.

16. Module

This field indicates the name of the module that generates the log. The module name is an upper-case string of 2 to 20 characters, which contain upper-case letters, digits, or underscores. The module field is mandatory in the log-type information, and optional in the debug-type information.

17. Level

Eight syslog levels from 0 to 7 are defined. The level of syslogs generated by each module is fixed and cannot be modified.

18. Mnemonic

This field indicates the brief information about the log. The mnemonic is an upper-case string of 4 to 32 characters, which contain upper-case letters, digits, or underscores. The Mnemonic field is mandatory in the log-type information, and optional in the debug-type information.

19. Structured-Data

Structured-data introduced in RFC5424 is parsed as a whole string containing parameter information. Each log may contain 0 or multiple parameters. If a parameter is null, replace this parameter with a placeholder (-). The format of this field is as follows:

```
[SD_ID@enterpriseID PARAM-NAME=PARAM-VALUE]
```

The following table describes each parameter of the structured-data field.

Parameter in structured-data	Description	Remarks
SD_ID	Parameter information name	The parameter information name is capitalized, and must be unique in a log.
@	Separator	"@enterpriseID" is added only to the customized parameter information, not to the parameter information defined in RFC5424.
enterpriseID	Enterprise ID	The enterprise ID is maintained by the Internet Assigned Numbers Authority (IANA). FS Networks' enterprise ID is 4881. You can query the enterprise ID on the official website of IANA. http://www.iana.org/assignments/enterprise-numbers
PARAM-NAME	Parameter name	The parameter name is capitalized, and must be unique in the structured-data of a log.
PARAM-VALUE	Parameter value	The parameter value must be enclosed in double quotation marks. Values of the IP address or MAC address must be capitalized, and other types of values are capitalized as required.

20. description

This field indicates the content of the syslog.

Overview

Feature	Description
Logging	Enable or disable the system logging functions.
Syslog Format	Configure the syslog format.
Logging Direction	Configure the parameters to send syslogs in different directions.
Syslog Filtering	Configure parameters of the syslog filtering function.
Featured Logging	Configure parameters of the featured logging function.
Syslog Monitoring	Configure parameters of the syslog monitoring function.

6.3.1 Logging

Enable or disable the logging, log redirection, and log statistics functions.

Related Configuration

Enable Logging

By default, logging is enabled.

Run the **logging on** command to enable logging in global configuration mode. After logging is enabled, logs generated by the system are sent in various directions for the administrator to monitor the performance of the system.

Enabling Log Redirection

By default, log redirection is enabled on the Stacking Status .

Run the **logging rd on** command to enable log redirection in global configuration mode. After log redirection is enabled, logs generated by the standby device or standby supervisor module are redirected to the active device or active supervisor module on the Stacking to facilitate the administrator to manage logs.

Enabling Log Statistics

By default, log statistics is disabled.

Run the **logging count** command to enable log statistics in global configuration mode. After log statistics is enabled, the system records the number of times a log is generated and the last time when the log is generated.

6.3.2 Syslog Format

Configure the syslog format, including the RFC5424 log format, timestamp format, sysname, and sequence number.

Related Configuration

Enabling the RFC5424 Log Format

By default, the RFC5424 log format is disabled.

After the new format (RFC5424 log format) is enabled, the **service sequence-numbers, service sysname, service timestamps**, **service private-syslog**, and **service standard-syslog** that are applicable only to the old format (RFC3164 log format) lose effect and are hidden.

After log format switchover, the outputs of the **show logging** and **show logging config** commands change accordingly.

Configuring the Timestamp Format

By default, the syslog uses the datetime timestamp format, and the timestamp does not contain the year and millisecond.

Run the **service timestamps** command in global configuration mode to use the datetime timestamp format that contains the year and millisecond in the syslog, or change the datetime format to the uptime format.

Adding Sysname to the Syslog

By default, the syslog does not contain sysname.

Run the **service sysname** command in global configuration mode to add sysname to the syslog.

Adding the Sequence Number to the Syslog

By default, the syslog does not contain the sequence number.

Run the **service sequence-numbers** command in global configuration mode to add the sequence number to the syslog.

Enabling the Standard Log Format

By default, logs are displayed in the following format:

```
*timestamp: %module-level-mnemonic: content
```

Run the **service standard-syslog** command in global configuration mode to enable the standard log format and logs are displayed in the following format:

```
timestamp %module-level-mnemonic: content
```

Compared with the default log format, an asterisk (*) is missing in front of the timestamp, and a colon (:) is missing at the end of the timestamp in the standard log format.

Enabling the Private Log Format

By default, logs are displayed in the following format:

```
*timestamp: %module-level-mnemonic: content
```

Run the **service private-syslog** command in global configuration mode to enable the private log format and logs are displayed in the following format:

```
timestamp module-level-mnemonic: content
```

Compared with the default log format, an asterisk (*) is missing in front of the timestamp, a colon (:) is missing at the end of the timestamp, and a percent sign (%) is missing at the end of the module name in the private log format.

6.3.3 Logging Direction

Configure parameters for sending syslogs in different directions, including the Console, monitor terminal, buffer, the log server, and log files.

Related Configuration

Synchronizing User Input with Log Output

By default, this function is disabled.

Run the **logging synchronous** command in line configuration mode to synchronize user input with log output. After this function is enabled, user input will not be interrupted.

Configuring the Log Rate Limit

By default, no log rate limit is configured.

Run the **logging rate-limit** { *number* | **all** *number* | **console** {*number* | **all** *number* } } [**except** [*severity*]] command in global configuration mode to configure the log rate limit.

Configuring the Log Redirection Rate Limit

By default, a maximum of 200 logs are redirected from the standby device to the active device of Stacking per second.

Run the **logging rd rate-limit** *number* [**except** *severity*] command in global configuration mode to configure the log redirection rate limit, that is, the maximum number of logs that are redirected from the standby device to the active device or from the standby supervisor module to the active supervisor module per second.

Configuring the Level of Logs Sent to the Console

By default, the level of logs sent to the Console is debugging (Level 7).

Run the **logging console** [*level*] command in global configuration mode to configure the level of logs that can be sent to the Console.

Sending Logs to the Monitor Terminal

By default, it is not allowed to send logs to the monitor terminal.

Run the **terminal monitor** command in the privileged EXEC mode to send logs to the monitor terminal.

Configuring the Level of Logs Sent to the Monitor Terminal

By default, the level of logs sent to the monitor terminal is debugging (Level 7).

Run the **logging monitor** [*level*] command in global configuration mode to configure the level of logs that can be sent to the monitor terminal.

Writing Logs into the Memory Buffer

By default, logs are written into the memory buffer, and the default level of logs is debugging (Level 7).

Run the **logging buffered** [*buffer-size*] [*level*] command in global configuration mode to configure parameters for writing logs into the memory buffer, including the buffer size and log level.

Sending Logs to the Log Server

By default, logs are not sent to the log server.

Run the **logging server**{ *ip-address* | **ipv6** *ipv6-address* } [**udp-port** *port*] [**vrf** *vrf-name*] command in global configuration mode to send logs to a specified log server.

Configuring the Level of Logs Sent to the Log Server

By default, the level of logs sent to the log server is informational (Level 6).

Run the **logging trap** [*level*] command in global configuration mode to configure the level of logs that can be sent to the log server.

Configuring the Facility Value of Logs Sent to the Log Server

If the RFC5424 log format is disabled, the facility value of logs sent to the log server is local7 (23) by default. If the RFC5424 log format is enabled, the facility value of logs sent to the log server is local0 (16) by default.

Run the **logging facility** *facility-type* command in global configuration mode to configure the facility value of logs sent to the log server.

Configuring the Source Address of Logs Sent to the Log Server

By default, the source address of logs sent to the log server is the IP address of the interface sending logs.

Run the **logging source** [**interface**] *interface-type interface-number* command to configure the source interface of logs. If this source interface is not configured, or the IP address is not configured for this source interface, the source address of logs is the IP address of the interface sending logs.

Run the **logging source** { **ip** *ip-address* | **ipv6** *ipv6-address* } command to configure the source IP address of logs. If this IP address is not configured on the device, the source address of logs is the IP address of the interface sending logs.

Writing Logs into Log Files

By default, logs are not written into log files. After the function of writing logs into log files is enabled, the level of logs written into log files is informational (Level 6) by default.

Run the **logging file** { **flash:***filename* | **usb0**:*filename* } [*max-file-size*] [level] command in global configuration mode to configure parameters for writing logs into log files, including the type of device where the file is stored, file name, file size, and log level.

Configuring the Number of Log Files

By default, the number of log files is 16.

Run the **logging file numbers** *numbers* command in global configuration mode to configure the number of log files.

Configuring the Interval at Which Logs Are Written into Log Files

By default, logs are written into log files at the interval of 3600s (one hour).

Run the **logging flash interval** *seconds* command in global configuration mode to configure the interval at which logs are written into log files.

Configuring the Storage Time of Log Files

By default, the storage time is not configured.

Run the **logging life-time level** *level days* command in global configuration mode to configure the storage time of logs. The administrator can specify different storage days for logs of different levels.

Immediately Writing Logs in the Buffer into Log Files

By default, syslogs are stored in the syslog buffer and then written into log files periodically or when the buffer is full.

Run the **logging flash flush** command in global configuration mode to immediately write logs in the buffer into log files so that you can collect logs conveniently.

6.3.4 Syslog Filtering

By default, logs generated by the system are sent in all directions.

Working Principle

Filtering Direction

Five log filtering directions are defined:

- **buffer**: Filters out logs sent to the log buffer, that is, logs displayed by the **show logging** command.
- **file**: Filters out logs written into log files.
- **server**: Filters out logs sent to the log server.
- **terminal**: Filters out logs sent to the Console and monitor terminal (including Telnet and SSH).

The four filtering directions can be used either in combinations to filter out logs sent in various directions, or separately to filter out logs sent in a single direction.

Filtering Mode

Two filtering modes are available:

- **contains-only**: Indicates that only logs that contain keywords specified in the filtering rules are output. You may be interested in only a specified type of logs. In this case, you can apply the contains-only mode on the device to display only logs that match filtering rules on the terminal, helping you check whether any event occurs.
- **filter-only**: Indicates that logs that contain keywords specified in the filtering rules are filtered out and will not be output. If a module generates too many logs, spamming may occur on the terminal interface. If you do not care about this type of logs, you can apply the filter-only mode and configure related filtering rules to filter out logs that may cause spamming.

The two filtering modes are mutually exclusive, that is, you can configure only one filtering mode at a time.

Filter Rule

Two filtering rules are available:

- **exact-match**: If exact-match is selected, you must select all the three filtering options (module, level, and mnemonic). If you want to filter out a specified log, use the exact-match filtering rule.
- **single-match**: If exact-match is selected, you only need to select one of the three filtering options (module, level, and mnemonic). If you want to filter out a specified type of logs, use the single-match filtering rule.

If the same module, level, or mnemonic is configured in both the single-match and exact-match rules, the single-match rule prevails over the exact-match rule.

Related Configuration

Configuring the Log Filtering Direction

By default, the log filtering direction is all, that is, logs sent in all directions are filtered.

Run the **logging filter direction** { **all** | **buffer** | **file** | **server** | **terminal** } command in global configuration mode to configure the log filtering direction to filter out logs in the specified directions.

Configuring the Log Filtering Mode

By default, the log filtering mode is filter-only.

Run the **logging filter type** { **contains-only** | **filter-only** } command in global configuration mode to configure the log filtering mode.

Configuring the Log Filtering Rule

By default, no log filtering rule is configured on a device, that is, logs are not filtered out.

Run the **logging filter rule exact-match module** *module-name* **mnemonic** *mnemonic-name* **level** *level* command in global configuration mode to configure the exact-match rule.

Run the **logging filter rule single-match** { **level** *level* | **mnemonic** *mnemonic-name* | **module** *module-name* } command in global configuration mode to configure the single-match rule.

6.3.5 Syslog Monitoring

After syslog monitoring is enabled, the system monitors the access attempts of users and generates the related logs.

Working Principle

After logging of login/exit attempts is enabled, the system records the access attempts of users. The log contains user name and source address.

After logging of operations is enabled, the system records changes in device configurations, The log contains user name, source address, and operation.

Related Configuration

Enabling Logging of Login or Exit Attempts

By default, a device does not generate logs when users access or exit the device.

Run the **logging userinfo** command in global configuration mode to enable logging of login/exit attempts. After this function is enabled, the device displays logs when users access the devices through Telnet, SSH, or HTTP so that the administrator can monitor the device connections.

Enabling Logging of Operations

By default, a device does not generate logs when users modify device configurations.

Run the **logging userinfo command-log** command in global configuration mode to enable logging of operations. After this function is enabled, the system displays related logs to notify the administrator of configuration changes.

6.4 Configuration

Configuration	Description and Command	
Configuring Syslog Format	(Optional) It is used to configure the syslog format.	
	service timestamps [*message-type* [**uptime**| **datetime** [**msec**] [**year**]]]	Configures the timestamp format of syslogs.
	service sysname	Adds the sysname to the syslog.
	service sequence-numbers	Adds the sequence number to the syslog.
	service standard-syslog	Enables the standard syslog format.
	service private-syslog	Enables the private syslog format.
	service log-format rfc5424	Enables the RFC5424 syslog format.
Sending Syslogs to the Console	(Optional) It is used to configure parameters for sending syslogs to the Console.	
	logging on	Enables logging.
	logging count	Enables log statistics.
	logging console [*level*]	Configures the level of logs displayed on the Console.
	logging rate-limit { *number* | **all** *number* | **console** {*number* | **all** *number* } } [**except** [*severity*]]	Configures the log rate limit.

Configuration	Description and Command	
Sending Syslogs to the Monitor Terminal	(Optional) It is used to configure parameters for sending syslogs to the monitor terminal.	
	terminal monitor	Enables the monitor terminal to display logs.
	logging monitor [*level*]	Configures the level of logs displayed on the monitor terminal.
Writing Syslogs into the Memory Buffer	(Optional) It is used to configure parameters for writing syslogs into the memory buffer.	
	logging buffered [*buffer-size*] [*level*]	Configures parameters for writing syslogs into the memory buffer, including the buffer size and log level.
Sending Syslogs to the Log Server	(Optional) It is used to configure parameters for sending syslogs to the log server.	
	logging server [**oob**] { *ip-address* | **ipv6** *ipv6-address* } [**udp-port** *port*] [**vrf** *vrf-name*]	Sends logs to a specified log server.
	logging trap [*level*]	Configures the level of logs sent to the log server.
	logging facility *facility-type*	Configures the facility value of logs sent to the log server.
	logging source [**interface**] *interface-type interface-number*	Configures the source interface of logs sent to the log server.
	logging source { **ip** *ip-address* | **ipv6** *ipv6-address* }	Configures the source address of logs sent to the log server.
Writing Syslogs into Log Files	(Optional) It is used to configure parameters for writing syslogs into a file.	
	logging file { **flash:***filename* | **usb0:***filename* } [*max-file-size*] [level]	Configures parameters for writing syslogs into a file, including the file storage type, file name, file size, and log level.
	logging file numbers *numbers*	Configures the number of files which logs are written into. The default value is 16.
	logging flash interval *seconds*	Configures the interval at which logs are written into log files. The default value is 3600.
	logging life-time level *level days*	Configures the storage time of log files.
Configuring Syslog Filtering	(Optional) It is used to enable the syslog filtering function.	
	logging filter direction { **all** | **buffer** | **file** | **server** | **terminal** }	Configures the log filtering direction.
	logging filter type { **contains-only** | **filter-only** }	Configures the log filtering mode.
	logging filter rule exact-match module *module-name* **mnemonic** *mnemonic-name* **level** *level*	Configures the exact-match filtering rule.

Configuration	Description and Command	
	logging filter rule single-match { **level** *level* | **mnemonic** *mnemonic-name* | **module** *module-name* }	Configures the single-match filtering rule.
Configuring Level-based Logging	(Optional) It is used to configure logging policies to send the syslogs based on module and severity level .	
	logging policy module *module-name* [**not-lesser-than**] *level* **direction** { **all** | **server** | **file** | **console** | **monitor** | **buffer** }	Sends logs to different destinations by module and severity level
Configuring Delayed Logging	(Optional) It is used to enable the delayed logging function.	
	logging delay-send terminal	Enables delayed display of logs on the Console and remote terminal.
	logging delay-send file flash:*filename*	Configures the name of the file on the local device where logs are buffered.
	logging delay-send interval *seconds*	Configures the interval at which logs are sent to the log server.
	logging delay-send server { [**oob**] *ip-address* | **ipv6** *ipv6-address* } [**vrf** *vrf-name*] **mode** { **ftp user** *username* **password** [**0** | **7**] *password* | **tftp** }	Configures the server address and delayed logging mode.
Configuring Periodical Logging	(Optional) It is used to enable the periodical logging function.	
	logging statistic enable	Enables the periodical logging function .
	logging statistic terminal	Enables periodical display of logs on the Console and remote terminal.
	logging statistic mnemonic *mnemonic* **interval** *minutes*	Configures the interval at which logs of a performance statistic object are sent to the server .
Configuring Syslog Redirection	(Optional) It is used to enable the log redirection function.	
	logging rd on	Enables the log redirection function.
	logging rd rate-limit *number* [**except** *severity*]	Configures the log redirection rate limit.
Configuring Syslog Monitoring	(Optional) It is used to configure parameters of the syslog monitoring function .	
	logging userinfo	Enables logging of login/exit attempts.
	logging userinfo command-log	Enables logging of operations.
Synchronizing User Input with Log Output	(Optional) It is used to synchronize the user input with log output.	
	logging synchronous	Synchronizes user input with log output.

6.4.1 Configuring Syslog Format

Configuration Effect

- Configure the format of syslogs.

Notes

RFC3164 Log Format

- If the device does not have the real time clock (RTC), which is used to record the system absolute time, the device uses its startup time (uptime) as the syslog timestamp by default. If the device has the RTC, the device uses its absolute time (datetime) as the syslog timestamp by default.
- The log sequence number is a 6-digit integer. Each time a log is generated, the sequence number increases by one. Each time the sequence number increases from 000000 to 1,000,000, or reaches 2^32, the sequence number starts from 000000 again.

RFC5424 Log Format

- After the RFC5424 log format is enabled, the timestamp is uniform.
- In the RFC5424 log format, the timestamp may or may not contain the time zone. Currently, only the timestamp without the time zone is supported.

Configuration Steps

Configuring the Timestamp Format of Syslogs

- (Optional) By default, the datetime timestamp format is used.
- Unless otherwise specified, perform this configuration on the device to configure the timestamp format.

Adding the Sysname to the Syslog

- (Optional) By default, the syslog does not contain the sysname.
- Unless otherwise specified, perform this configuration on the device to add the sysname to the syslog.

Adding the Sequence Number to the Syslog

- (Optional) By default, the syslog does not contain the sequence number.
- Unless otherwise specified, perform this configuration on the device to add the sequence number to the syslog.

Enabling the Standard Log Format

- (Optional) By default, the default log format is used.
- Unless otherwise specified, perform this configuration on the device to enable the standard log format.

Enabling the Private Log Format

- (Optional) By default, the default log format is used.
- Unless otherwise specified, perform this configuration on the device to enable the private log format.

Enabling the RFC5424 Log Format

- (Optional) By default, the RFC5424 log format is disabled.
- Unless otherwise specified, perform this configuration on the device to enable the RFC5424 log format.

Verification

- Generate a syslog, and check the log format.

Related Commands

Configuring the Timestamp Format of Syslogs

Command	**service timestamps** [*message-type* [**uptime** | **datetime** [**msec**] [**year**]]]
Parameter Description	*message-type*: Indicates the log type. There are two log types: log and debug. **uptime**: Indicates the device startup time in the format of dd:hh:mm:ss, for example, 07:00:10:41. **datetime**: Indicates the current device time in the format of MM DD hh:mm:ss, for example, Jul 27 16:53:07. **msec**: Indicates that the current device time contains millisecond. **year**: Indicates that the current device time contains year.
Command Mode	Global configuration mode
Configuration Usage	Two syslog timestamp formats are available, namely, uptime and datetime. You can select a timestamp format as required.

Adding the Sysname to the Syslog

Command	**service sysname**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	This command is used to add the sysname to the log to enable you to learn about the device that sends syslogs to the server.

Adding the Sequence Number to the Syslog

Command	**service sequence-numbers**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	This command is used to add the sequence number to the log. The sequence number starts from 1. After the sequence number is added, you can learn clearly whether any log is lost and the generation sequence of logs.

Enabling the Standard Syslog Format

Command	**service standard-syslog**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	By default, logs are displayed in the following format (default format): *timestamp: %module-level-mnemonic: content If the standard syslog format is enabled, logs are displayed in the following format: timestamp %module-level-mnemonic: content Compared with the default format, an asterisk (*) is missing in front of the timestamp, and a colon (:) is missing at the end of the timestamp in the standard log format.

Enabling the Private Syslog Format

Command	**service private-syslog**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	By default, logs are displayed in the following format (default format): *timestamp: %module-level-mnemonic: content If the private syslog format is enabled, logs are displayed in the following format: timestamp module-level-mnemonic: content Compared with the default format, an asterisk (*) is missing in front of the timestamp, a colon (:) is missing at the end of the timestamp, and a percent sign (%) is missing in front of the module name in the private log format.

Enabling the RFC5424 Syslog Format

Command	**service log-format rfc5424**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	After the new format (RFC5424 log format) is enabled, the **service sequence-numbers**, **service sysname**, **service timestamps**, **service private-syslog**, and **service standard-syslog** commands that are applicable only to the old format (RFC3164 log format) loss effect and are hidden. After log format switchover, the outputs of the **show logging** and **show logging config** commands change accordingly.

Configuration Example

Enabling the RFC3164 Log Format

Scenario	It is required to configure the timestamp format as follows: 1. Enable the RFC3164 format. 2. Change the timestamp format to datetime and add the millisecond and year to the timestamp. 3. Add the sysname to the log. 4. Add the sequence number to the log.
Configuration Steps	● Configure the syslog format.
	FS# configure terminal FS(config)# no service log-format rfc5424 FS(config)# service timestamps log datetime year msec FS(config)# service timestamps debug datetime year msec FS(config)# service sysname FS(config)# service sequence-numbers
Verification	After the timestamp format is configured, verify that new syslogs are displayed in the RFC3164 format. ● Run the **show logging config** command to display the configuration. ● Enter or exit global configuration mode to generate a new log, and check the format of the timestamp in the new log.
	FS(config)#exit 001302: *Jun 14 2013 19:01:40.293: FS %SYS-5-CONFIG_I: Configured from console by admin on console FS#show logging config Syslog logging: enabled Console logging: level informational, 1306 messages logged Monitor logging: level informational, 0 messages logged Buffer logging: level informational, 1306 messages logged File logging: level informational, 121 messages logged File name:syslog_test.txt, size 128 Kbytes, have written 5 files Standard format:false Timestamp debug messages: datetime Timestamp log messages: datetime Sequence-number log messages: enable Sysname log messages: enable Count log messages: enable

	Trap logging: level informational, 121 message lines logged,0 fail

Enabling the RFC5424 Log Format

Scenario	It is required to enable the RFC5424 format.
Configuration Steps	● Configure the syslog format.
	FS# configure terminal FS(config)# service log-format rfc5424
Verification	Verify that new syslogs are displayed in the RFC5424 format. ● Run the **show logging config** command to display the configuration. ● Enter or exit global configuration mode to generate a new log, and check the format of the new log.
	FS(config)#exit <133>1 2013-07-24T12:19:33.130290Z fs SYS 5 CONFIG - Configured from console by console FS#show logging config Syslog logging: enabled Console logging: level debugging, 4740 messages logged Monitor logging: level debugging, 0 messages logged Buffer logging: level debugging, 4745 messages logged Statistic log messages: disable Statistic log messages to terminal: disable Delay-send file name:syslog_ftp_server, Current write index:3, Current send index:3, Cycle:10 seconds Count log messages: enable Trap logging: level informational, 2641 message lines logged,4155 fail logging to 192.168.23.89 logging to 2000::1

6.4.2 Sending Syslogs to the Console

Configuration Effect

- Send syslogs to the Console to facilitate the administrator to monitor the performance of the system.

Notes

- If too many syslogs are generated, you can limit the log rate to reduce the number of logs displayed on the Console.

Configuration Steps

Enabling Logging

- (Optional) By default, the logging function is enabled.

Enabling Log Statistics

- (Optional) By default, log statistics is disabled.
- Unless otherwise specified, perform this configuration on the device to enable log statistics.

Configuring the Level of Logs Displayed on the Console

- (Optional) By default, the level of logs displayed on the Console is debugging (Level 7).
- Unless otherwise specified, perform this configuration on the device to configure the level of logs displayed on the Console.

Configuring the Log Rate Limit

- (Optional) By default, the no rate limit is configured.
- Unless otherwise specified, perform this configuration on the device to limit the log rate.

Verification

- Run the **show logging config** command to display the level of logs displayed on the Console.

Related Commands

Enabling Logging

Command	**logging on**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	By default, logging is enabled. Do not disable logging in general cases. If too many syslogs are generated, you can configure log levels to reduce the number of logs.

Enabling Log Statistics

Command	**logging count**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	By default, log statistics is disabled. If log statistics is enabled, syslogs will be classified and counted. The system records the number of times a log is generated and the last time when the log is generated.

Configuring the Level of Logs Displayed on the Console

Command	**logging console** [*level*]
Parameter Description	*level*: Indicates the log level.
Command Mode	Global configuration mode
Configuration Usage	By default, the level of logs displayed on the Console is debugging (Level 7). You can run the **show logging config** command in privileged EXEC mode to display the level of logs displayed on the Console.

Configuring the Log Rate Limit

Command	**logging rate-limit** { *number* | **all** *number* | **console** {*number* | **all** *number* } } [**except** [*severity*]]
Parameter Description	*number*: Indicates the maximum number of logs processed per second. The value ranges from 1 to 10,000. **all**: Indicates that rate limit is applied to all logs ranging from Level 0 to Level 7. **console**: Indicates the number of logs displayed on the Console per second. **except** *severity*: Rate limit is not applied to logs with a level equaling to or lower than the specified severity level. By default, the severity level is error (Level 3), that is, rate limit is not applied to logs of Level 3 or lower.
Command Mode	Global configuration mode
Configuration Usage	By default, no rate limit is configured.

Configuration Example

Sending Syslogs to the Console

Scenario	It is required to configure the function of displaying syslogs on the Console as follows: 1. Enable log statistics. 2. Set the level of logs that can be displayed on the Console to informational (Level 6). 3. Set the log rate limit to 50.
Configuration Steps	● Configure parameters for displaying syslogs on the Console.
	FS# configure terminal FS(config)# logging count FS(config)# logging console informational FS(config)# logging rate-limit console 50
Verification	● Run the **show logging config** command to display the configuration.
	FS(config)#show logging config Syslog logging: enabled Console logging: level informational, 1303 messages logged

Scenario	It is required to configure the function of displaying syslogs on the Console as follows: 1. Enable log statistics. 2. Set the level of logs that can be displayed on the Console to informational (Level 6). 3. Set the log rate limit to 50.
Configuration Steps	● Configure parameters for displaying syslogs on the Console.
	FS# configure terminal FS(config)# logging count FS(config)# logging console informational FS(config)# logging rate-limit console 50
Verification	● Run the **show logging config** command to display the configuration.
	Monitor logging: level debugging, 0 messages logged Buffer logging: level debugging, 1303 messages logged File logging: level informational, 118 messages logged File name:syslog_test.txt, size 128 Kbytes, have written 5 files Standard format:false Timestamp debug messages: datetime Timestamp log messages: datetime Sequence-number log messages: enable Sysname log messages: enable Count log messages: enable Trap logging: level informational, 118 message lines logged,0 fail

6.4.3 Sending Syslogs to the Monitor Terminal

Configuration Effect

- Send syslogs to a remote monitor terminal to facilitate the administrator to monitor the performance of the system.

Notes

- If too many syslogs are generated, you can limit the log rate to reduce the number of logs displayed on the monitor terminal.
- By default, the current monitor terminal is not allowed to display logs after you access the device remotely. You need to manually run the **terminal monitor** command to allow the current monitor terminal to display logs.

Configuration Steps

Allowing the Monitor Terminal to Display Logs

- (Mandatory) By default, the monitor terminal is not allowed to display logs.

- Unless otherwise specified, perform this operation on every monitor terminal connected to the device.

Configuring the Level of Logs Displayed on the Monitor Terminal

- (Optional) By default, the level of logs displayed on the monitor terminal is debugging (Level 7).
- Unless otherwise specified, perform this configuration on the device to configure the level of logs displayed on the monitor terminal.

Verification

- Run the **show logging config** command to display the level of logs displayed on the monitor terminal.

Related Commands

Allowing the Monitor Terminal to Display Logs

Command	**terminal monitor**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Configuration Usage	By default, the current monitor terminal is not allowed to display logs after you access the device remotely. You need to manually run the **terminal monitor** command to allow the current monitor terminal to display logs.

Configuring the Level of Logs Displayed on the Monitor Terminal

Command	**logging monitor** [*level*]
Parameter Description	*level*: Indicates the log level.
Command Mode	Global configuration mode
Configuration Usage	By default, the level of logs displayed on the monitor terminal is debugging (Level 7). You can run the **show logging config** command in privileged EXEC mode to display the level of logs displayed on the monitor terminal.

Configuration Example

Sending Syslogs to the Monitor Terminal

Scenario	It is required to configure the function of displaying syslogs on the monitor terminal as follows: 1. Display logs on the monitor terminal. 2. Set the level of logs that can be displayed on the monitor terminal to informational (Level 6).
Configuration Steps	● Configure parameters for displaying syslogs on the monitor terminal.

	FS# configure terminal FS(config)# logging monitor informational FS(config)# line vty 0 4 FS(config-line)# monitor
Verification	● Run the **show logging config** command to display the configuration.
	FS#show logging config Syslog logging: enabled Console logging: level informational, 1304 messages logged Monitor logging: level informational, 0 messages logged Buffer logging: level debugging, 1304 messages logged File logging: level informational, 119 messages logged File name:syslog_test.txt, size 128 Kbytes, have written 5 files Standard format:false Timestamp debug messages: datetime Timestamp log messages: datetime Sequence-number log messages: enable Sysname log messages: enable Count log messages: enable Trap logging: level informational, 119 message lines logged,0 fail

Common Errors

- To disable this function, run the **terminal no monitor** command, instead of the **no terminal monitor** command.

6.4.4 Writing Syslogs into the Memory Buffer

Configuration Effect

- Write syslogs into the memory buffer so that the administrator can view recent syslogs by running the **show logging** command.

Notes

- If the buffer is full, old logs will be overwritten by new logs that are written into the memory buffer.

Configuration Steps

Writing Logs into the Memory Buffer

- (Optional) By default, the system writes logs into the memory buffer, and the default level of logs is debugging (Level 7).
- Unless otherwise specified, perform this configuration on the device to write logs into the memory buffer.

Verification

- Run the **show logging config** command to display the level of logs written into the memory buffer.
- Run the **show logging** command to display the level of logs written into the memory buffer.

Related Commands

Writing Logs into the Memory Buffer

Command	**logging buffered** [*buffer-size*] [*level*]
Parameter Description	*buffer-size*: Indicates the size of the memory buffer. *level*: Indicates the level of logs that can be written into the memory buffer.
Command Mode	Global configuration mode
Configuration Usage	By default, the level of logs written into the memory buffer is debugging (Level 7). Run the **show logging** command in privileged EXEC mode to display the level of logs written into the memory buffer and the buffer size.

Configuration Example

Writing Syslogs into the Memory Buffer

Scenario	It is required to configure the function of writing syslogs into the memory buffer as follows: 1. Set the log buffer size to 128 KB (131,072 bytes). 2. Set the information level of logs that can be written into the memory buffer to informational (Level 6).
Configuration Steps	● Configure parameters for writing syslogs into the memory buffer.
	FS# configure terminal FS(config)# logging buffered 131072 informational
Verification	● Run the **show logging config** command to display the configuration and recent syslogs.
	FS#show logging Syslog logging: enabled Console logging: level informational, 1306 messages logged Monitor logging: level informational, 0 messages logged Buffer logging: level informational, 1306 messages logged File logging: level informational, 121 messages logged File name:syslog_test.txt, size 128 Kbytes, have written 5 files Standard format:false Timestamp debug messages: datetime Timestamp log messages: datetime

Scenario	It is required to configure the function of writing syslogs into the memory buffer as follows: 1. Set the log buffer size to 128 KB (131,072 bytes). 2. Set the information level of logs that can be written into the memory buffer to informational (Level 6).
Configuration Steps	● Configure parameters for writing syslogs into the memory buffer.
	FS# configure terminal FS(config)# logging buffered 131072 informational
Verification	● Run the **show logging config** command to display the configuration and recent syslogs.
	Sequence-number log messages: enable Sysname log messages: enable Count log messages: enable Trap logging: level informational, 121 message lines logged,0 fail Log Buffer (Total 131072 Bytes): have written 4200 001301: *Jun 14 2013 19:01:09.488: FS %SYS-5-CONFIG_I: Configured from console by admin on console 001302: *Jun 14 2013 19:01:40.293: FS %SYS-5-CONFIG_I: Configured from console by admin on console //Logs displayed are subject to the actual output of the **show logging** command.

6.4.5 Sending Syslogs to the Log Server

Configuration Effect

- Send syslogs to the log server to facilitate the administrator to monitor logs on the server.

Notes

- If the device has a MGMT interface and is connected to the log server through the MGMT interface, you must add the **oob** option (indicating that syslogs are sent to the log server through the MGMT interface) when configuring the **logging server** command.
- To send logs to the log server, you must add the timestamp and sequence number to logs. Otherwise, the logs are not sent to the log server.

Configuration Steps

Sending Logs to a Specified Log Server

- (Mandatory) By default, syslogs are not sent to any log server.
- Unless otherwise specified, perform this configuration on every device.

Configuring the Level of Logs Sent to the Log Server

- (Optional) By default, the level of logs sent to the log server is informational (Level 6).
- Unless otherwise specified, perform this configuration on the device to configure the level of logs sent to the log server.

Configuring the Facility Value of Logs Sent to the Log Server

- (Optional) If the RFC5424 format is disabled, the facility value of logs sent to the log server is local7 (23) by default. If the RFC5424 format is enabled, the facility value of logs sent to the log server is local0 (16) by default.
- Unless otherwise specified, perform this configuration on the device to configure the facility value of logs sent to the log server.

Configuring the Source Interface of Logs Sent to the Log Server

- (Optional) By default, the source interface of logs sent to the log server is the interface sending the logs.
- Unless otherwise specified, perform this configuration on the device to configure the source interface of logs sent to the log server.

Configuring the Source Address of Logs Sent to the Log Server

- (Optional) By default, the source address of logs sent to the log server is the IP address of the interface sending the logs.
- Unless otherwise specified, perform this configuration on the device to configure the source address of logs sent to the log server.

Verification

- Run the **show logging config** command to display the configurations related to the log server.

Related Commands

Sending Logs to a Specified Log Server

Command	**logging server** [**oob**] { *ip-address* | **ipv6** *ipv6-address* } [**udp-port** *port*] [**vrf** *vrf-name*] Or **logging** { *ip-address* | **ipv6** *ipv6-address* } [**udp-prot** *port*] [**vrf** *vrf-name*]
Parameter Description	**oob**: Indicates that logs are sent to the log server through the MGMT interface. *ip-address*: Specifies the IP address of the host that receives logs. **ipv6** *ipv6-address*: Specifies the IPv6 address of the host that receives logs. **vrf** *vrf-name*: Specifies the VPN routing and forwarding (VRF) instance connected to the log server. **udp-port** *port*: Specifies the port ID of the log server. The default port ID is 514.
Command Mode	Global configuration mode
Configuration Usage	This command is used to specify the address of the log server that receives logs. You can specify multiple log servers, and logs will be sent simultaneously to all these log servers. You can configure up to five log servers on a FS product.

Configuring the Level of Logs Sent to the Log Server

Command	**logging trap** [*level*]
Parameter Description	*level*: Indicates the log level.
Command Mode	Global configuration mode
Configuration Usage	By default, the level of logs sent to the log server is informational (Level 6). You can run the **show logging config** command in privileged EXEC mode to display the level of logs sent to the log

	server.

Configuring the Facility Value of Logs Sent to the Log Server

Command	**logging facility** *facility-type*
Parameter Description	*facility-type*: Indicates the facility value of logs.
Command Mode	Global configuration mode
Configuration Usage	If the RFC5424 format is disabled, the facility value of logs sent to the server is local7 (23) by default. If the RFC5424 format is enabled, the facility value of logs sent to the server is local0 (16) by default.

Configuring the Source Interface of Logs Sent to the Log Server

Command	**logging source** [**interface**] *interface-type interface-number*
Parameter Description	*interface-type*: Indicates the interface type. *interface-number*: Indicates the interface number.
Command Mode	Global configuration mode
Configuration Usage	By default, the source interface of logs sent to the log server is the interface sending the logs. To facilitate management, you can use this command to set the source interface of all logs to an interface so that the administrator can identify the device that sends the logs based on the unique address.

Configuring the Source Address of Logs Sent to the Log Server

Command	**logging source** { **ip** *ip-address* | **ipv6** *ipv6-address* }
Parameter Description	**ip** *ip-address*: Specifies the source IPv4 address of logs sent to the IPv4 log server. **ipv6** *ipv6-address*: Specifies the source IPv6 address of logs sent to the IPv6 log server.
Command Mode	Global configuration mode
Configuration Usage	By default, the source IP address of logs sent to the log server is the IP address of the interface sending the logs. To facilitate management, you can use this command to set the source IP address of all logs to the IP address of an interface so that the administrator can identify the device that sends the logs based on the unique address..

Configuration Example

Sending Syslogs to the Log Server

Scenario	It is required to configure the function of sending syslogs to the log server as follows: 1. Set the IPv4 address of the log server to 10.1.1.100. 2. Set the level of logs that can be sent to the log server to debugging (Level 7). 3. Set the source interface to Loopback 0.

Configuration Steps	● Configure parameters for sending syslogs to the log server.
	FS# configure terminal FS(config)# logging server 10.1.1.100 FS(config)# logging trap debugging FS(config)# logging source interface Loopback 0
Verification	● Run the **show logging config** command to display the configuration.
	FS#show logging config Syslog logging: enabled Console logging: level informational, 1307 messages logged Monitor logging: level informational, 0 messages logged Buffer logging: level informational, 1307 messages logged File logging: level informational, 122 messages logged File name:syslog_test.txt, size 128 Kbytes, have written 5 files Standard format:false Timestamp debug messages: datetime Timestamp log messages: datetime Sequence-number log messages: enable Sysname log messages: enable Count log messages: enable Trap logging: level debugging, 122 message lines logged,0 fail logging to 10.1.1.100

6.4.6 Writing Syslogs into Log Files

Configuration Effect

- Write syslogs into log files at the specified interval so that the administrator can view history logs anytime on the local device.

Notes

- Sylsogs are not immediately written into log files. They are first buffered in the memory buffer, and then written into log files either periodically (at the interval of one hour by default) or when the buffer is full.

Configuration Steps

Writing Logs into Log Files

- (Mandatory) By default, syslogs are not written to any log file.
- Unless otherwise specified, perform this configuration on every device.

Configuring the Number of Log Files

- (Optional) By default, syslogs are written to 16 log files.
- Unless otherwise specified, perform this configuration on the device to configure the number of files which logs are written into.

Configuring the Interval at Which Logs Are Written into Log Files

- (Optional) By default, syslogs are written to log files every hour.
- Unless otherwise specified, perform this configuration on the device to configure the interval at which logs are written into log files.

Configuring the Storage Time of Log Files

- (Optional) By default, no storage time is configured.
- Unless otherwise specified, perform this configuration on the device to configure the storage time of log files.

Immediately Writing Logs in the Buffer into Log Files

- (Optional) By default, syslogs are stored in the buffer and then written into log files periodically or when the buffer is full.
- Unless otherwise specified, perform this configuration to write logs in the buffer into log files immediately. This command takes effect only once after it is configured.

Verification

- Run the **show logging config** command to display the configurations related to the log server.

Related Commands

Writing Logs into Log Files

Command	**logging file** { **flash:***filename* | **usb0**:*filename* } [*max-file-size*] [*level*]
Parameter Description	**flash**: Indicates that log files will be stored on the extended Flash. **usb0**: Indicates that log files will be stored on USB 0. This option is supported only when the device has one USB port and a USB flash drive is inserted into the USB port. *filename*: Indicates the log file name, which does not contain a file name extension. The file name extension is always txt. *max-file-size*: Indicates the maximum size of a log file. The value ranges from 128 KB to 6 MB. The default value is 128 KB. *level*: Indicates the level of logs that can be written into a log file.
Command Mode	Global configuration mode
Configuration Usage	This command is used to create a log file with the specified file name on the specified file storage device. The file size increases with the amount of logs, but cannot exceed the configured maximum size. If not specified, the maximum size of a log file is 128 KB by default. After this command is configured, the system saves logs to log files. A log file name does not contain any file name extension. The file name extension is always txt, which cannot be changed. After this command is configured, logs will be written into log files every hour. If you run the **logging flie**

	flash:syslog command, a total of 16 log files will be created, namely, **syslog.txt**, **syslog_1.txt**, **syslog_2.txt**, …, **syslog_14.txt**, and **syslog_15.txt**. Logs are written into the 16 log files in sequence. For example, the system writes logs into **syslog_1.txt** after **syslog.txt** is full. When **syslog_15.txt** is full, logs are written into **syslog.txt** again,

Configuring the Number of Log Files

Command	**logging file numbers** *numbers*
Parameter Description	*numbers*: Indicates the number of log files. The value ranges from 2 to 32.
Command Mode	Global configuration mode
Configuration Usage	This command is used to configure the number of log files. If the number of log files is modified, the system will not delete the log files that have been generated. Therefore, you need to manually delete the existing log files to save the space of the extended flash. (Before deleting existing log files, you can transfer these log files to an external server through TFTP.) For example, after the function of writing logs into log files is enabled, 16 log files will be created by default. If the device has generated 16 log files and you change the number of log files to 2, new logs will be written into **syslog.txt** and **syslog_1.txt** by turns. The existing log files from **syslog_2.txt** to **syslog_15.txt** will be preserved. You can manually delete these log files.

Configuring the Interval at Which Logs Are Written into Log Files

Command	**logging flash interval** *seconds*
Parameter Description	*seconds*: Indicates the interval at which logs are written into log files. The value ranges from 1s to 51,840s.
Command Mode	Global configuration mode
Configuration Usage	This command is used to configure the interval at which logs are written into log files. The countdown starts after the command is configured.

Configuring the Storage Time of Log Files

Command	**logging life-time level** *level days*
Parameter Description	*level*: Indicates the log level. *days*: Indicates the storage time of log files. The unit is day. The storage time is not less than seven days.
Command Mode	Global configuration mode
Configuration Usage	After the log storage time is configured, the system writes logs of the same level that are generated in the same day into the same log file. The log file is named **yyyy-mm-dd_filename_level.txt**, where **yyyy-mm-dd** is the absolute time of the day when the logs are generated, **filename** is the log file named configured by the **logging file flash** command, and **level** is the log level. After you specify the storage time for logs of a certain level, the system deletes the logs after the storage time expires. Currently, the storage time ranges from 7days to 365 days.

	If the log storage time is not configured, logs are stored based on the file size to ensure compatibility with old configuration commands.

Immediately Writing Logs in the Buffer into Log Files

Command	**logging flash flush**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	After this command is configured, syslogs are stored in the buffer and then written into log files periodically or when the buffer is full. You can run this command to immediately write logs into log files. The **logging flash flush** command takes effect once after it is configured. That is, after this command is configured, logs in the buffer are immediately written to log files.

Configuration Example

Writing Syslogs into Log Files

Scenario	It is required to configure the function of writing syslogs into log files as follows: 1. Set the log file name to syslog. 2. Set the level of logs sent to the Console to debugging (Level 7). 3. Set the interval at which device logs are written into files to 10 minutes (600s).
Configuration Steps	● Configure parameters for writing syslogs into log files.
	FS# configure terminal FS(config)# logging file flash:syslog debugging FS(config)# logging flash interval 600
Verification	● Run the **show logging config** command to display the configuration.
	FS(config)#show logging config Syslog logging: enabled Console logging: level informational, 1307 messages logged Monitor logging: level informational, 0 messages logged Buffer logging: level informational, 1307 messages logged File logging: level debugging, 122 messages logged File name:syslog.txt, size 128 Kbytes, have written 1 files Standard format:false Timestamp debug messages: datetime

Scenario	It is required to configure the function of writing syslogs into log files as follows: 1. Set the log file name to syslog. 2. Set the level of logs sent to the Console to debugging (Level 7). 3. Set the interval at which device logs are written into files to 10 minutes (600s).
Configuration Steps	● Configure parameters for writing syslogs into log files.
	FS# configure terminal FS(config)# logging file flash:syslog debugging FS(config)# logging flash interval 600
Verification	● Run the **show logging config** command to display the configuration.
	Timestamp log messages: datetime Sequence-number log messages: enable Sysname log messages: enable Count log messages: enable Trap logging: level debugging, 122 message lines logged,0 fail logging to 10.1.1.100

6.4.7 Configuring Syslog Filtering

Configuration Effect

- Filter out a specified type of syslogs if the administrator does not want to display these syslogs.
- By default, logs generated by all modules are displayed on the Console or other terminals. You can configure log filtering rules to display only desired logs.

Notes

- Two filtering modes are available: contains-only and filter-only. You can configure only one filtering mode at a time.
- If the same module, level, or mnemonic is configured in both the single-match and exact-match rules, the single-match rule prevails over the exact-match rule.

Configuration Steps

Configuring the Log Filtering Direction

- (Optional) By default, the filtering direction is all, that is, all logs are filtered out.
- Unless otherwise specified, perform this configuration on the device to configure the log filtering direction.

Configuring the Log Filtering Mode

- (Optional) By default, the log filtering mode is filter-only.
- Unless otherwise specified, perform this configuration on the device to configure the log filtering mode.

Configuring the Log Filtering Rule

- (Mandatory) By default, no filtering rule is configured.
- Unless otherwise specified, perform this configuration on the device to configure the log filtering rule.

Verification

- Run the **show running** command to display the configuration.

Related Commands

Configuring the Log Filtering Direction

Command	**logging filter direction** { **all** | **buffer** | **file** | **server** | **terminal** }
Parameter Description	**all**: Filters out all logs. **buffer**: Filters out logs sent to the log buffer, that is, the logs displayed by the **show logging** command. **file**: Filters out logs written into log files. **server**: Filters out logs sent to the log server. **terminal**: Filters out logs sent to the Console and VTY terminal (including Telnet and SSH).
Command Mode	Global configuration mode
Configuration Usage	The default filtering direction is **all**, that is, all logs are filtered out. Run the **default logging filter direction** command to restore the default filtering direction.

Configuring the Log Filtering Mode

Command	**logging filter type** { **contains-only** | **filter-only** }
Parameter Description	**contains-only**: Indicates that only logs that contain keywords specified in the filtering rules are displayed. **filter-only**: Indicates that logs that contain keywords specified in the filtering rules are filtered out and will not be displayed.
Command Mode	Global configuration mode
Configuration Usage	Log filtering modes include contains-only and filter-only. The default filtering mode is filter-only.

Configuring the Log Filtering Rule

Command	**logging filter rule** { **exact-match module** *module-name* **mnemonic** *mnemonic-name* **level** *level* | **single-match** { **level** *level* | **mnemonic** *mnemonic-name* | **module** *module-name* } }
Parameter Description	**exact-match**: If exact-match is selected, you must specify all three filtering options. **single-match**: If single-match is selected, you may specify only one of the three filtering options. **module** *module-name*: Indicates the module name. Logs of this module will be filtered out. **mnemonic** *mnemonic-name*: Indicates the mnemonic. Logs with this mnemonic will be filtered out. **level** *level*: Indicates the log level. Logs of this level will be filtered out.

Command Mode	Global configuration mode
Configuration Usage	Log filtering rules include exact-match and single-match. The **no logging filter rule exact-match** [**module** *module-name* **mnemonic** *mnemonic-name* **level** *level*] command is used to delete the exact-match filtering rules. You can delete all exact-match filtering rules at a time or one by one. The **no logging filter rule single-match** [**level** *level* | **mnemonic** *mnemonic-name* | **module** *module-name*] command is used to delete the single-match filtering rules. You can delete all single-match filtering rules at a time or one by one.

Configuration Example

Configuring Syslog Filtering

Scenario	It is required to configure the syslog filtering function as follows: 1. Set the filtering directions of logs to **terminal** and **server.** 2. Set the log filtering mode to filter-only. 3. Set the log filtering rule to single-match to filter out logs that contain the module name "SYS".
Configuration Steps	● Configure the syslog filtering function.
	FS# configure terminal FS(config)# logging filter direction server FS(config)# logging filter direction terminal FS(config)# logging filter type filter-only FS(config)# logging filter rule single-match module SYS
Verification	● Run the **show running-config | include loggging** command to display the configuration. ● Enter and exit global configuration mode, and verify that the system displays logs accordingly.
	FS#configure Enter configuration commands, one per line. End with CNTL/Z. FS(config)#exit FS# FS#show running-config | include logging logging filter direction server logging filter direction terminal logging filter rule single-match module SYS

6.4.8 Configuring Syslog Redirection

Configuration Effect

- On the Stacking, logs on the secondary or standby device are displayed on its Console window, and redirected to the active device for display on the Console or VTY window, or stored in the memory buffer, extended flash, or syslog server.
- On a box-type Stacking, after the log redirection function is enabled, logs on the secondary or standby device will be redirected to the active device, and the role flag (*device ID) will be added to each log to indicate that the log is redirected. Assume that four devices form a Stacking. The ID of the active device is 1, the ID of the secondary device is 2, and the IDs of two standby devices are 3 and 4. The role flag is not added to logs generated by the active device. The role flag (*2) is added to logs redirected from the secondary device to the active device. The role flags (*3) and (*4) are added respectively to logs redirected from the two standby devices to the active device.
- On a card-type Stacking, after the log redirection function is enabled, logs on the secondary or standby supervisor module will be redirected to the active supervisor module, and the role flag "(device ID/supervisor module name) will be added to each log to indicate that the log is redirected. If four supervisor modules form a Stacking, the role flags are listed as follows: (*1/M1), (*1/M2), (*2/M1), and (*2/M2).

Notes

- The syslog redirection function takes effect only on the Stacking.
- You can limit the rate of logs redirected to the active device to prevent generating a large amount of logs on the secondary or standby device.

Configuration Steps

Enabling Log Redirection

- (Optional) By default, log redirection is enabled on the Stacking.
- Unless otherwise specified, perform this configuration on the active device of Stacking or active supervisor module.

Configuring the Rate Limit

- (Optional) By default, a maximum of 200 logs can be redirected from the standby device to the active device of Stacking per second.
- Unless otherwise specified, perform this configuration on the active device of Stacking or active supervisor module.

Verification

- Run the **show running** command to display the configuration.

Related Commands

Enabling Log Redirection

Command	**logging rd on**
Parameter Description	N/A
Command Mode	Global configuration mode

Configuration Usage	By default, log redirection is enabled on the Stacking.

Configuring the Rate Limit

Command	**logging rd rate-limit** *number* [**except** *level*]
Parameter Description	**rate-limit** *number*: Indicates the maximum number of logs redirected per second. The value ranges from 1 to 10,000. **except** *level*: Rate limit is not applied to logs with a level equaling to or lower than the specified severity level. By default, the severity level is error (Level 3), that is, rate limit is not applied to logs of Level 3 or lower.
Command Mode	Global configuration mode
Configuration Usage	By default, a maximum of 200 logs can be redirected from the standby device to the active device of Stacking per second.

Configuration Example

Configuring Syslog Redirection

Scenario	It is required to configure the syslog redirection function on the Stacking as follows: 1. Enable the log redirection function. 2.Set the maximum number of logs with a level higher than critical (Level 2) that can be redirected per second to 100.
Configuration Steps	● Configure the syslog redirection function.
	FS# configure terminal FS(config)# logging rd on FS(config)# logging rd rate-limit 100 except critical
Verification	● Run the **show running-config | include logging** command to display the configuration. ● Generate a log on the standby device, and verify that the log is redirected to and displayed on the active device.
	FS#show running-config | include logging logging rd rate-limit 100 except critical

6.4.9 Configuring Syslog Monitoring

Configuration Effect

● Record login/exit attempts. After logging of login/exit attempts is enabled, the related logs are displayed on the device when users access the device through Telnet or SSH. This helps the administrator monitor the device connections.

● Record modification of device configurations. After logging of operations is enabled, the related logs are displayed on the device when users modify the device configurations. This helps the administrator monitor the changes in device configurations.

Notes

- If both the **logging userinfo** command and the **logging userinfo command-log** command are configured on the device, only the configuration result of the **logging userinfo command-log** command is displayed when you run the **show running-config** command.

Configuration Steps

Enabling Logging of Login/Exit Attempts

- (Optional) By default, logging of login/exit attempts is disabled.
- Unless otherwise specified, perform this configuration on every line of the device to enable logging of login/exit attempts.

Enabling logging of Operations

- (Optional) By default, logging of operations is disabled.
- Unless otherwise specified, perform this configuration on every line of the device to enable logging of operations.

Verification

- Run the **show running** command to display the configuration.

Related Commands

Enabling Logging of Login/Exit Attempts

Command	**logging userinfo**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	By default, a device does not generate related logs when users log into or exit the device.

Enabling Logging of Operations

Command	**logging userinfo command-log**
Parameter Description	N/A
Command Mode	Global configuration mode
Configuration Usage	The system generates related logs when users run configuration commands. By default, a device does not generate logs when users modify device configurations.

Configuration Example

Configuring Syslog Monitoring

Scenario	It is required to configure the syslog monitoring function as follows: 1. Enable logging of login/exit attempts. 2. Enable logging of operations.
Configuration Steps	● Configure the syslog monitoring function.
	FS# configure terminal FS(config)# logging userinfo FS(config)# logging userinfo command-log
Verification	● Run the **show running-config | include logging** command to display the configuration. ● Run a command in global configuration mode, and verify that the system generates a log.
	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/0 *Jun 16 15:03:43: %CLI-5-EXEC_CMD: Configured from console by admin command: interface GigabitEthernet 0/0 FS#show running-config | include logging logging userinfo command-log

6.4.10 Synchronizing User Input with Log Output

Configuration Effect

● By default, the user input is not synchronized with the log output. After this function is enabled, the content input during log output is displayed after log output is completed, ensuring integrity and continuity of the input.

Notes

● This command is executed in line configuration mode. You need to configure this command on every line as required.

Configuration Steps

Synchronizing User Input with Log Output

● (Optional) By default, the synchronization function is disabled.

● Unless otherwise specified, perform this configuration on every line to synchronize user input with log output.

Verification

● Run the **show running** command to display the configuration.

Related Commands

Synchronizing User Input with Log Output

Command	**logging synchronous**
Parameter Description	N/A
Command Mode	Line configuration mode
Configuration Usage	This command is used to synchronize the user input with log output to prevent interrupting the user input.

Configuration Example

Synchronizing User Input with Log Output

Scenario	It is required to synchronize the user input with log output as follows: 1. Enable the synchronization function.
Configuration Steps	● Configure the synchronization function.
	FS# configure terminal FS(config)# line console 0 FS(config-line)# logging synchronous
Verification	● Run the **show running-config | begin line** command to display the configuration.
	FS#show running-config | begin line line con 0 logging synchronous login local As shown in the following output, when a user types in "vlan", the state of interface 0/1 changes and the related log is output. After log output is completed, the log module automatically displays the user input "vlan" so that the user can continue typing. FS(config)#vlan *Aug 20 10:05:19: %LINK-5-CHANGED: Interface GigabitEthernet 0/1, changed state to up *Aug 20 10:05:19: %LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet 0/1, changed state to up FS(config)#vlan

6.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears logs in the memory buffer.	**clear logging**

Displaying

Description	Command
Displays log statistics and logs in the memory buffer based on the timestamp from oldest to latest.	**show logging**
Displays log statistics and logs in the memory buffer based on the timestamp from latest to oldest.	**show logging reverse**
Displays syslog configurations and statistics.	**show logging config**
Displays log statistics of each module in the system.	**show logging count**

7 Configuring CWMP

7.1 Overview

CPE WAN Management Protocol (CWMP) provides a general framework of unified device management, related message specifications, management methods, and data models, so as to solve difficulties in unified management and maintenance of dispersed customer-premises equipment (CPEs), improve troubleshooting efficiency, and save O&M costs.

CWMP provides the following functions:

- **Auto configuration and dynamic service provisioning.** CWMP allows an Auto-Configuration Server (ACS) to automatically provision CPEs who initially access the network after start. The ACS can also dynamically re-configure running CPEs.
- **Firmware management.** CWMP manages and upgrades the firmware and its files of CPEs.
- **Software module management.** CWMP manages modular software according to data models implemented.
- **Status and performance monitoring.** CWMP enables CPEs to notify the ACE of its status and changes, achieving real-time status and performance monitoring.
- **Diagnostics.** The ACE diagnoses or resolves connectivity or service problems based on information from CPEs, and can also perform defined diagnosis tests.

Protocols and Standards

For details about TR069 protocol specifications, visit http://www.broadband-forum.org/technical/trlist.php.

Listed below are some major CWMP protocol specifications:

- TR-069_Amendment-4.pdf: CWMP standard
- TR-098_Amendment-2.pdf: Standard for Internet gateway device data model
- TR-106_Amendment-6.pdf: Standard for CPE data model
- TR-181_Issue-2_Amendment-5.pdf: Standard for CPE data model 2
- tr-098-1-4-full.xml: Definition of Internet gateway device data model
- tr-181-2-4-full.xml: Definition 2 of CPE data model 2

7.2 Applications

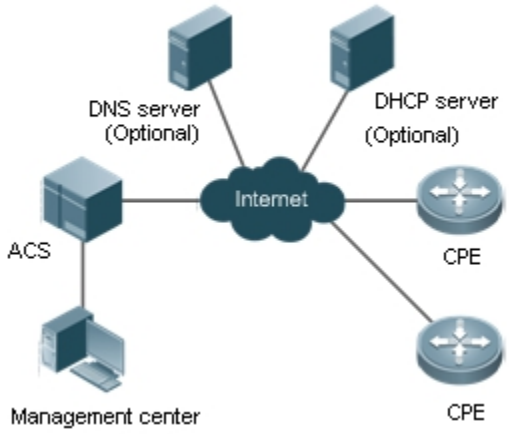
Typical Application	Scenario
CWMP Network Application Scenario	Initiate CPE-ACS connection, so as to upgrade the CPE firmware, upload the configuration files, restore the configuration, and realize other features.

7.2.1 CWMP Network Application Scenario

Application Scenario

The major components of a CWMP network architecture are CPEs, an ACS, a management center, a DHCP server, and a Domain Name System (DNS) server. The management center manages a population of CPEs by controlling the ACS on a Web browser.

Figure 7- 1



Note	● If the Uniform Resource Locator (URL) of the ACS is configured on CPEs, the DHCP server is optional. If not, the DHCP is required to dynamically discover the ACS URL. ● If the URLs of the ACS and CPEs contain IP addresses only, the DNS server is optional. If their URLs contain domain names, the DNS server is required to resolves the names.

Functional Deployment

HTTP runs on both CPEs and the ACS.

7.3 Features

Basic Concept

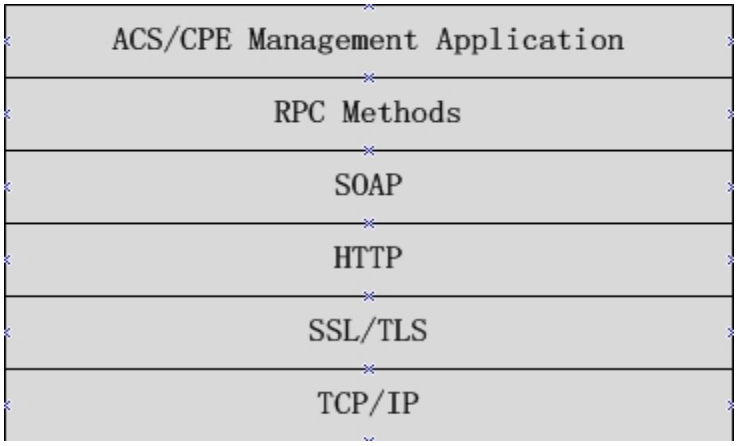
↘ Major Terminologies

- **CPE**: Customer Premises Equipment
- **ACS**: Auto-Configuration Server
- **RPC**: Remote Procedure Call
- **DM**: Data Model

↘ Protocol Stack

Figure 7- 2 shows the protocol stack of CWMP.

Figure 7- 2 CWMP Protocol Stack



As shown in Figure 7-2, CWMP defines six layers with respective functions as follows:

- ACS/CPE Application

The application layer is not a part of CWMP. It is the development performed by various modules of the CPEs/ACS to support CWMP, just like the Simple Network Management Protocol (SNMP), which does not cover the MIB management of functional modules.

- RPC Methods

This layer provides various RPC methods for interactions between the ACS and the CPEs.

- SOAP

The Simple Object Access Protocol (SOAP) layer uses a XML-based syntax to encode and decode CWMP messages.. Thus, CWMP messages must comply with the XML-based syntax.

- HTTP

All CWMP messages are transmitted over Hypertext Transfer Protocol (HTTP). Both the ACS and the CPEs can behave in the role of HTTP clients and servers. The server function is used to monitor reverse connections from the peer.

- SSL/TLS

The Secure Sockets Layer (SSL) or Transport Layer Security (TLS) layer guarantees CWMP security, including data integrity, confidentiality, and authentication.

- TCP/IP

This layer is the (Transmission Control Protocol/Internet Protocol (TCP/IP) protocol stack.

↘ RPC Methods

The ACS manages and monitors CPEs by calling mostly the following RPC methods:

- Get RPC Methods

The Get methods enable the ACS to remotely obtain the set of RPC methods, as well as names, values and attributes of the DM parameters supported on CPEs.

- Set RPC Methods

The Set methods enable the ACS to remotely set the values and attributes of the DM parameters supported on CPEs.

- Inform RPC Methods

The Inform methods enable CPEs to inform the ACS of their device identifiers, parameter information, and events whenever sessions are

established between them.

- Download RPC Methods

The Download method enables the ACS to remotely control the file download of CPEs, including firmware management, upgrade, and Web package upgrade.

- Upload RPC Methods

The Upload method enables the ACS to remotely control the file upload of CPEs, including upload of firmware and logs.

- Reboot RPC Methods

The Reboot method enables the ACS to remotely reboot the CPEs.

Session Management

CWMP sessions or interactions are the basis for CWMP. All CWMP interactions between the ACS and CPEs rely on their sessions. CWMP helps initiate and maintain ACS-CPE sessions to link them up for effective management and monitoring. An ACS-CPE session is a TCP connection, which starts from the Inform negotiation to TCP disconnection. The session is classified into CPE Initiated Session and ACS Initiated Session according to the session poster.

DM Management

CWMP operates based on CWMP Data Model (DM). CWMP manages all functional modules by a set of operations performed on DM. Each functional module registers and implements a respective data model, just like the MIBs implemented by various functional modules of SNMP.

A CWMP data model is represented in the form of a character string. For a clear hierarchy of the data model, a dot (.) is used as a delimiter to distinguish an upper-level data model node from a lower-level data model node. For instance, in the data model **InternetGatewayDevice.LANDevice**, **InternetGatewayDevice** is the parent data model node of **LANDevice**, and **LANDevice** is the child data model node of **InternetGatewayDevice**.

DM nodes are classified into two types: object nodes and parameter nodes. The parameter nodes are also known as leaf nodes. An object node is a node under which there are child nodes, and a parameter node is a leaf node under which there is no any child node. Object nodes are further classified into single-instance object nodes and multi-instance object nodes. A single-instance object node is an object node for which there is only one instance, whereas a multi-instance object node is an object node for which there are multiple instances.

DM nodes can also be classified into readable nodes and readable-and-writable nodes. A readable node is a node whose parameter values can be read but cannot be modified, and a readable-and-writable node is a node whose parameter values can be both read and modified.

A data model node has two attributes. One attribute relates to a notification function; that is, whether to inform the ACS of changes (other than changes caused by CWMP) to parameter values of the data model. The other attribute is an identifier indicating that the parameters of the data model node can be written using other management modes (than the ACS); that is, whether the values of the parameters can be modified using other management modes such as Telnet. The ACS can modify the attributes of the data models using RPC methods.

CWMP manages the data models using corresponding RPC methods.

Event Management

When some events concerned by the ACS occur on the CPE, the CPE will inform the ACS of these events. The ACS monitors these events to monitor the working status of the CPE. The CWMP events are just like Trap messages of SNMP or product logs. Using RPC methods, to the ACS filters out the unconcerned types of events. CWMP events are classified into two types: single or (not cumulative) events and

multiple (cumulative) events. A single event means that there is no quantitative change to the same event upon re-occurrence of the event, with the old discarded and the newest kept. A multiple event means that the old are not discarded and the newest event is kept as a complete event when an event re-occurs for multiple times later; that is, the number of this event is incremented by 1.

All events that occur on the CPE are notified to the ACS using the INFORM method.

Features

Feature	Description
Upgrading the Firmware	The ACS controls the upgrade of the firmware of a CPE using the Download method.
Upgrading the Configuration Files	The ACS controls the upgrade of the configuration files of a CPE using the Download method.
Uploading the Configuration Files	The ACS controls the upload of the configuration files of a CPE using the Upload method.
Backing up and Restoring a CPE	When a CPE breaks away from the management center, this feature can remotely restore the CPE to the previous status.

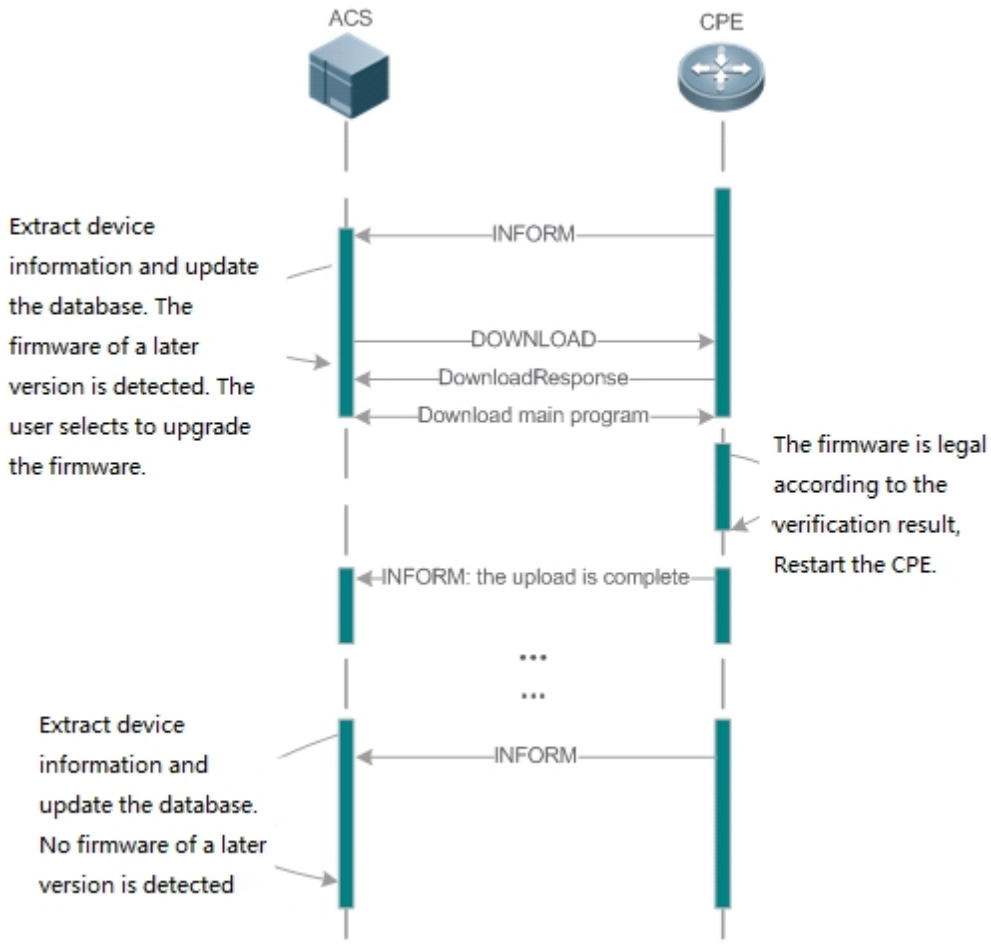
7.3.1 Upgrading the Firmware

Upgrading the Firmware means the firmware of a network element (NE) can be upgraded, so as to implement device version upgrade or replacement.

Working Principle

Sequence Diagram of Upgrading the Firmware

Figure 7- 3



Users specify a CPE for the ACS to deliver the Download method for upgrading the firmware. The CPE receives the request and starts to download the latest firmware from the destination file server, upgrade the firmware, and then reboot. After restart, the CPE will indicate the successful or unsuccessful completion of the method application.

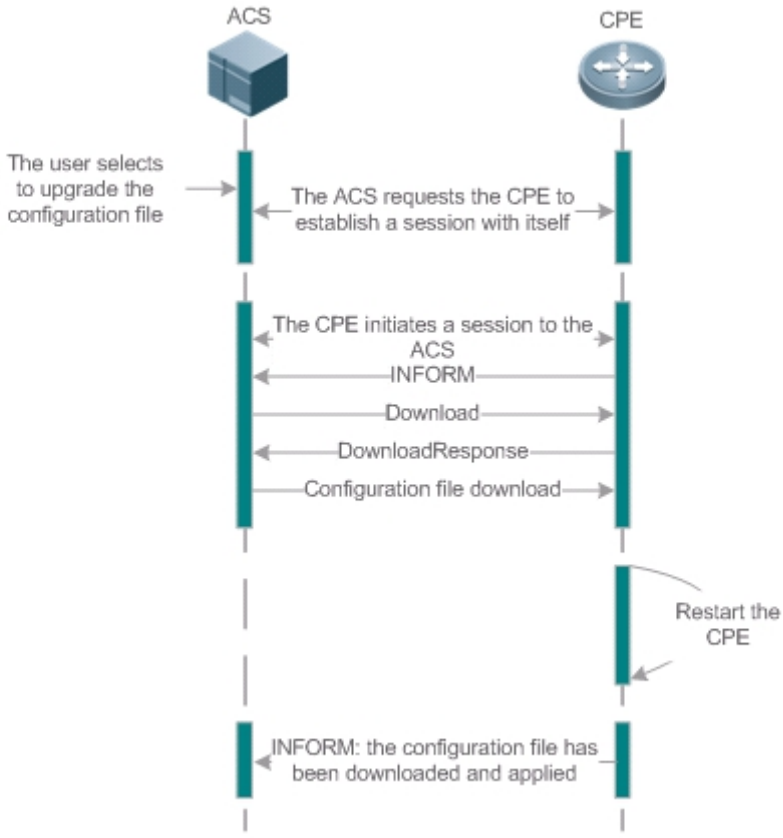
The file server can be ACS or separately deployed.

7.3.2 Upgrading the Configuration Files

Upgrading the Configuration Files means the current configuration files of a CPE can be replaced with specified configuration files, so that the new configuration files act on the CPE after reset.

Working Principle

Figure 7- 4



Users specify a CPE for the ACS to deliver the Download methods for upgrading its configuration files. The CPE downloads the configuration files from the specified file server, upgrade configuration files, and then reboot. After that, the CPE will indicate successful or unsuccessful completion of the method application.

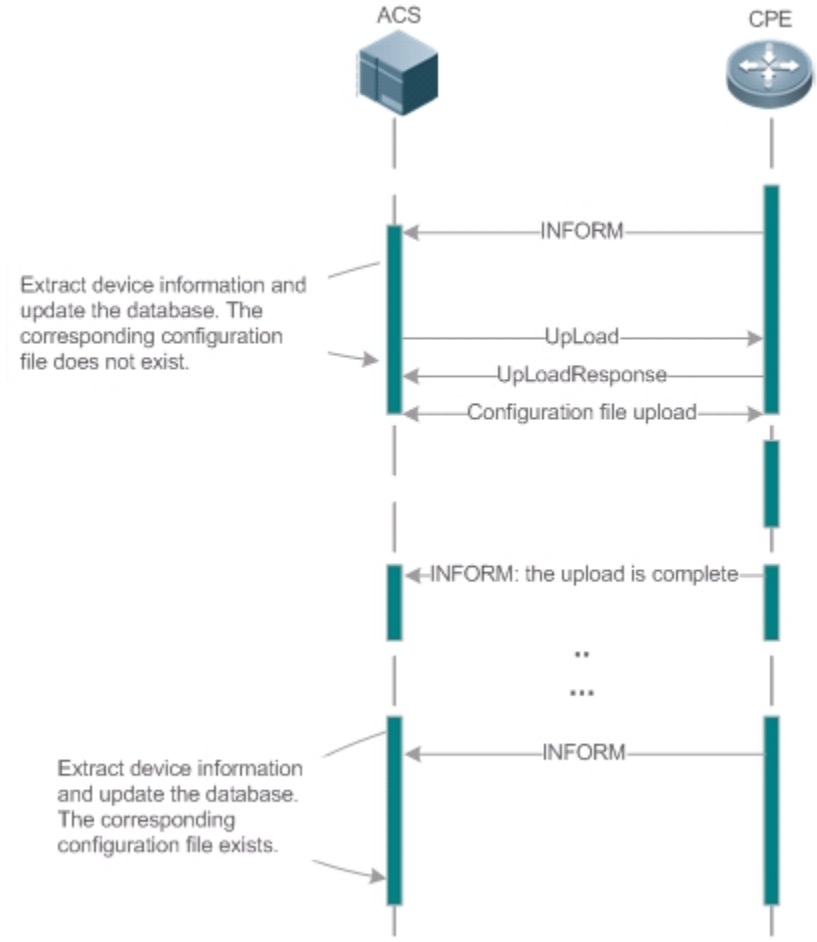
The file server can be ACS or separately deployed.

7.3.3 Uploading the Configuration Files

Uploading the Configuration Files means the ACS controls the configuration files of CPEs by using the Upload method.

Working Principle

Figure 7- 5



When a CPE initially accesses the ACS, the ACS attempts to learn the configuration files of the CPE in the following sequence:

- When the ACS initially receives an Inform message from the CPE, it locates the corresponding database information according to device information carried in the message.
- If the database does not contain the configuration files of the CPE, the ACS delivers the Upload method to the CPE for uploading the configuration files.
- The CPE uploads its current configuration files to the ACS.
- The CPE returns a successful or unsuccessful response to the Upload request.

7.3.4 Backing Up and Restoring a CPE

When a remote CPE breaks away from the management center due to abnormal operations, the CPE backup and restoration feature helps restore the CPE to the previous status, so that the management center can resume the supervision of the CPE as necessary.

Working Principle

You can configure the restoration function on a CPE, so that the CPE can restore itself from exceptions of its firmware or configuration files. Then when the CPE fails to connect to the ACS and breaks away from the management center after its firmware or configuration files are upgraded, the previous firmware or configuration files of the CPE can be restored in time for the ACS to manage the CPE. This kind of exception is generally caused by delivery of a wrong version or configuration file.

Before the CPE receives a new firmware or configuration files to upgrade, the CPE will back up its current version and configuration files.

In addition, there is a mechanism for determining whether the problem described in the preceding scenario has occurred. If the problem has occurred, the CPE is restored to the previous manageable status.

7.4 Configuration

Action	Suggestions and Related Commands	
Establishing a Basic CWMP Connection	(Mandatory) You can configure the ACS or CPE usernames and passwords to be authenticated for CWMP connection.	
	cwmp	Enables CWMP and enters CWMP configuration mode.
	acs username	Configures the ACS username for CWMP connection.
	acs password	Configures the ACS password for CWMP connection.
	cpe username	Configures the CPE username for CWMP connection.
	cpe password	Configures the CPE password for CWMP connection.
	（Optional）You can configure the URLs of the CPE and the ACS.	
	acs url	Configures the ACS URL.
	cpe url	Configures the CPE URL.
Configuring CWMP-Related Attributes	(Optional) You can configure the basic functions of the CPE, such as upload, backup and restoration of firmware, configuration files or logs.	
	cpe inform	Configures the periodic notification function of the CPE.
	cpe back-up	Configures the backup and restoration of the firmware and configuration file of the CPE.
	disable download	Disables the function of downloading firmware and configuration files from the ACS.
	disable upload	Disables the function of uploading configuration and log files to the ACS.
	timer cpe- timeout	Configures the ACS response timeout on CPEs.

7.4.1 Establishing a Basic CWMP Connection

Configuration Effect

- A session connection is established between the ACS and the CPE.

Precautions

- N/A

Configuration Method

Enabling CWMP and Entering CWMP Configuration Mode

- (Mandatory) The CWMP function is enabled by default.

Command	**cwmp**
Parameter Description	N/A
Defaults	CWMP is enabled by default.
Command Mode	Global configuration guide
Usage Guide	N/A

Configuring the ACS Username for CWMP Connection

- This configuration is mandatory on the ACS.
- Only one username can be configured for the ACS. If multiple are configured, the latest configuration is applied.

Command	**acs username** *username*
Parameter Description	**username** *username*: The ACS username for CWMP connection
Defaults	The ACS username is not configured by default.
Command Mode	CWMP configuration mode
Usage Guide	N/A

Configuring the ACS Password for CWMP Connection

- This configuration is mandatory on the ACS.
- The password of the ACS can be in plaintext or encrypted form. Only one password can be configured for the ACS. If multiple are configured, the latest configuration is applied.

Command	**acs password** {*password* | *encryption-type encrypted-password*}
Parameter Description	*password:* ACS password *encryption-type:* 0 (no encryption) or 7 (simple encryption) *encrypted-password:* Password text
Defaults	*encryption-type*: 0 *encrypted-password*: N/A
Command	CWMP configuration mode

Mode	
Usage Guide	N/A

Configuring the CPE Username for CWMP Connection

- This configuration is mandatory on the CPE.
- Only one username can be configured for the CPE. If multiple are configured, the latest configuration is applied.

Command	**cpe username** *username*
Parameter Description	*username*: CPE username
Defaults	No CPE username is configured by default.
Command Mode	CWMP configuration mode
Usage Guide	N/A

Configuring the CPE Password for CWMP Connection

- This configuration is mandatory on the CPE.
- The password of the CPE can be in plaintext or encrypted form. Only one password can be configured for the CPE. If multiple are configured, the latest configuration is applied.

Command	**cpe password** {*password* | *encryption-type encrypted-password*}
Parameter Description	*password:* CPE password *encryption-type:* 0 (no encryption) or 7 (simple encryption) *encrypted-password:* Password text
Defaults	*encryption-type*: 0 *encrypted-password*: N/A
Command Mode	CWMP configuration mode
Usage Guide	Use this command to configure the CPE user password to be authenticated for the ACS to connect to the CPE. In general, the encryption type does not need to be specified. The encryption type needs to be specified only when copying and pasting the encrypted password of this command. A valid password should meet the following format requirements: ● Contain 1 to 26 characters including letters and figures. ● The leading spaces will be ignored, while the trailing and middle are valid. ● If 7 (simple encryption) is specified, the valid characters only include 0 to 9 and a (A) to f (F).

Configuring the ACS URL for CMWP Connection

- This configuration is optional on the CPE.
- Only one ACS URL can be configured. If multiple are configured, the latest configuration is applied. The ACS URL must be in HTTP format.

Command	**acs url** { *url* | **macc** }
Parameter Description	*url*: ACS URL
Defaults	No ACS URL is configured by default.
Command Mode	CWMP configuration mode
Usage Guide	If the ACS URL is not configured but obtained through DHCP, CPEs will use this dynamic URL to initiate connection to the ACS. The ACS URL must: ● Be in format of **http://host[:port]/path** or **https://host[:port]/path**. ● Contain 256 characters at most. cpe inform interval 30

Configuring the CPE URL for CWMP Connection

- This configuration is optional on the CPE.
- Only one CPE URL can be configured. If multiple are configured, the latest configuration is applied. The CPE URL must be in HTTP format instead of domain name format.

Command	**cpe url** *url*
Parameter Description	*url:* CPE URL
Defaults	No CPE URL is configured by default.
Command Mode	CWMP configuration mode
Usage Guide	If CPE URL is not configured, it is obtained through DHCP. The CPE URL must: ● Be in format of **http://ip [: port]/**. ● Contain 256 characters at most.

Verification

- Run the **show cwmp configuration** command.

Command	**show cwmp configuration**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	N/A
Configuration Examples	The following example displays the CWMP configuration. FS(config-cwmp)#show cwmp configuration CWMP Status : enable

	ACS URL : http://www.fs.com/acs ACS username : admin ACS password : ****** CPE URL : http://10.10.10.2:7547/ CPE username : fs CPE password : ****** CPE inform status : disable CPE inform interval : 60s CPE inform start time : 0:0:0 0 0 0 CPE wait timeout : 50s CPE download status : enable CPE upload status : enable CPE back up status : enable CPE back up delay time : 60s

Configuration Examples

The following configuration examples describe CWMP-related configuration only.

Configuring Usernames and Passwords on the CPE

Network Environment Figure 7- 6	
Configuration Method	● Enable CWMP. ● On the CPE, configure the ACS username and password to be authenticated for the CPE to connect to the ACS. ● On the CPE, configure the CPE username and password to be authenticated for the ACS to connect to the CPE.
CPE	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# cwmp FS(config-cwmp)# acs username USERB FS(config-cwmp)# acs password PASSWORDB FS(config-cwmp)# cpe username USERB FS(config-cwmp)# cpe password PASSWORDB
Verification	● Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.

CPE	FS # show cwmp configuration CWMP Status : enable ACS URL : http://10.10.10.1:7547/acs ACS username : USERA ACS password : ****** CPE URL : http://10.10.10.2:7547/ CPE username : USERB CPE password : ******

Configuring the URLs of the ACS and the CPE

Network Environment	See Figure 7-6.
Configuration Method	● Configure the ACS URL. ● Configure the CPE URL.
CPE	FS# configure terminal FS(config)# cwmp FS(config-cwmp)# acs url http://10.10.10.1:7547/acs FS(config-cwmp)# cpe url http://10.10.10.1:7547/
Verification	Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.
CPE	FS #show cwmp configuration CWMP Status : enable ACS URL : http://10.10.10.1:7547/acs ACS username : USERA ACS password : ****** CPE URL : http://10.10.10.2:7547/

Common Errors

- The user-input encrypted password is longer than 254 characters, or the length of the password is not an even number.
- The user-input plaintext password is longer than 100 characters.
- The user-input plaintext password contains illegal characters.
- The user-input encrypted password contains illegal characters (the legitimate characters includes only 0~9, a~f and A~F)
- The URL of the ACS is set to **NULL**.
- The URL of the CPE is set to **NULL**.

7.4.2 Configuring CWMP-Related Attributes

Configuration Effect

- You can configure common functions of the CPE, such as the backup and restoration of its firmware or configuration file, whether to enable the CPE to download firmware and configuration files from the ACS, and whether to enable the CPE to upload its configuration and log files to the ACS.

Configuration Method

Configuring the Periodic Notification Function of the CPE

- (Optional) The value range is from 30 to 3,600 in seconds. The default value is 600 seconds.
- Perform this configuration to reset the periodical notification interval of the CPE.

Command	**cpe inform** [**interval** *seconds*] [**start-time** *time*]
Parameter Description	*seconds*: Specifies the periodical notification interval of the CPE. The value range is from 30 to 3,600 in seconds. *time*: Specifies the date and time for starting periodical notification in *yyyy-mm-ddThh:mm:ss* format.
Command Mode	CWMP configuration mode
Defaults	The default value is 600 seconds.
Usage Guide	Use this command to configure the periodic notification function of the CPE. ● If the time for starting periodical notification is not specified, periodical notification starts after the periodical notification function is enabled. The notification is performed once within every notification interval. ● If the time for starting periodical notification is specified, periodical notification starts at the specified start time. For instance, if the periodical notification interval is set to 60 seconds and the start time is 12:00 am next day, periodical notification will start at 12:00 am next day and once every 60 seconds.

Disabling the Function of Downloading Firmware and Configuration Files from the ACS

- (Optional) The CPE can download firmware and configuration files from the ACS by default.
- Perform this configuration if the CPE does not need to download firmware and configuration files from the ACS.

Command	**disable download**
Parameter Description	N/A
Defaults	The CPE can download firmware and configuration files from the ACS by default.
Command Mode	CWMP configuration mode
Usage Guide	Use this command to disable the function of downloading main program and configuration files from the ACS. ● This command does not act on configuration script files. The configuration scripts can still be executed even if this function is disabled.

Disabling the Function of Uploading Configuration and Log Files to the ACS

- (Optional.) The CPE can upload configuration and log files to the ACS by default.
- Perform this configuration if the CPE does not need to upload configuration and log files to the ACS.

Command	**disable upload**
Parameter Description	N/A
Defaults	The CPE can upload configuration and log files to the ACS by default.
Command Mode	CWMP configuration mode
Usage Guide	Use this command to disable the function of uploading configuration and log files to the ACS.

Configuring the Backup and Restoration of the Firmware and Configuration Files of the CPE

- (Optional) The backup and restoration of the firmware and configuration files of the CPE is enabled by default. The value range is from 30 to 10,000 in seconds. The default value is 60 seconds.
- The longer the delay-time is, the longer the reboot will be complete.
- Perform this configuration to modify the function of backing up and restoring the firmware and configuration files of the CPE.

Command	**cpe back-up** [**delay-time** *seconds*]
Parameter Description	*seconds*: Specifies the delay for backup and restoration of the firmware and configuration file of the CPE.
Defaults	The default value is 60 seconds.
Command Mode	CWMP configuration mode
Usage Guide	N/A

Configuring the ACS Response Timeout

- (Optional) The value range is from 10 to 600 in seconds. The default value is 30 seconds.
- Perform this configuration to modify the ACS response timeout period on the CPE.

Command	**timer cpe- timeout** *seconds*
Parameter Description	*seconds*: Specifies the timeout period in seconds. The value range is from 10 to 600.
Defaults	The default value is 30 seconds.
Command Mode	CWMP configuration mode
Usage Guide	N/A

Verification

- Run the show cwmp configuration command.

Command	**show cwmp configuration**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	N/A
Configuration Examples	The following example displays the CWMP configuration. FS(config-cwmp)#show cwmp configuration CWMP Status : enable ACS URL : http://www.fs.com/acs ACS username : admin ACS password : ****** CPE URL : http://10.10.10.2:7547/ CPE username : fs CPE password : ****** CPE inform status : disable CPE inform interval : 60s CPE inform start time : 0:0:0 0 0 0 CPE wait timeout : 50s CPE download status : enable CPE upload status : enable CPE back up status : enable CPE back up delay time : 60s

Configuration Examples

Configuring the Periodical Notification Interval of the CPE

Network Environment	See Figure 7-6.
Configuration Steps	● Enable the CWMP function and enter CWMP configuration mode. ● Set the periodical notification interval of the CPE to 60 seconds.
CPE	FS#config Enter configuration commands, one per line. End with CNTL/Z. FS(config)#cwmp FS(config-cwmp)#cpe inform interval 60

Verification	Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.
CPE	FS #show cwmp configuration CWMP Status : enable …… CPE inform interval : 60s

Disabling the Function of Downloading Firmware and Configuration Files from the ACS

Network Environment	See Figure 7-6.
Steps	● Enable the CWMP function and enter CWMP configuration mode. ● Disable the function of downloading firmware and configuration files from the ACS.
CPE	FS#config Enter configuration commands, one per line. End with CNTL/Z. FS(config)#cwmp FS(config-cwmp)#disable download
Verification	Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.
CPE	FS #show cwmp configuration CWMP Status : enable …… CPE download status : disable

Disabling the Function of Uploading Configuration and Log Files to the ACS

Network Environment	See Figure 7-6.
Configuration Steps	● Enable the CWMP function and enter CWMP configuration mode. ● Disable the CPE's function of uploading configuration and log files to the ACS.
CPE	FS#config Enter configuration commands, one per line. End with CNTL/Z. FS(config)#cwmp FS(config-cwmp)# disable upload

Verification	Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.
CPE	FS #show cwmp configuration CWMP Status : enable …… CPE upload status : disable

Configuring the Backup and Restoration Delay

Network Environment	See Figure 7-6.
Configuration Steps	● Enable the CWMP function and enter CWMP configuration mode. ● Set the backup and restoration delay to 100 seconds.
CPE	FS#config Enter configuration commands, one per line. End with CNTL/Z. FS(config)#cwmp FS(config-cwmp)# cpe back-up Seconds 30
Verification	● Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.
CPE	FS #show cwmp configuration CWMP Status : enable …… CPE back up delay time : 30s

Configuring the ACS Response Timeout of the CPE

Network Environment	See Figure 7-6.
Configuration Steps	● Enable the CWMP function and enter CWMP configuration mode. ● Set the response timeout of the CPE to 100 seconds.
CPE	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# cwmp FS(config-cwmp)# timer cpe-timeout 100
Verification	● Run the **show** command on the CPE to check whether the configuration commands have been successfully applied.

CPE	FS#show cwmp configuration CWMP Status : enable …… CPE wait timeout : 100s

Common Errors

N/A

7.5 Monitoring

Displaying

Command	Function
show cwmp configuration	Displays the CWMP configuration.
show cwmp status	Displays the CWMP running status.

8 Configuring Monitoring

8.1 Overview

Intelligent monitoring is the intelligent hardware management of FS Network devices, including intelligent fan speed adjustment, and intelligent temperature monitoring. The intelligent monitoring performs the following tasks:

- Automatic fan speed adjustment based on ambient temperature changes
- Real-time temperature monitoring of boards to alert users

By default, the intelligent monitoring function is enabled after the device is powered on. It does not require any manual configuration.

Protocol Specification

N/A

8.2 Features

Basic Concepts

N/A

Features

Feature	Function
Intelligent Speed Adjustment of Fans	The rotating speed of fans is automatically adjusted as the temperature changes to address the heat dissipation needs of the system.
Intelligent Temperature Monitoring	The system automatically monitors the temperature. When the temperature exceeds a certain threshold, the system automatically generates an alarm.
Power Monitoring	The system automatically monitors the power. When the power is insufficient or cannot be identified, the system automatically generates an alarm.

8.2.1 Intelligent Speed Adjustment of Fans

As the ambient temperature rises or drops, the fans automatically raise or reduce their rotating speed to dissipate heat and ensure that the noise is low.

Working Principle

Verification

- Run the **show fan** command to display working status of all fans.
- Run the **show fan speed command** to dislay rotating speed.

8.2.2 Intelligent Temperature Monitoring

The system automatically monitors the temperature. When the temperature changes, the system automatically notify users.

Working Principle

The system monitors the temperature once per minute. When the temperature exceeds a certain threshold, the system takes a certain action. The temperature and action vary with different devices.

Verification

Run the **show temperature** command to display system temperature.

8.2.3 Power Monitoring

The system automatically monitors the power. When the power is insufficient or cannot be identified, the system automatically generates an alarm.

Working Principle

The system monitors the power once per minutes. If the system finds the power insufficient, the alarm LED becomes yellow and a Syslog message is generated. Once the alarm event is resolved, the system recovers. If the system cannot identify the inserted power, the alarm LED becomes yellow. After you remove the power, the system recovers.

Verification

Run the **show power** command to display power information.

9 Configuring Module Hot Swapping

9.1 Overview

Module Hot Swapping is a common maintenance function provided by chassis-based devices.

Module Hot Swapping automates the installation, uninstallation, reset, and information check of hot-swappable modules (management cards, line cards, cross-connect and synchronous timing boards [XCSs], and multi-service cards) after they are inserted into chassis-based devices.

9.2 Applications

Application	Description
Resetting Online Modules	During routine maintenance, you can reset an abnormally running module to troubleshoot the fault.
Clearing the Configuration of a Module	During routine maintenance, you can replace the module in a slot with a different type of module.
Clearing the Configuration of a Stacking Member Device	During routine maintenance, you can clear the configuration of all modules on a Stacking member device and then reconfigure the modules.
Deleting a MAC Address from the Configuration File	During routine maintenance, you can delete the MAC addresses of Stacking member devices to perform MAC address reelection.
Modifying a MAC Address in the Configuration File	When you replace a switch with a new one in gateway mode, you can configure the MAC address of the new switch to be the same as that of the replaced switch to retain the MAC address of the bound gateway on downstream devices.

9.2.1 Resetting Online Modules

Scenario

During routine maintenance, you can reset an abnormally running module in a slot to troubleshoot the fault.

Deployment

Run the **reset module** command on the console to reset a module.

9.2.2 Clearing the Configuration of a Module

Scenario

During routine maintenance, you can replace the module in a slot on a chassis-based device with a different type of module without affecting other modules.

Deployment

Perform the following operations in sequence:

1. Remove the module from the target slot.

2. Run the **remove configuration module** command on the device to remove the module configuration.
3. Insert a new module into the slot.

9.2.3 Clearing the Configuration of a Stacking Member Device

Scenario

In Stacking mode, to meet service change requirements, you need to clear all configurations on a member device and reconfigure the device. You can run the **remove configuration device** command to clear configurations all at once, rather than clear the configuration of individual modules one by one on the member device.

Deployment

Perform the following operations in sequence:

1. Run the **remove configuration device** command on the target device.
2. Save the configuration.
3. Restart the Stacking and check whether the configuration of the device is cleared.

9.2.4 Deleting the MAC Address from the Configuration File

Scenario

In general, the MAC address used by a system is written in the management card or the flash memory of the chassis. In Stacking mode, to avoid service interruption due to the change of the MAC address, the system automatically saves the MAC address to the configuration file. After the system restarts, the valid MAC address (if any) in the configuration file is used in preference. The **no sysmac** command can be used to delete the MAC address from the configuration file. Then the MAC address written in the flash memory is used by default.

Deployment

Perform the following operations in sequence:

1. Run the **no sysmac** command on the target device to delete its MAC address.
2. Save the configuration.
3. Restart the Stacking and check whether the MAC address of the device is reelected.

9.2.5 Modifying a MAC Address in the Configuration File

Scenario

In gateway mode (the **auth-mode gateway** command is configured), some peripheral devices are configured with the MAC address of the bound gateway. If the gateway is replaced, you can use the **sysmac** command to configure the MAC address of the new gateway to be the same as that of the replaced gateway to retain the MAC address of the bound gateway on downstream devices. The **sysmac** command is valid only in gateway mode.

Deployment

Perform the following operations in sequence:

1. Run the **sysmac** command in gateway mode on the target device.

2. Save the configuration.
3. Restart the device and check whether its MAC address is modified.

9.3 Features

Feature

Feature	Description
Automatically Installing the Inserted Module	After a new module is inserted into a chassis-based device, the device's management software will automatically install the module driver.
Resetting Online Modules	Online modules can be reset.

10.3.1 Automatically Installing the Inserted Module

You can hot-swap (insert and remove) a module on a device in running state without impact on other modules. After the module is inserted into a slot, the device's management software will automatically install the module driver. The configuration of the removed module is retained for subsequent configuration. If the removed module is inserted again, the module will be automatically started with its configuration effective.

The module mentioned here can be a management card, a line card, an XCS, or a multi-service card. A management card can only be inserted in a management card slot (M1 or M2). A line card or multi-service card can be inserted in a line card slot. An XCS can only be inserted in an XCS slot.

Working Principle

After a module is inserted, the device's management software will automatically install the module driver and save the module information (such as the quantity of ports on the module and port type) to the device, which will be used for subsequent configuration. After the module is removed, its information is not cleared by the management software. You can continue to configure the module information. When the module is inserted again, the management software assigns the user's module configuration to the module and make it take effect.

10.3.2 Resetting Online Modules

The management software of a device provides the online module reset feature for module software troubleshooting.

Resetting an online module may interrupt some services on the device.

Working Principle

After you run the **reset module** command, the device's management software uses a hardware or software interface of the device to restart the software on the target module and restores the hardware chip to the post-power-on state. The software failure of the module will be rectified after the module is reset.

9.4 Configuration

The module Hot Swapping feature is automatically implemented without manual configuration.

Configuration	Description and Command	
Clearing Module and Device Configuration	(Optional) It is used to clear configuration in global configuration mode. After you run the following commands, you need to save the command configuration so that it can take effect after system restart.	
	remove configuration device *device-id*	Clears the configuration of a Stacking member device.
	no sysmac	Deletes a MAC address from the configuration file.
	sysmac	Modifies a MAC address in the configuration file.

10.4.1 Clearing Module and Device Configuration

Configuration Effect

- Clear the configuration of a module.
- Clear the configuration of a Stacking member device.
- Delete a MAC address from the configuration file.

Configuration Steps

Clearing the Configuration of a Stacking Member Device

- (Optional) Perform this configuration when you need to clear the configuration of a Stacking member device.

Command	**remove configuration device** *device-id*
Parameter Description	*device-id:* Indicates the ID of a chassis.
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command to clear the configuration of a Stacking member device.

Deleting a MAC Address from the Configuration File

- (Optional) Perform this configuration when you need to change the MAC address of a system to the reelected MAC address.
- In general, the MAC address used by a system is written in the management card or the flash memory of the chassis. In Stacking mode, to avoid service interruption due to the change of the MAC address, the system automatically saves the MAC address to the configuration file. After the system restarts, the valid MAC address (if any) in the configuration file is used in preference.

Command	**no sysmac**

Parameter Description	N/A
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command to delete a MAC address from the configuration file. Then the MAC address written in the flash memory is used by default.

Modifying a MAC Address in the Configuration File

- (Optional) Perform this configuration when you need to modify the MAC address of a device.
- In gateway mode (the **auth-mode gateway** command is configured), some peripheral devices are configured with the MAC address of the bound gateway. If the gateway is replaced, you can use the **sysmac** command to configure the MAC address of the new gateway to be the same as that of the replaced gateway to retain the MAC address of the bound gateway on downstream devices. The **sysmac** command is valid only in gateway mode.

Command	**sysmac** *mac-address*
Parameter Description	*mac-address* : Indicates the new MAC address.
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command to configure the MAC address of a device. To make the MAC address take effect, save the configuration and restart the device.

Verification

Run the **show version slot** command to display the installation information of a line card.

Command	**show version slots** [*device-id / slot-num*]
Parameter Description	*device-id*: (Optional) Indicates the ID of a chassis (in Stacking mode, when you input a slot number, you also need to input the ID of the chassis where the module is located). *slot-num*: (Optional) Indicates the number of a slot.
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to display the online state of a module. The **Configured Module** column shows the information of the installed module. After you run the **remove configuration module** command, the installation information of the removed module is deleted from this column.

9.5 Monitoring

Displaying

Description	Command
Displays the details of a module.	**show version module detail** [*slot-num*] **show version module detail** [device-id/*slot-num*] (in Stacking mode)
Displays the online state of a module.	**show version slots** [*slot-num*] **show version slots** [*device-id/slot-num*] (in Stacking mode)
Displays the current MAC address of a device.	**show sysmac**

10 Configuring Supervisor Module Redundancy

10.1 Overview

Supervisor module redundancy is a mechanism that adopts real-time backup (also called hot backup) of the service running status of supervisor modules to improve the device availability.

In a network device with the control plane separated from the forwarding plane, the control plane runs on a supervisor module and the forwarding plane runs on cards. The control plane information of the master supervisor module is backed up to the slave supervisor module in real time during device running. When the master supervisor module is shut down as expected (for example, due to software upgrade) or unexpectedly (for example, due to software or hardware exception), the device can automatically and rapidly switch to the slave supervisor module without losing user configuration, thereby ensuring the normal operation of the network. The forwarding plane continues with packet forwarding during switching. The forwarding is not stopped and no topology fluctuation occurs during the restart of the control plane.

The supervisor module redundancy technology provides the following conveniences for network services:

1. Improving the network availability

The supervisor module redundancy technology sustains data forwarding and the status information about user sessions during switching.

2. Preventing neighbors from detecting link flaps

The forwarding plane is not restarted during switching. Therefore, neighbors cannot detect the status change of a link from Down to Up.

3. Preventing route flaps

The forwarding plane sustains forwarding communication during switching, and the control plane rapidly constructs a new forwarding table. The process of replacing the old forwarding table with the new one is unobvious, preventing route flaps.

4. Preventing loss of user sessions

Thanks to real-time status synchronization, user sessions that are created prior to switching are not lost.

10.2 Applications

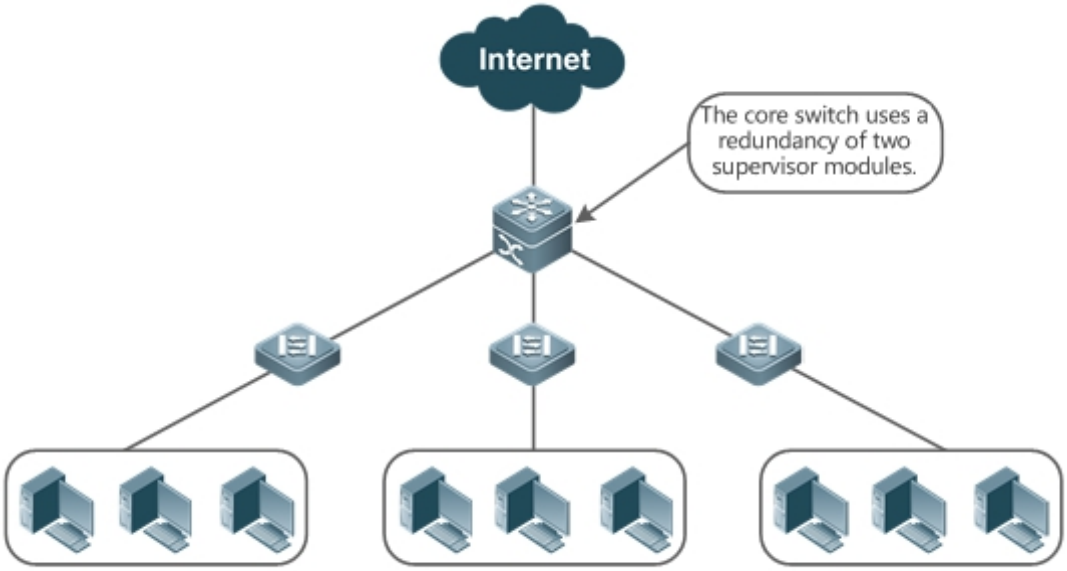
Application	Description
Redundancy of Supervisor Modules	On a core switch where two supervisor modules are installed, the redundancy technology can improve the network stability and system availability.

10.2.1 Redundancy of Supervisor Modules

Scenario

As shown in the following figure, in this network topology, if the core switch malfunctions, networks connected to the core switch break down. In order to improve the network stability, two supervisor modules need to be configured on the core switch to implement redundancy. The master supervisor module manages the entire system and the slave supervisor module backs up information about service running status of the master supervisor module in real time. When manual switching is performed or forcible switching is performed due to a failure occurring on the master supervisor module, the slave supervisor module immediately takes over functions of the master supervisor module. The forwarding plane can proceed with data forwarding and the system availability is enhanced.

Figure 11- 1



Deployment

For chassis-type devices, the system is equipped with the master/slave backup mechanism. The system supports plug-and-play as long as master and slave supervisor modules conform to redundancy conditions.

For case-type devices, each device is equivalent to one supervisor module and one line card. The Stacking composed of multiple case-type devices also has the master/slave backup mechanism.

10.3 Features

Basic Concepts

Master Supervisor Module, Slave Supervisor Module

On a device where two supervisor modules are installed, the system elects one supervisor module as active, which is called the master supervisor module. The other supervisor module functions as a backup supervisor module. When the master supervisor module malfunctions or actively requests switching, the backup supervisor module takes over the functions of the master supervisor module and becomes the new master supervisor module, which is called the slave supervisor module. In general, the slave supervisor module does not participate in switch management but monitors the running status of the master supervisor module.

Globally Master Supervisor Module, Globally Slave Supervisor Module, Globally Candidate Supervisor Module

In a Stacking system composed of two or more chassis-type devices, each chassis has two supervisor modules, with the master supervisor module managing the entire chassis and the slave supervisor module functioning as a backup. For the entire Stacking system, there are two or more supervisor modules. One master supervisor module is elected out of the supervisor modules to manage the entire Stacking system, one slave supervisor module is elected as the backup of the Stacking system, and other supervisor modules are used as candidate supervisor modules. A candidate supervisor module replaces the master or slave supervisor module and runs as the master or slave supervisor module when the original master or slave supervisor module malfunctions. In general, candidate supervisor modules do not participate in backup. To differentiate master and slave supervisor modules in a chassis from those in a Stacking system, the master, slave, and candidate supervisor modules in a Stacking system are called "globally master supervisor module", "globally slave supervisor module," and "globally candidate supervisor module" respectively. The redundancy mechanism of supervisor modules takes effect on the

globally master supervisor module and globally slave supervisor module. Therefore, the master and slave supervisor modules in the Stacking environment are the globally master supervisor module and globally slave supervisor module.

In a Stacking system composed of two or more case-type devices, each case-type device is equivalent to one supervisor module and one line card. The system elects one device as the globally master supervisor module and one device as the globally slave supervisor module, and other devices serve as globally candidate supervisor modules.

Prerequisites for Redundancy of Supervisor Modules

In a device system, the hardware and software of all supervisor modules must be compatible so that the redundancy of supervisor modules functions properly.

Batch synchronization is required between the master and slave supervisor modules during startup so that the two supervisor modules are in the same state. The redundancy of supervisor modules is ineffective prior to synchronization.

Redundancy Status of Supervisor Modules

The master supervisor module experiences the following status changes during master/slave backup:

- alone state: In this state, only one supervisor module is running in the system, or the master/slave switching is not complete, and redundancy is not established between the new master supervisor module and the new slave supervisor module.
- batch state: In this state, redundancy is established between the master and slave supervisor modules and batch backup is being performed.
- realtime state: The master supervisor module enters this state after the batch backup between the master and slave supervisor modules is complete. Real-time backup is performed between the master and slave supervisor modules, and manual switching can be performed only in this state.

Overview

Feature	Description
Election of Master and Slave Supervisor Modules	The device can automatically select the master and slave supervisor modules based on the current status of the system. Manual selection is also supported.
Information Synchronization of Supervisor Modules	In the redundancy environment of supervisor modules, the master supervisor module synchronizes status information and configuration files to the slave supervisor module in real time.

10.3.1 Election of Master and Slave Supervisor Modules

Working Principle

Automatically Selecting Master and Slave Supervisor Modules for Chassis-type Devices

Users are allowed to insert or remove supervisor modules during device running. The device, based on the current condition of the system, automatically selects an engine for running, without affecting the normal data switching. The following cases may occur and the master supervisor module is selected accordingly:

- If only one supervisor module is inserted during device startup, the device selects this supervisor module as the master supervisor module regardless of whether it is inserted into the M1 slot or M2 slot.
- If two supervisor modules are inserted during device startup, by default, the supervisor module in the M1 slot is selected as the master supervisor module and the supervisor module in the M2 slot is selected as the slave supervisor module to serve as a backup, and

relevant prompts are output.

- If one supervisor module is inserted during device startup and another supervisor module is inserted during device running, the supervisor module that is inserted later is used as the slave supervisor module to serve as a backup regardless of whether it is inserted into the M1 slot or M2 slot, and relevant prompts are output.

- Assume that two supervisor modules are inserted during device startup and one supervisor module is removed during device running (or one supervisor module malfunctions). If the removed supervisor module is the slave supervisor module prior to removal (or failure), only a prompt is displayed after removal (or malfunction), indicating that the slave supervisor module is removed (or fails to run). If the removed supervisor module is the master supervisor module prior to removal (or failure), the other supervisor module becomes the master supervisor module and relevant prompts are output.

Manually Selecting the Master and Slave Supervisor Modules

Users can manually make configuration to select the master and slave supervisor modules, which are selected based on the environment as follows:

- In standalone mode, users can manually perform master/slave switching. The supervisor modules take effect after reset.

- In Stacking mode, users can manually perform master/slave switching to make the globally slave supervisor module become the globally master supervisor module. If a Stacking system has only two supervisor modules, the original globally master supervisor module becomes the new globally slave supervisor module after reset. If there are more than two supervisor modules, one globally candidate supervisor module is elected as the new globally slave supervisor module and the original globally master supervisor module becomes a globally candidate supervisor module after reset.

Related Configuration

Manually Performing Master/Slave Switching

- By default, the device can automatically select the master supervisor module.

- In both the standalone and Stacking modes, users can run the **redundancy forceswitch** command to perform manual switching.

10.3.2 Information Synchronization of Supervisor Modules

Working Principle

- Status synchronization

The master supervisor module synchronizes its running status to the slave supervisor module in real time so that the slave supervisor module can take over the functions of the master supervisor module at any time, without causing any perceivable changes.

- Configuration synchronization

There are two system configuration files during device running: running-config and startup-config. running-config is a system configuration file dynamically generated during running and changes with the service configuration. startup-config is a system configuration file imported during device startup. You can run the **write** command to write running-config into startup-config or run the **copy** command to perform the copy operation.

For some functions that are not directly related to non-stop forwarding, the synchronization of system configuration files can ensure consistent user configuration during switching.

In the case of redundancy of dual supervisor modules, the master supervisor module periodically synchronizes the startup-config and running-config files to the slave supervisor module and all candidate supervisor modules. The configuration synchronization is also

triggered in the following operations:

1. The running-config file is synchronized when the device switches from the global configuration mode to privileged EXEC mode.
2. The startup-config file is synchronized when the **write** or **copy** command is executed to save the configuration.
3. Information configured over the Simple Network Management Protocol (SNMP) is not automatically synchronized and the synchronization of the running-config file needs to be triggered by running commands on the CLI.

Related Configuration

- By default, the startup-config and running-config files are automatically synchronized once per hour.
- Run the **auto-sync time-period** command to adjust the interval for the master supervisor module to synchronize configuration files.

10.4 Configuration

Configuration	Description and Command	
Configuring Manual Master/Slave Switching	Optional.	
	show redundancy states	Displays the hot backup status.
	redundancy forceswitch	Manually performs master/slave switching.
Configuring the Automatic Synchronization Interval	Optional.	
	redundancy	Enters the redundancy configuration mode.
	auto-sync time-period	Configures the automatic synchronization interval of configuration files in the case of redundancy of dual supervisor modules.
Resetting Supervisor Modules	Optional.	
	redundancy reload	Resets the slave supervisor module or resets both the master and slave supervisor modules at the same time.

10.4.1 Configuring Manual Master/Slave Switching

Configuration Effect

The original master supervisor module is reset and the slave supervisor module becomes the new master supervisor module.

If there are more than two supervisor modules in the system, the original slave supervisor module becomes the master supervisor module, one supervisor module is elected out of candidate supervisor modules to serve as the new slave supervisor module, and the original master supervisor module becomes a candidate supervisor module after reset.

Notes

To ensure that data forwarding is not affected during switching, batch synchronization needs to be first performed between the master and slave supervisor modules so that the two supervisor modules are in the same state. That is, manual switching can be performed only when the redundancy of supervisor modules is in the real-time backup state. In addition, to ensure synchronization completeness of configuration files, service modules temporarily forbid manual master/slave switching during synchronization. Therefore, the following

conditions need to be met simultaneously for manual switching:

- Manual master/slave switching is performed on the master supervisor module and a slave supervisor module is available.
- All virtual switching devices (VSDs) in the system are in the real-time hot backup state.
- The hot-backup switching of all VSDs in the system is not temporarily forbidden by service modules.

If devices are virtualized as multiple VSDs, manual switching can be successfully performed only when the supervisor modules of all the VSDs are in the real-time backup state.

Configuration Steps

- Optional.
- Make the configuration on the master supervisor module.

Verification

Run the **show redundancy states** command to check whether the master and slave supervisor modules are switched.

Related Commands

Checking the Hot Backup Status

Command	**show redundancy states**
Parameter Description	N/A
Command Mode	Privileged EXEC mode or global configuration mode
Usage Guide	N/A

Manually Performing Master/Slave Switching

Command	**redundancy forceswitch**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuration Example

Manually Performing Master/Slave Switching

Configuration Steps	In the VSD environment where the name of one VSD is staff, perform master/slave switching.
	FS> enable FS# show redundancy states

	Redundancy role: master Redundancy state: realtime Auto-sync time-period: 3600 s VSD staff redundancy state: realtime FS# redundancy forceswitch This operation will reload the master unit and force switchover to the slave unit. Are you sure to continue? [N/y] y
Verification	On the original slave supervisor module, run the **show redundancy states** command to check the redundancy status.
	FS# show redundancy states Redundancy role: master Redundancy state: realtime Auto-sync time-period: 3600 s VSD staff redundancy state: realtime

10.4.2 Configuring the Automatic Synchronization Interval

Configuration Effect

Change the automatic synchronization interval of the startup-config and running-config files. If the automatic synchronization interval is set to a smaller value, changed configuration is frequently synchronized to other supervisor modules, preventing the configuration loss incurred when services and data are forcibly switched to the slave supervisor module when the master supervisor module malfunctions.

Configuration Steps

- Optional. Make the configuration when the synchronization interval needs to be changed.
- Make the configuration on the master supervisor module.

Verification

- View the output syslogs to check whether timed synchronization is performed.

Related Commands

Entering the Redundancy Configuration Mode

Command	**redundancy**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Automatic Synchronization Interval of Configuration Files

Command	**Auto-sync time-period** *value*
Parameter Description	**time-period *value***: Indicates the automatic synchronization interval, with the unit of seconds. The value ranges from 1 second to 1 month (2,678,400 seconds).
Command Mode	Redundancy configuration mode
Usage Guide	Configure the automatic synchronization interval of the startup-config and running-config files in the case of redundancy of dual supervisor modules.

Configuration Example

Configuring the Automatic Synchronization Interval

Configuration Steps	In redundancy configuration mode of the master supervisor module, configure the automatic synchronization interval to 60 seconds.
	FS(config)# redundancy FS(config-red)# auto-sync time-period 60 Redundancy auto-sync time-period: enabled (60 seconds). FS(config-red)# exit
Verification	Run the **show redundancy states** command to check the configuration.
	FS# show redundancy states Redundancy role: master Redundancy state: realtime Auto-sync time-period: 3600 s

10.4.3 Resetting Supervisor Modules

Configuration Effect

Resetting only the slave supervisor module does not affect data forwarding, and the forwarding is not interrupted or user session information is not lost during reset of the slave supervisor module.

In standalone mode, running the **redundancy reload shelf** command will cause simultaneous reset of all supervisor modules and line cards in the chassis. In Stacking mode, the device of a specified ID is reset when this command is executed. If there are two or more devices in the system and the device to be reset is the device where the globally master supervisor module resides, the system performs master/slave switching.

Notes

In Stacking mode, if the supervisor modules of the system do not enter the real-time backup state, resetting the device where the globally master supervisor module resides will cause the reset of the entire Stacking system.

Configuration Steps

- Optional. Perform the reset when the supervisor modules or device runs abnormally.

Related Commands

Command	**redundancy reload** {**peer** | **shelf** [*switchid*] }
Parameter Description	**peer**: Only resets the slave supervisor module. **shelf** [*switchid*]: Indicates that the master and slave supervisor modules are set in standalone mode, and the ID of the device to be reset needs to be specified in Stacking mode.
Command Mode	Privileged EXEC mode
Usage Guide	In standalone mode, the device reset command is **redundancy reload shelf**, that is, the entire device is reset. In Stacking mode, the device reset command is **redundancy reload shelf** *switchid*, that is, the device of a specified device ID is reset.

Configuration Example

Resetting a Device in Stacking Mode

Configuration Steps	In privileged EXEC mode of the globally master supervisor module, reset the device with the ID of 2.
	FS# redundancy reload shelf 2 This operation will reload the device 2. Are you sure to continue? [N/y] y Preparing to reload device 2!
Verification	Check whether the relevant supervisor module or device is restarted.

10.5 Monitoring

Displaying

Description	Command
Displays the current redundancy status of dual supervisor modules.	**show redundancy states**

11 Configuring Package Management

11.1 Overview

Package management (pkg_mgmt) is a package management and upgrade module. This module is responsible for installing, upgrading/degrading, querying and maintaining various components of the device, among which upgrade is the main function. Through upgrade, users can install new version of software that is more stable or powerful. Adopting a modular structure, the FSOS system not only supports overall upgrade and subsystem upgrade but also supports separate upgrade of a feature package.

Component upgrade described in this document applies to both the box-type device and rack-type device. In addition, this document is for only version 11.0 and later, excluding those upgraded from earlier versions.

Protocols and Standards

N/A

11.2 Applications

Application	Scenario
Upgrading/Degrading Subsystem	Upgrade subsystem firmware like boot, kernel, and rootfs on the box-type device and rack-type device.
Upgrading/Degrading a Single Feature Package	Upgrade a single feature package on the box-type device and rack-mount device.
Auto-Sync for Upgrade	Configure the auto sync policy, range and path.

11.2.1 Upgrading/Degrading Subsystem

Scenario

After the upgrade of a subsystem firmware is complete, all system software on the device is updated, and the overall software is enhanced. Generally, the subsystem firmware of the box-type device is called main package.

The main features of this upgrade mode are as follows: All software on the device is updated after the upgrade is completed; all known software bugs are fixed. It takes a long time to finish upgrade.

Deployment

You can store the main package in the root directory of the TFTP server, download the package to the device, and then run an upgrade command to upgrade the package locally.

11.2.2 Upgrading/Degrading a Single Feature Package

Scenario

Device software consists of several components, and each component is an independent feature module. After an independent feature package is upgraded, only the feature bug corresponding to this package is fixed. Besides, this feature is enhanced with the other features unchanged.

The features of this upgrade mode are as follows: Generally, a feature package is small and the upgrade speed is high. After the upgrade is completed, only the corresponding functional module is improved, and other functional modules remain unchanged.

Deployment

You can store this package in the root directory of the TFTP server, download the package to the local device, and then complete the upgrade.

11.2.3 Auto-Sync for Upgrade

Scenario

Auto-sync upgrade aims to ensure the coordination of multiple modules within a system on a Stacking. Specifically, the upgrade firmware is pushed to all target members automatically and the software version of new members is upgraded automatically based on the auto-sync policy.

Deployment

- Configure the policy for auto-sync upgrade.
- Configure the path of firmware for auto-sync upgrade.

11.3 Features

Basic Concepts

Subsystem

A subsystem exists on a device in the form of images. The subsystems of the FSOS include:

- boot: After being powered on, the device loads and runs the boot subsystem first. This subsystem is responsible for initializing the device, and loading and running system images.
- kernel: kernel is the OS core part of the system. This subsystem shields hardware composition of the system and provides applications with abstract running environment.
- rootfs: rootfs is the collection of applications in the system.

Main Package and Rack Package

- Main package is often used to upgrade/degrade a subsystem of the box-type device. The main package is a combination package of the boot, kernel, and rootfs subsystems. The main package can be used for overall system upgrade/degradation.

Feature Package of FSOS

- The feature package of FSOS refers to a collection which enables a certain feature. When the device is delivered, all supported functions are contained in the rootfs subsystem. You can upgrade only a specific feature by upgrading a single feature package.

"Firmware" in this document refers to an installation file that contains a subsystem or feature module.

Overview

Feature	Description
Upgrading/Degrading and Managing Subsystem Components	Upgrades/degrades a subsystem.
Upgrading/Degrading and Managing Functional Components	Upgrades/degrades a functional component.
Auto-Sync for Upgrade	Ensures uniform upgrade upon member change.

11.3.1 Upgrading/Degrading and Managing Subsystem Components

Subsystem upgrade/degradation aims to upgrade the software by replacing the subsystem components of the device with the subsystem components in the firmware. The subsystem component contains redundancy design. Subsystems of the device are not directly replaced with the subsystems in the package during upgrade/degradation in most cases. Instead, subsystems are added to the device and then activated during upgrade/degradation.

Working Principle

Upgrade/Degradation

Various subsystems exist on the device in different forms. Therefore, upgrade/degradation varies with different subsystems.

- boot: Generally, this subsystem exists on the norflash device in the form of images. Therefore, upgrading/degrading this subsystem is to write the image into the norflash device.
- kernel: This subsystem exists in a specific partition in the form of files. Therefore, upgrading/degrading this subsystem is to write the file.
- rootfs: Generally, this subsystem exists on the nandflash device in the form of images. Therefore, upgrading/degrading this subsystem is to write the image into the nandflash device.

Management

Query the subsystem components that are available currently and then load subsystem components as required.

Each subsystem component contains redundancy design. During the upgrade/degradation:

- boot: The boot subsystem always contains a master boot subsystem and a slave boot subsystem. Only the master boot subsystem is involved in the upgrade, and the slave boot subsystem serves as the redundancy backup all along.
- kernel: as the kernel subsystem contains at least one redundancy backup. More redundancy backups are allowed if there is enough space.
- rootfs: The rootfs subsystem always contains a redundancy backup.

The boot component is not included in the scope of subsystem management due to its particularity. During upgrade of the kernel or rootfs subsystem component, the upgrade/degradation module always records the subsystem component in use, the redundant subsystem component, and management information about various versions.

Relevant Configuration

Upgrade

- Store the upgrade file on the local device, and then run the **upgrade** command for upgrade.

11.3.2 Upgrading/Degrading and Managing Functional Components

Working Principle

In fact, upgrading a feature is replacing feature files on the device with the feature files in the package.

Managing feature components is aimed at recording the information of feature components by using a database. In fact, installing, displaying and uninstalling a component is the result of performing the Add, Query and Delete operation on the database.

Relevant Configuration

Upgrade

- Store the upgrade file on the local device, and then run the **upgrade** command for upgrade.

11.3.3 Auto-Sync for Upgrade

Working Principle

Auto-sync upgrade aims to ensure the coordination of multiple modules (line cards and chassis) within a system. Specifically, the upgrade firmware is pushed to all target members automatically and the software version of new members is upgraded automatically based on the auto-sync policy.

There are three policies available.

None: No auto-sync upgrade.

Compatible: Performs auto-synchronization based on the sequential order of versions.

Coordinate: Synchronizes with the version based on the firmware stored on the supervisor module.

Auto-sync is performed in the following scenarios:

If no upgrade target is specified, the firmware is pushed to all matching members(including line cards and chassis) for auto-sync.

Every member is checked when the device is restarted and auto-sync is performed accordingly.

Every new member is checked when added into the system and auto-sync is performed accordingly.

Management

Auto-upgrade policy, range and path should be configured in advance.

Relevant Configuration

Configuring Auto-Sync Policy

To perform upgrade as expected, check the configuration in advance, such as the path.

If some line cards are not checked for upgrade because the system is not configured with auto-sync policy . You can upgrade them manually.

11.4 Configuration

Configuration	Description and Command	
Upgrading/Degrading a Firmware	The basic function of the configuration is installing and upgrading/degrading a subsystem firmware, and feature package.	
	upgrade *url* [**force**]	*url* is a local path where the firmware is stored. This command is used to upgrade the firmware stored on the device.
	upgrade download tftp:/ *path* [**force**]	*path* is the path of the firmware on the server. This command is used to download a firmware from the server and upgrade the package automatically.
	upgrade download ftp://*path* [***force***]	*path* is the path of the firmware on the server. This command is used to download a firmware from the server and upgrade the package automatically.
Auto-Sync for Upgrade	(Optional) Configures auto-sync policy.	
	upgrade auto-sync policy [**none** | **compatible** | **coordinate**]	Configures the auto-sync policy.
	upgrade auto-sync range [**chassis** | **vsu**]	Configures the auto-sync range.
	upgrade auto-sync package *url*	Configures the auto-sync path.

11.4.1 Upgrading/Degrading a Firmware

Configuration Effect

Available firmwares include the main package, rack package, and various feature packages.

- After the upgrade of the main package is complete, all system software on the line card is updated, and the overall software is enhanced.
- After an independent feature package is upgraded, only the feature bug corresponding to this package is fixed. Besides, this feature is enhanced, with other features remain unchanged.

Generally a main package is released to upgrade a box-type device.

Notes

N/A

Configuration Steps

Upgrading the Main Package for a Single Device

- Optional configuration. This configuration is required when all system software on the device needs to be upgraded.

- Download the firmware to the local device and run the **upgrade** command.

Generally a main package is pushed to upgrade a box-type device.

Upgrading Each Feature Package

- Optional configuration. The configuration is used to fix bugs of a certain feature and enhance the function of this feature.
- Download the firmware to the local device and run the **upgrade** command.

Verification

- After upgrading a subsystem component, you can run the **show upgrade history** command to check whether the upgrade is successful.
- After upgrading a feature component, you can run the **show component** command to check whether the upgrade is successful.

Commands

Upgrade

Command	**upgrade** *url* [**force**]
Parameter Description	**force** indicates forced upgrade.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Command	**upgrade download tftp:/***path* [**force**]
Parameter Description	**force** indicates forced upgrade.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Command	**upgrade download ftp:/***path* [**force**]
Parameter Description	**force** indicates forced upgrade.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Displaying the Firmware Stored on the Device

Command	**show upgrade file** *url*
Parameter Description	*url* indicates the path of the firmware in the device file system.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Displaying Upgrade History

Command	**show upgrade history**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Displaying the Feature Components Already Installed

Command	**show component**
Parameter Description	[*component_name*]: component name When this parameter value is N/A, the command is used to display all components already installed on the device and basic information of these components. When this parameter value is not N/A, the command is used to display detailed information of the corresponding component, check whether the component is intact, and check whether this component works properly.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuration Example

Example of Upgrading a Subsystem Firmware on the Box-Type Device

Network Environment	Before the upgrade, you must copy the firmware to the device. The upgrade module provides the following solutions. ● Run some file system commands like **copy tftp** to copy the firmware on the server to the device file system, and then run the **upgrade *url*** command to upgrade the firmware in the local file system. ● Run the **upgrade download tftp://*path*** command directly to upgrade the firmware file stored on the tftp server. ● Run the **upgrade download ftp://*path*** command directly to upgrade the firmware file stored on the ftp server.
Configuration	● Run the upgrade command.

Steps	● After upgrading the subsystem, restart the device.
Verification	● Check the system version on the current device. If the version information changes, the upgrade is successful.

Example of Upgrading a Feature Package on the Box-Type Device

Network Environment	Before the upgrade, you must copy the firmware to the device. The upgrade module provides the following solutions. ● Run some file system commands like **copy tftp** to copy the firmware on the server to the device file system, and then run the **upgrade *url*** command to upgrade the firmware in the local file system. ● Run the **upgrade download tftp://*path*** command directly to upgrade the firmware file stored on the tftp server. ● Run the **upgrade download ftp://*path*** command directly to upgrade the firmware file stored on the ftp server.
Configuration Steps	● Run the **upgrade** command. ● Check whether the device needs to be restarted based on the prompt displayed after the upgrade.
Verification	● Check the version of the feature component on the current device. If the version information changes, the upgrade is successful.
	FS# show component Package :sysmonit Version:1.0.1.23cd34aa Build time: Wed Dec 7 00:58:56 2011 Size:12877 Install time :Wed Mar 5 14:23:12 2012 Description:this is a system monit package Required packages: None ---------- package:bridge Version: 2.3.1.1252ea Build time: Wed Dec 7 00:54:56 2011 Size:26945 Install time : Wed Mar 19:23:15 2012 Description:this is a bridge package Required packages: None

Common Errors

If an error occurs during the upgrade, the upgrade module displays an error message. The following provides an example:

```
Upgrade info [ERR]
        Reason:creat config file err(217)
```

The following describes several types of common error messages:

- Invalid firmware: The cause is that the firmware may be damaged or incorrect. It is recommended to obtain the firmware again and perform the upgrade operation.
- Firmware not supported by the device: The cause is that you may use the firmware of other devices by mistake. It is recommended to obtain the firmware again, verify the package, and perform the upgrade operation.

11.4.2 Auto-Sync for Upgrade

Configuration Effect

Auto-sync policy, range and path is configured.

Notes

N/A

Configuration Steps

Configuring Auto-Sync Policy

Run the **upgrade auto-sync policy command** to configure the auto-sync policy. There are three modes available:

None: No auto-sync upgrade.

Compatible: Performs auto-synchronization based on the sequential order of versions.

Coordinate: Synchronizes with the version based on the firmware stored on the supervisor module.

Configuring Auto-Sync Range

Run the **upgrade auto-sync range** command to configure the auto-sync range. There are two ranges available:

chassis: Performs auto-sync on a chassis.

vsd: Performs auto-sync in the Stacking system.

Configuring Auto-Sync Path

Every time the system is upgraded, the firmware path is recorded automatically for later auto-sync upgrade. Alternatively, use the **upgrade auto-sync package** command to set a path.

Verification

Run the **upgrade auto-sync** command to check the configuration.

Commands

Configuring Auto-Sync Policy

command	**upgrade auto-sync policy** [**none** | **compatible** | **coordinate**]
Parameter Description	**none**: No auto-sync upgrade **compatible**: Performs auto-synchronization based on the sequential order of versions. **coordinate**: Synchronizes with the version based on the firmware stored on the supervisor module.
Command	Privileged EXEX mode

Mode	
Usage Guide	It is recommended to set **coordinate**.

Configuring Auto-Sync Range

command	**upgrade auto-sync range** [VSU]
Parameter Description	**VSU;** Performs auto-sync in the Stacking system.
Command Mode	Privileged EXEX mode
Usage Guide	It is recommended to set **Stacking** to ensure uniformity.

Configuring Auto-Sync Path

command	**upgrade auto-sync package** *url*
Parameter Description	*url* indicates the path of the firmware in the device file system.
Command Mode	Privileged EXEX mode
Usage Guide	The path is not set generally.

Configuration Example

Configuring Auto-Sync Policy

Configuration Steps	Configure the auto-sync policy.
	FS# upgrade auto-sync policy coordinate Upgrade auto-sync policy is set as coordinate
Verification	Check the auto-sync policy.

Configuring Auto-Sync Range

Configuration Steps	Configure the auto-sync range.

	FS# upgrade auto-sync range Stacking Upgrade auto-sync range is set as Stacking.
Verification	Check the auto-sync range.

Common Errors

url is not valid.

11.5 Monitoring

Displaying

Function	Command
Displays all components already installed on the current device and their information.	**show component** [*component _name*]
Displays the upgrade history.	**show upgrade history**

12 Configuring PoE

12.1 Overview

Power over Ethernet (PoE) is a technology that can transmit electricity and data to devices through twisted pairs over Ethernet. This technology enables various devices such as VOIP, WIFI APs, network cameras, hubs and computers to obtain electricity through twisted pairs.

The largest distance that can be powered by a PoE switch is 100 m as defined by the standards. A PoE switch can collect statistics about the power supplies of all ports and the entire device, which can be displayed by a query command.

Protocols and Standards

Currently, PoE complies with the IEEE 802.3af and IEEE 802.3at standards. The following table lists the main characteristics of and differences between the two standards:

Parameter	802.3af	802.3at
Available Power for PD	12.95 W	25.50 W
Maximum Power Provided by PSE	15.4 W	30 W
Voltage Range of PSE	44.0-57.0 V	50.0-57.0 V
Voltage Range of PD	37.0-57.0 V	42.5-57.0 V
Maximum Resistance of Network Cables	20 Ω	12.5 Ω
Power Management Mode	Classify power levels during line initialization.	Classify the power supply into 4 levels during line initialization or dynamically adjust the power supply in the unit of 0.1 W.
Supported Cables	Cat-3 or Cat-5 twisted pairs	Cat-5 twisted pairs

12.2 Applications

Application	Description
PoE Power Supply Scenario	In the scenario, a PoE switch powers powered devices (PDs) and implements data exchange.

12.2.1 PoE Power Supply Scenario

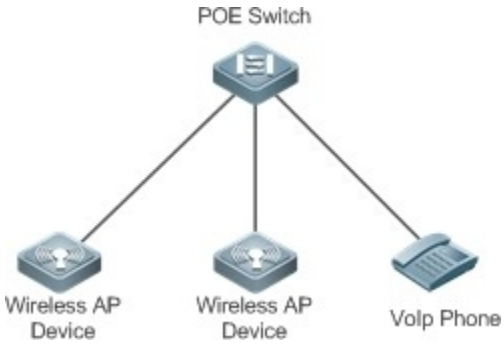
Scenario

In a PoE system set up with a PoE switch, the PoE switch combines the PoE power supply with the PSE. In addition to providing normal network data exchange, the PoE switch also provides the power supply function. The main PDs in the system include the APs of a WLAN and VoIP telephones.

The PoE switch provides power management, including power supply enabling for ports, power supply priority management, over-temperature protection for ports, and power supply status query for devices and ports.

A PoE switch enabling PoE+ supports LLDP correlation with PDs for dynamically managing the power supply power of ports.

Figure 12-1



With constantly increasing functions, terminals consume more and more power. The power consumption of certain terminals exceeds the maximum power 30 W that can be provided by POE+. Consequently, these terminals have to return to the traditional power supply mode.

Deployment

- By default, a PoE switch port is enabled with the power supply function and can start the power supply after detecting an accessed device.
- If the total power of the PoE system is insufficient, you can manually configure the power priority for ports to ensure that the ports are powered first.
- LLDP correlation is enabled by default.

12.3 Features

Basic Concepts

PoE Power Supply

The PoE power supply powers the entire PoE system and is classified into external and internal power supplies. Cassette PoE switches of FS often have internal power supplies and certain products also support external power supplies. External power supplies are called RPS.

PSE

Power Sourcing Equipment (PSE) queries and detects PDs on PoE ports, classifies PDs into different levels, and supplies power for the PDs. After detecting that a PD is removed, the PSE stops supplying power.

PD

PDs are devices powered by PSE and are classified into standard PDs and non-standard PDs. Standard PDs are PDs that comply with the IEEE 802.3af and 802.3at standards. Common non-standard PDs include non-standard PDs with featured resistance, Cisco pre-standard PDs, PDs supporting only signal cable power supplies, and PDs supporting only idle cable power supplies. FS switches use signal cable power supplies and do not support PDs supporting only idle cable power supplies.

When being powered by a PoE power supply, a PD can also connect to other power supplies for redundant backup of the power supplies.

Overview

Feature	Description
Power Supply Management for the PoE System	Manages the power supply policies of the system, such as the power supply mode and disconnection detection mode, and supports monitoring on the power supply of the PoE system, such as the system alarm limit and trap sending enabling/disabling.
Power Supply Management for PoE Ports	Manages the power supply policies of PoE ports, such as port enabling and power supply prioritization.
Auxiliary PoE Power Supply Functions	Provides auxiliary power supply management functions for the system, such as the power alarm limit of the system and PD descriptor configuration of ports.

12.3.1 Power Supply Management for the PoE System

Working Principle

Power supply management for the PoE system supports:

You can switch the power supply mode (namely, the method for allocating power for PDs connected to the PoE switch). The PoE switch supports the auto mode, energy-saving mode and static mode for power supply management.

In the auto mode, the system allocates power based on the detected PD classes and types on ports. A PoE switch allocates power for PDs of classes 0 to 4 as follows: 15.4 W for Class0, 4 W for Class1, 7 W for Class2, 15.4 W for Class3, and 30 W for Class4. In this mode, even if there is a device of Class3 that consumes only 11 W, the PoE switch allocates a power of

15.4 W for the port connecting to this device. The auto mode is the default power supply management mode of the PoE switch.

In the energy-saving mode, the PoE switch dynamically adjusts allocated power based on actual consumption of PDs. In this mode, the PoE switch can power more PDs, but the power fluctuation of certain PDs may affect the power supply of other PDs. The energy-saving mode is an optional mode of the PoE switch. If the switch does not support this mode, corresponding prompt information will be displayed during configuration.

In the energy-saving mode, the PoE switch calculates the power consumption of the system based on the actual power consumption of the PDs. If certain PDs have a large power fluctuation in this mode, overload may occur on the PoE switch, which causes damage of the PoE device. The PoE switch provides a command for setting the reserved power of the PoE system to ensure that the PoE switch always has "rich" power and that the consumed power will not exceed the limit of the PoE switch.

In the static mode, the switch allocates power to each port as configured. If the power is insufficient, it will be allocated to each port based on port ID from low to high. If the switch does not support this mode, a prompt message will be displayed. The PoE switch provides uninterruptible power supply during hot startup. When the system is restarted, PDs that are being powered will not be powered

off during hot startup of the PoE switch. After the hot startup is completed, the system recovers the status saved in the configuration file.

FS devices provide PoE-compatible commands to support non-standard PoE devices.

Related Configuration

Configuring the Power Supply Management Mode

By default, the power supply management mode is **energy-saving**.

You can run the **poe mode** { **auto** | **energy-saving** | **static** } command to configure the power supply management mode. Since different power management modes provide different methods for allocating power to PDs, mode switching may affect the PDs that can be powered.

Configuring Reserved Power

By default, the reserved power is 0.

You can run the poe reserve-power int command to configure the reserved power. When the system switches to the energy-saving mode, this function takes effect. Configuring Uninterruptible Power Supply During Hot

Startup

By default, the system disables the uninterruptible power supply function during hot startup.

You can run the **poe uninterruptible-power** command to enable the uninterruptible power supply function during hot startup. The configuration takes effect after being saved. During hot startup of the system, the PoE system supplies stable power for PDs.

Configuring Compatibility with Non-standard PDs

By default, non-PoE devices are not compatible.

You can run the **poe legacy** command to configure compatibility with non-standard PDs.

12.3.2 Power Supply Management for PoE Ports

Working Principle

Power supply management for the PoE ports supports:

You can enable or disable the PoE function for ports.

You can configure power supply priorities for ports of a PoE switch. The priorities are Critical, High and Low in a descending order. In the auto and energy-saving modes, ports with high priorities are powered first. When the system power of the PoE switch is insufficient, ports with low priorities are powered off first. The default priorities of all ports are low.

Ports with the same priority are sorted by the port number. A smaller port number means a higher priority. For example, the priority of port 1 is higher than those of ports 2 and 3.

For ports with the same priority, newly inserted ports do not preempt the power of ports that are being powered. For ports with different priorities, ports with higher priorities can preempt the power of ports with lower priorities.

You can configure a switch to manage the power-on/off of a port based on time ranges. The time range can be configured by the **time-range** command in the global configuration mode.

You can configure the maximum power of a port to restrict the maximum output power of the port. In the auto and energy-saving modes, configuring the maximum power can restrict the maximum output power of ports. When the power of a port is greater than the configured maximum power for 10 seconds, the port is powered off, the device connected to the port is powered off, a log indicates power overload for the port, and the LED indicator of the port is displayed in yellow. 10 seconds later, the port is powered on again. If the power of the port is still greater than the maximum power for 10 seconds, the port will be powered off again. This process repeats constantly.

Related Configuration

Enabling the Power Supply Function for a Port

By default, ports are enabled with the PoE power supply function.

You can run the **no poe enable** command to disable the PoE function for ports.

Configuring Power Supply Priorities for Ports

By default, the power supply priorities of ports arelow.

You can run the **poe priority** { **low | high | critical** } command to configure the power supply priority of a port. If the power is insufficient, ports with high priorities preempt the power of ports with low priorities. In this case, certain ports with low priorities may be powered off due to insufficient power.

Configuring the Power Allocated to a Port

By default, the power allocated to a port is 0.

You can run the **poe alloc-power** *int* command to configure the power allocated to a port.

Configuring the Maximum Power for Ports

By default, there is no power restriction on ports.

You can run the **poe max-power** *int* command to configure the maximum power for a port. In the static mode, the maximum power configured for a port does not take effect. If the maximum power configured for a port is 15.4 W but the power consumed by the PD connected to the port is greater than 1.1 times of the maximum power, over-current occurs on the port.

Configuring the Regular Power-off Function for a Port

By default, ports do not have the regular power-off function.

You can run the **poe power-off time-range** *range-name* command to configure the regular power-off function for a port. In the clock period specified by **time-range**, the PoE switch does not supply power for connected PDs.

Configuring Compatibility with Non-standard PDs

By default, non-PoE devices are not compatible.

You can run the **poe legacy** command to configure compatibility with non-standard PDs.

12.3.3 Auxiliary PoE Power Supply Functions

Working Principle

The PoE MIB (RFC3621) standard provides **pethMainPseUsageThreshold** to set the power alarm threshold of the system.

PoE switches provide the CLI to set this value. The function of this CLI is the same as **pethMainPseUsageThreshold MIB**, which is setting the power alarm limit of the system. If the **pethNotificationControlEnable** switch is enabled in the MIB, the MIB receives notifications on the alarm power.

In actual application, whether the system sends trap notifications in case of power change and port power-on/off needs to be controlled. The **pethNotificationControlEnable** item is provided in the PoE standard MIB RFC3621, which is used to set whether to send trap notifications.

In actual application, you often have to record the PD connected to a specific PoE port. RFC3621 provides

pethPsePortType to set the PD description. PoE switches provide the CLI to set this value.

Related Configuration

Configuring the Power Alarm Threshold of the System

By default, the power alarm threshold of the system is 90.

You can run the **poe warning-power** *int* command to configure the power alarm threshold of the system.

Configuring the Trap Notification Sending Switch of the System

By default, the system disables sending of trap notifications.

You can run the **poe notification-control enable** command to enable trap notification sending of the system.

Configuring the PD Descriptor of a Port

By default, a port has no PD descriptor.

You can run the **poe pd-description** *pd-name* command to configure the PD descriptor for the port.

12.4 Configuration

Configuration	Description and Command	
Configuring Power	(Mandatory) It is used to manage the PoE power supply of the system.	
Supply of the PoE System	**poe mode**	Configures the power supply management mode.
	poe uninterruptible-power	Configures uninterruptible power supply during hot startup.
	reserve-power	Configures the reserved power.
Configuring Power Supply on PoE Ports	(Mandatory) It is used to manage the PoE power supply of a specific port.	
	poe enable	Enables the power supply function for a port.
	poe priority	Configures the power supply priority for the port.

	poe max-power	Configures the maximum power allocated to the port.
	poe power-off time-range name	Configures the regular power-off function for the port.
	poe legacy	Configures compatibility with non-standard PDs.
Configuring Auxiliary PoE Power Supply Functions	(Optional) It facilitates PoE system management.	
	poe warning-power	Configures the power alarm threshold of the system.
	poe notification-control enable	Configures the trap notification sending switch of the system.
	poe pd-description	Configures the PD descriptor of a port.
Enabling the LLDP Classification	(Optional) It is used to manage the LLDP classification between the PoE and PDs.	
	poe class-lldp enable	Uses the LLDP classification.

12.4.1 Configuring Power Supply of the PoE System

Configuration Effect

- Configure **mode** and change the power allocation mode for PDs. In the auto mode, power is allocated based on PD classes. In the energy-saving mode, power is allocated based on actual consumption. In the static mode, power is allocated based on the **alloc-power** command.
- Configure **reserve-power** to reserved power.
- Configure **uninterruptible-power**, which maintains the PoE power supply function during hot startup.

Configuration Steps

Configuring the Power Supply Management Mode

- (Mandatory) It is energy-saving by default.
- Switch the power supply management mode, power off all PoE ports and then power on them based on the new power supply management mode.
- If the port requires stable power supply, configure the **static** mode. To ensure that the PoE switch powers more ports, you can use the energy-saving mode and allocate power to the ports based on actual power consumption.
- Support the global configuration and port-based configuration.

Configuring the Reserved Power of the System

- (Mandatory) It takes effect only in the energy-saving mode.
- Set the system reserved power command, which takes effect only when the current PoE switch is in the energy-saving mode.
- Setting the reserved power in the energy-saving mode may cause power-off of ports that have been powered on.
- Support the global configuration.

Configuring Uninterruptible Power Supply During Hot Startup

- (Optional) It is disabled by default.
- In actual application, switches may need to be upgraded. For example, after the management software is upgraded, a PoE switch needs to be restarted. However, many PDs are normally powered by the PoE switch in this case. Direct restart may cause power-off and then power-on of the PDs, that is, the PDs may be interrupted for a period of time.
- After this function is enabled or disabled, the configuration will take effect upon next reset only after being saved.
- Support the global configuration.

Verification

View the power supply status of the PoE system to check whether the configuration is correct and whether the configuration takes effect for the power supply.

Related Commands

Configuring the Power Supply Management Mode

Command	**poe mode** { **auto** | **energy-saving** | **static** }
Parameter Description	{ **auto** | **energy-saving** | **static** }: Indicates the auto, energy-saving and static mode.
Command Mode	Global configuration mode
Usage Guide	-

Configuring Reserved Power of System

Command	**poe reserve-power** *int*
Parameter Description	*Int*: Indicates the percentage of the reserved power. <0~50>
Command Mode	Global configuration mode
Usage Guide	-

Configuring Uninterruptible Power Supply During Hot Startup

Command	**poe uninterruptible-power**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	-

Configuration Example

Configuring the Power Supply Management Policies for the System

Scenario	● Each of the connected PDs consumes low power, but the number of the connected PDs is large and all ports are occupied. ● The PDs should not be disconnected during hot startup.
Configuration Steps	● Switch the mode to the energy-saving mode. ● Sets the reserved power to 20%. ● Support uninterruptible power supply during hot startup.
	FS# configure terminal FS(config)# poe mode energy-saving FS(config)# poe reserve-power 20 FS(config)# poe uninterruptible-power FS(config)# exit FS# write
Verification	Run the **show poe powersupply** command to view the configurations and the power supply information.
	FS#show poe powersupply Device member : 1 Power management : energy-saving PSE total power : 125.0 W PSE total power consumption : 15.0 W PSE available power : 100.0 W PSE total remain power : 110.0 W PSE peak power : 15.0 W PSE average power : 13.9 W

	PSE powered port : 1 PSE disconnect mode : dc PSE reserve power : 20 PSE available reserve power : 25.0 W PSE warning power : 90

12.4.2 Configuring Power Supply on PoE Ports

Configuration Effect

- Configure **time-range** to ensure that ports are not powered off within the time-range.
- Configure **priority** for ports. If the power is insufficient, ports with high priorities can preempt the power of ports with low priorities but ports with the same priority do not preempt the power from each other.
- Configure **legacy** to configure compatibility with non-standard PD devices.
- Configure **max-power** for ports. If the power consumed by a port exceeds 1.1 times of the max-power, the power is powered off. After a penalty period of 10 seconds, the port is powered on again.
- Configure **alloc-power** to allocate power in static mode.

Configuration Steps

Enabling the Power Supply Function for a Port

- (Mandatory) It is enabled by default.
- To enable or disable the PoE function for a port, you must enable or disable the power supply function of the port.
- By default, the PoE function of the port for connecting a convergence switch is enabled and the PoE function for a core switch is disabled.
- If you run the **interface range** command to configure the PoE function for ports in batches, the enabling or disabling of the PoE function for a port may affect the global power supply management because the **range** command is configured for ports one after another. Therefore, ports may be powered on and then off during the configuration process, which is normal.
- Support port-based configuration.

Configuring the Regular Power-off Function for a Port

- Optional.
- When the power supply function is enabled for a port, configure **time-range** and then manage the power-on/off of the port based on the period of time specified by range-name.
- The accuracy of the regular power supply function for a PoE port is one minute and 30 seconds.
- Configure the regular power-off function for a PoE port. **range-name** indicates the name of the time range, consisting of up to 32 characters.
- Support port-based configuration.

Configuring the Power Supply Priority for a Port

- (Mandatory) The priority of a port is low by default.
- In scenarios with insufficient power, in order to supply stable power for certain ports, you can configure priorities for the ports.
- Support the global configuration and port-based configuration.

Configuring Compatibility with Non-standard PDs

- (Optional) It is not supported by default.
- If connected PDs do not meet the PoE standard, the function of being compatible with non-standard PDs can be enabled to supply power for the PDs.
- Running this command for ports not connected to PDs may cause burning of peer devices due to incorrect power-on. Therefore, you must run this command when PDs are connected to ports.
- The powered-on non-standard DSs will not be powered off if the **no poe legacy** command is run.
- If this command is not configured, non-standard PDs connected will not be powered on and the system will not display any prompt information.

Configuring the Maximum Power Allocated to a Port

- (Optional) There is no maximum power restriction on a port by default.
- This command may take effect in the auto and energy-saving modes.
- When max-power is set to 0, the port is powered off and will not be powered on again.
- Only support 802.3af PoE switch. The max-power ranges from 0 to 15.4.
- Configure the maximum power of a port. The maximum power cannot exceed 1.1 times of the configured power to reduce the impact of high power consumed by a single port on power management.
- Support port-based configuration.

Configuring the Power Allocated to a Port

- (Optional) This command is active only in static mode.
- This command is required static mode.
- If the system power is insufficient when power allocated to a port is configured, the system prompts that the configuration fails.
- When alloc-power is set to 0 in static mode, the port is powered off and will not be powered on again.
- Support port-based configuration.

Verification

View the PoE information of PoE ports to check whether the configuration is correct and whether the configuration takes effect for the power supply.

Related Commands

Enabling the Power Supply Function for a Port

Command	**poe enable**

Parameter Description	-
Command Mode	Interface configuration mode
Usage Guide	-

Configuring the Regular Power-off Function for a Port

Command	**poe power-off time-range** *name*
Parameter Description	*name***:** Indicates the descriptor of **time-range**.
Command Mode	Interface configuration mode
Usage Guide	-

Configuring Power Supply Priorities for Ports

Command	**poe priority { low | high | critical }**
Parameter Description	{ low | high | critical }: Indicates the priority. The value can be **Low**, **High** or **Critical**.
Command Mode	Interface configuration mode
Usage Guide	-

Configuring Compatibility with Non-standard PDs

Command	**poe legacy**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	-

Configuring the Maximum Power Allocated to a Port

Command	**poe max-power** *int*
Parameter Description	*Int*: Indicates the maximum power, ranging from 0 to 30-95 W. The value ranges from 0 to 15.4 for a system supporting only 802.3af.
Command Mode	Interface configuration mode
Usage Guide	-

Configuring the Power Allocated to a Port

Command	**poe alloc-power int**
Parameter Description	*Int*: Indicates the maximum power, ranging from 0 to 30 W.
Command Mode	Interface configuration mode
Usage Guide	-

Configuration Example

Configuring the Power Supply Management Policies for a Port

Scenario	● The port g0/1 requires a stable power supply not affected by the network environment. ● The power is powered off from 8:00 to 12:00 and is powered on in other time. ● The maximum power of the port does not exceed 17 W.
Configuration Steps	● Set the priority of the port g0/1 to critical. ● Configure **time-range** and associate the port **time-range** configuration of the PoE. ● Set the maximum power of the port g0/1 to 15.4 W.
	FS# configure terminal FS(config)# time-range poe-time FS(config-time-range)# periodic daily 8:00 to 12:00 FS(config-time-range)# exit FS(config)# interface gigabitEthernet 0/1 FS(config-if)# poe power-off time-range poe-time FS(config-if)# poe priority critical FS(config-if)# poe max-power 15.4

Verification	Run the **show poe interface gigabitEthernet 0/1** command to view the configurations and the power supply information.
	FS#show poe interface gigabitEthernet 0/1 Interface : gi0/1 Power enabled : enable Power status : on Max power : 15.4W Allocate power : N/A Current power : 14.8 W Average power : 14.8 W Peak power : 14.8 W LLDP requested power : N/A LLDP allocated power : N/A Voltage : 53.5 V Current : 278 mA PD class : 4 Trouble cause : None Priority : critical Legacy : off Power-off time-range : poe-time Power management : auto

11.4.3 Configuring Auxiliary PoE Power Supply Functions

Configuration Effect

- Configure **warning-power** to display a warning when the power used by the system exceeds the alarm threshold.
- Configure **notification-control** to control whether the system sends trap notifications in case of power change and port power-on/off.
- Configure **pd-description** to identify the PD connected to a port.

Configuration Steps

Configuring the Power Alarm Threshold of the System

- (Mandatory) It is 99 by default, which is consistent with that specified in the RFC3621 MIB.
- Configure the power alarm threshold of the system. When the power used by the system exceeds the threshold, the system displays a warning.

- If you set the power alarm threshold of the system by using **pethMainPseUsageThreshold** provided by the PoE MIB, the CLI will be configured as well.
- Support the global configuration.

Configuring the Trap Notification Sending Switch of the System

- (Mandatory) It is disabled by default.
- When trap notification sending is enabled, trap notifications will be sent when the alarm power notification and power on/off notification of the system are enabled and disabled.
- This CLI command can control only sending of trap notifications defined in the RFC3621 and does not take effect for trap notifications not defined in the RFC3621.
- When sending of trap notifications defined in the RFC3621 is enabled, a notification is sent if the alarm power changes from being lower than or equal to the system power to being higher than the system power. If the alarm power is always higher than the system power, no trap notification will be sent. If the alarm power changes from being higher than or equal to the system power to being lower than the system power, no trap notification will be sent if the alarm power is always lower than the system power subsequently.
- Support the global configuration and port-based configuration.

Configuring the PD Descriptor of a Port

- (Optional) A port has no PD descriptor by default.
- Configure the PD descriptor of a port to easily identify the PD connected to the port.
- If you set the PD by using **pethPsePortType** provided by the MIB, the CLI will be configured as well.
- Support port-based configuration.

Verification

Check whether alarm information is output when the power used by the system fluctuates on the alarm power threshold to check whether the alarm power configuration takes effect.

Connect the PoE to the SNMP server and power on and off a port to check whether corresponding trap notifications are received from the server and check whether the trap configuration takes effect.

View the PoE information of the port to check whether the PD descriptor of the port is correct.

Related Commands

Configuring the Power Alarm Threshold of the System

Command	**poe warnig-power** *int*
Parameter Description	*Int*: Indicates the alarm power percentage, ranging from 0 to 99.
Command Mode	Global configuration mode
Usage Guide	-

Configuring the Trap Notification Sending Switch of the System

Command	**poe notification-control enable**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	-

Configuring the PD Descriptor of a Port

Command	**poe pd-description** *pd-name*
Parameter Description	*pd-name*: Indicates the PD descriptor name. The parameter value is a string and supports a maximum of 32 characters.
Command Mode	Interface configuration mode
Usage Guide	-

Configuration Example

Configuring the Power Supply Management Policies for the System

Scenario	● When the system power exceeds 80%, a warning should be displayed. ● When a port is powered on or off, trap notifications should be sent. ● PDs connected to ports can be identified.
Configuration Steps	● Set the alarm power threshold of the system to 80%. ● Enable the trap notification sending switch of the system. ● Configure the PD descriptor of the port g0/1 as ap220.
	FS# configure terminal FS(config)# poe poe warnig-power 80 FS(config)# poe notification-control enable FS(config)# interface gigabitEthernet 0/1 FS(config-if)# poe pd-description ap220
Verification	Run the **show running-config** command to view the configurations and power supply information.

12.5 Monitoring

Displaying

Description	Command
Displays the PoE configuration and status of a specified port.	**show poe interface**
Displays the PoE status orconfigurations of all ports.	**show poe interfaces**
Displays the power supply status of the current PoE system.	**show poe powersupply**

Chapter2 Ethernet Switching Configuration

1. Configuring Interfaces
2. Configuring MAC Addresses
3. Configuring Aggregated Port
4. Configuring VLAN
5. Configuring Super VLAN
6. Configuring MAC VLAN
7. Configuring Protocol VLAN
8. Configuring Private VLAN
9. Configuring Voice VLAN
10. Configuring MSTP
11. Configuring GVRP
12. Configuring LLDP
13. Configuring QinQ
14. Configuring ERPS

1 Configuring Interfaces

1.1 Overview

Interfaces are important in implementing data switching on network devices. FS devices support two types of interfaces: physical ports and logical interfaces. A physical port is a hardware port on a device, such as the 100M Ethernet interface and gigabit Ethernet interface. A logical interface is not a hardware port on the device. A logical interface, such as the loopback interface and tunnel interface, can be associated with a physical port or independent of any physical port. For network protocols, physical ports and logical interfaces serve the same function.

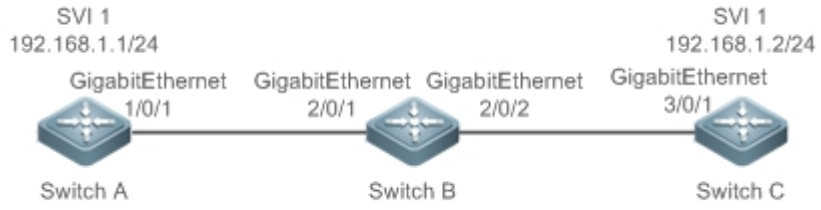
1.2 Applications

Application	Description
L2 Data Switching Through the Physical Ethernet Interface	Implement Layer-2 (L2) data communication of network devices through the physical L2 Ethernet interface.
L3 Routing Through the Physical Ethernet Interface	Implement Layer-3 (L3) data communication of network devices through the physical L3 Ethernet interface.

1.2.1 L2 Data Switching Through the Physical Ethernet Interface

Scenario

Figure 1- 1



As shown in Figure 1- 1 , Switch A, Switch B, and Switch C form a simple L2 data switching network.

Deployment

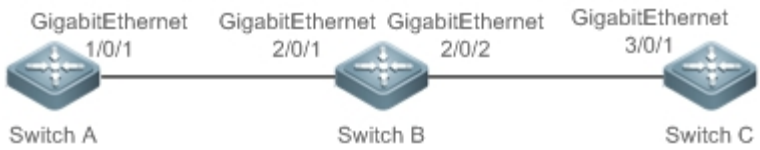
- Connect Switch A to Switch B through physical ports GigabitEthernet 1/0/1 and GigabitEthernet 2/0/1.
- Connect Switch B to Switch C through physical ports GigabitEthernet 2/0/2 and GigabitEthernet 3/0/1.
- Configure GigabitEthernet 1/0/1, GigabitEthernet 2/0/1, GigabitEthernet 2/0/2, and GigabitEthernet3/0/1 as Trunk ports.

- Create a switch virtual interface (SVI), SVI 1, on Switch A and Switch C respectively, and configure IP addresses from a network segment for the two SVIs. The IP address of SVI 1 on Switch A is 192.168.1.1/24, and the IP address of SVI 1 on Switch C is 192.168.1.2/24.
- Run the **ping 192.168.1.2** command on Switch A and the **ping 192.168.1.1** command on Switch C to implement data switching through Switch B.

1.2.2 L3 Routing Through the Physical Ethernet Interface

Scenario

Figure 1- 2



As shown in Figure 1- 2, Switch A, Switch B, and Switch C form a simple L3 data communication network.

Deployment

- Connect Switch A to Switch B through physical ports GigabitEthernet 1/0/1 and GigabitEthernet 2/0/1.
- Connect Switch B to Switch C through physical ports GigabitEthernet 2/0/2 and GigabitEthernet 3/0/1.
- Configure GigabitEthernet 1/0/1, GigabitEthernet 2/0/1, GigabitEthernet 2/0/2, and GigabitEthernet3/0/1 as L3 routed ports.
- Configure IP addresses from a network segment for GigabitEthernet 1/0/1 and GigabitEthernet 2/0/1. The IP address of GigabitEthernet 1/0/1 is 192.168.1.1/24, and the IP address of GigabitEthernet 2/0/1 is 192.168.1.2/24.
- Configure IP addresses from a network segment for GigabitEthernet 2/0/2 and GigabitEthernet 3/0/1. The IP address of GigabitEthernet 2/0/2 is 192.168.2.1/24, and the IP address of GigabitEthernet 3/0/1 is 192.168.2.2/24.
- Configure a static route entry on Switch C so that Switch C can directly access the network segment 192.168.1.0/24.
- Run the **ping 192.168.2.2** command on Switch A and the **ping 192.168.1.1** command on Switch C to implement L3 routing through Switch B.

1.3 Features

Basic Concepts

Interface Classification

Interfaces on FS devices fall into three categories:

- L2 interface (Switch or Gateway Bridge mode)
- L3 interface (supported by L3 devices)

4. Common L2 interfaces are classified into the following types:

- Switch port
- L2 aggregate port (AP)

5. Common L3 interfaces are classified into the following types:

- Routed port
- L3 AP port
- SVI
- Loopback interface
- Tunnel interface

Switch Port

A switch port is an individual physical port on the device, and implements only the L2 switching function. The switch port is used to manage physical ports and L2 protocols related to physical ports.

L2 AP Port

An AP port is formed by aggregating multiple physical ports. Multiple physical links can be bound together to form a simple logical link. This logical link is called an AP port.

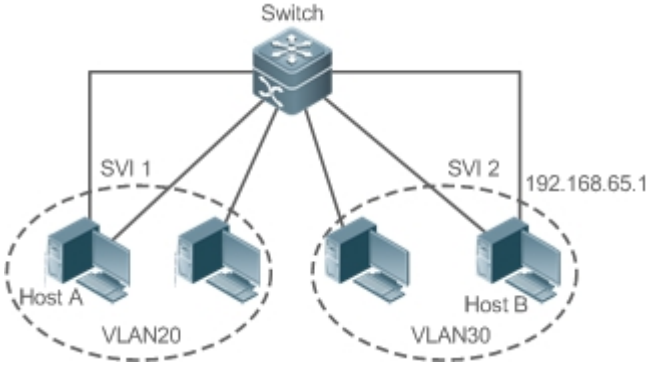
For L2 switching, an AP port is equivalent to a switch port that combines bandwidths of multiple ports, thus expanding the link bandwidth. Frames sent over the L2 AP port are balanced among the L2 AP member ports. If one member link fails, the L2 AP port automatically transfers the traffic on the faulty link to other member links, improving reliability of connections.

SVI

The SVI can be used as the management interface of the local device, through which the administrator can manage the device. You can also create an SVI as a gateway interface, which is mapped to the virtual interface of each VLAN to implement routing across VLANs among L3 devices. You can run the **interface vlan** command to create an SVI and assign an IP address to this interface to set up a route between VLANs.

As shown in Figure 1- 3, hosts in VLAN 20 can directly communicate with each other without participation of L3 devices. If Host A in VLAN 20 wants to communicate with Host B in VLAN 30, SVI 1 of VLAN 20 and SVI 2 of VLAN 30 must be used.

Figure 1- 3



Routed Port

A physical port on a L3 device can be configured as a routed port, which functions as the gateway interface for L3 switching. A routed port is not related with a specific VLAN. Instead, it is just an access port. The routed port cannot be used for L2 switching. You can run the **no switchport** command to change a switch port to a routed port and assign an IP address to this port to set up a route. Note that you must delete all L2 features of a switch port before running the **no switchport** command.

If a port is a L2 AP member port or a DOT1X port that is not authenticated, you cannot run the **switchport** or **no switchport**

command to configure the switch port or routed port.

L3 AP Port

Like the L2 AP port, a L3 AP port is a logical port that aggregates multiple physical member ports. The aggregated ports must be the L3 ports of the same type. The AP port functions as a gateway interface for L3 switching. Multiple physical links are combined into one logical link, expanding the bandwidth of a link. Frames sent over the L3 AP port are balanced among the L3 AP member ports. If one member link fails, the L3 AP port automatically transfers the traffic on the faulty link to other member links, improving reliability of connections.

A L3 AP port cannot be used for L2 switching. You can run the **no switchport** command to change a L2 AP port that does not contain any member port into a L3 AP port, add multiple routed ports to this L3 AP port, and then assign an IP address to this L3 AP port to set up a route.

Loopback Interface

The loopback interface is a local L3 logical interface simulated by the software that is always UP. Packets sent to the loopback interface are processed on the device locally, including the route information. The IP address of the loopback interface can be used as the device ID of the Open Shortest Path First (OSPF) routing protocol, or as the source address used by Border Gateway Protocol (BGP) to set up a TCP connection. The procedure for configuring a loopback interface is similar to that for configuring an Ethernet interface, and you can treat the loopback interface as a virtual Ethernet interface.

Tunnel interface

The Tunnel interface is used to implement the tunneling function, allowing the use of transport protocols (such as IP) to transmit network packets of any protocol. Like other logical interfaces, the Tunnel interface is a virtual interface to the system. the Tunnel interface does not specifically specify a transport protocol or load protocol, it provides a transport mode used to implement a standard point-to-point. Because Tunnel implements a point-to-point transport link, a Tunnel interface must be set up for each individual link.Translated with www.DeepL.com/Translator (free version)

Overview

Feature	Description
Interface Configuration Commands	You can configure interface-related attributes in interface configuration mode. If you enter interface configuration mode of a non-existing logical interface, the interface will be created.
Interface Description and Administrative Status	You can configure a name for an interface to identify the interface and help you remember the functions of the interface. You can also configure the administrative status of the interface.
MTU	You can configure the maximum transmission unit (MTU) of a port to limit the length of a frame that can be received or sent over this port.
Bandwidth	You can configure the bandwidth of an interface.
Load Interval	You can specify the interval for load calculation of an interface.
Carrier Delay	You can configure the carrier delay of an interface to adjust the delay after which the status of an interface changes from Down to Up or from Up to Down.
Link Trap Policy	You can enable or disable the link trap function on an interface.
Interface Index Persistence	You can enable the interface index persistence function so that the interface index remains unchanged after the device is restarted.

Feature	Description
Routed Port	You can configure a physical port on a L3 device as a routed port, which functions as the gateway interface for L3 switching.
L3 AP Port	You can configure an AP port on a L3 device as a L3 AP port, which functions as the gateway interface for L3 switching.
Selection of Interface Medium Type	You can select the medium type (fiber or copper) of a combo port as required.
Interface Speed, Duplex Mode, Flow Control Mode, and Auto Negotiation Mode	You can configure the speed, duplex mode, flow control mode, and auto negotiation mode of an interface.
Automatic Module Detection	If the interface speed is set to auto, the interface speed can be automatically adjusted based on the type of the inserted module.
Protected Port	You can configure some ports as protected ports to disable communication between these ports. You can also disable routing between protected ports.
Port Errdisable Recovery	After a port is shut down due to a violation, you can run the **errdisable recovery** command in global configuration mode to recover all the ports in errdisable state and enable these ports.
Optical Module Alarm Detection	You can configure the optical module alarm detection function to detect alarms generated upon optical module insertion or removal, optical module exceptions, optical module exception clearing, and other events.
EEE	You can configure the Energy Efficient Ethernet (EEE) function to enable the interface to work in low power consumption mode.
Port Flapping Protection	You can configure the port flapping protection function so that the system can automatically shut down a port when flapping occurs on the port.

1.3.1 Interface Configuration Commands

üRun the interface command in global configuration mode to enter interface configuration mode. You can configure interface-related attributes in interface configuration mode.

Working Principle

Run the interface command in global configuration mode to enter interface configuration mode. If you enter interface configuration mode of a non-existing logical interface, the interface will be created. You can also run the interface range or interface range macro command in global configuration mode to configure the range (IDs) of interfaces. Interfaces defined in the same range must be of the same type and have the same features.

You can run the **no interface** command in global configuration mode to delete a specified logical interface.

Interface Numbering Rules

In stand-alone mode, the ID of a physical port consists of two parts: slot ID and port ID on the slot. For example, if the slot ID of the port is 2, and port ID on the slot is 3, the interface ID is 2/3. In Stacking or stack mode, the ID of a physical port consists of three parts: device ID, slot ID, and port ID on the slot. For example, if the device ID is 1, slot ID of the port is 2, and port ID on the slot is 3, the interface ID is 1/2/3.

The device ID ranges from 1 to the maximum number of supported member devices.

The slot number rules are as follows: The static slot ID is 0, whereas the ID of a dynamic slot (pluggable module or line card) ranges from 1 to the number of slots. Assume that you are facing the device panel. Dynamic slot are numbered from 1 sequentially from front to rear, from left to right, and from top to bottom.

The ID of a port on the slot ranges from 1 to the number of ports on the slot, and is numbered sequentially from left to right.

You can select fiber or copper as the medium of a combo port. Regardless of the medium selected, the combo port uses the same port ID.

The ID of an AP port ranges from 1 to the number of AP ports supported by the device.

The ID of an SVI is the VID of the VLAN corresponding to this SVI.

Configuring Interfaces Within a Range

You can run the **interface range** command in global configuration mode to configure multiple interfaces at a time. Attributes configured in interface configuration mode apply to all these interfaces.

The **interface range** command can be used to specify several interface ranges.

The **macro** parameter is used to configure the macro corresponding to a range. For details, see "Configuring Macros of Interface Ranges."

Ranges can be separated by commas (,).

The types of interfaces within all ranges specified in a command must be the same.

Pay attention to the format of the **range** parameter when you run the **interface range** command.

The following interface range formats are valid:

- **FastEthernet** device/slot/{first port} - {last port};
- **GigabitEthernet** device/slot/{first port} - {last port};
- **TenGigabitEthernet** device/slot/{first port} - {last port};
- **FortyGigabitEthernet** device/slot/{first port} - {last port};
- **AggregatePort** *Aggregate-port ID* (The AP ID ranges from 1 to the maximum number of AP ports supported by the device.)
- **vlan** vlan-ID-vlan-ID (The VLAN ID ranges from 1 to 4,094.)
- **Loopback** loopback-ID (The loopback ID ranges from 1 to 2,147,483,647.)
- Tunnel tunnel-ID-tunnel-ID, the range is 0 to the maximum number of Tunnel ports supported by the device minus one.

Interfaces in an interface range must be of the same type, namely, FastEthernet, GigabitEthernet, AggregatePort, or SVI.

Configuring Macros of Interface Ranges

You can define some macros to replace the interface ranges. Before using the **macro** parameter in the **interface range** command, you must first run the **define interface-range** command in global configuration mode to define these macros.

Run the **no define interface-range macro_name** command in global configuration mode to delete the configured macros.

1.3.2 Interface Description and Administrative Status

You can configure a name for an interface to identify the interface and help you remember the functions of the interface.

You can enter interface configuration mode to enable or disable an interface.

Working Principle

Interface Description

You can configure the name of an interface based on the purpose of the interface. For example, if you want to assign GigabitEthernet 1/1 for exclusive use by user A, you can describe the interface as "Port for User A."

Interface Administrative Status

You can configure the administrative status of an interface to disable the interface as required. If the interface is disabled, no frame will be received or sent on this interface, and the interface will loss all its functions. You can enable a disabled interface by configuring the administrative status of the interface. Two types of interface administrative status are defined: Up and Down. The administrative status of an interface is Down when the interface is disabled, and Up when the interface is enabled.

1.3.3 MTU

You can configure the MTU of a port to limit the length of a frame that can be received or sent over this port.

Working Principle

When a large amount of data is exchanged over a port, frames greater than the standard Ethernet frame may exist. This type of frame is called jumbo frame. The MTU is the length of the valid data segment in a frame. It does not include the Ethernet encapsulation overhead.

If a port receives or sends a frame with a length greater than the MTU, this frame will be discarded.

1.3.4 Bandwidth

Working Principle

The **bandwidth** command can be configured so that some routing protocols (for example, OSPF) can calculate the route metric and the Resource Reservation Protocol (RSVP) can calculate the reserved bandwidth. Modifying the interface bandwidth will not affect the data transmission rate of the physical port.

The **bandwidth** command is a routing parameter, and does not affect the bandwidth of a physical link.

1.3.5 Load Interval

Working Principle

You can run the **load-interval** command to specify the interval for load calculation of an interface. Generally, the interval is 10s.

1.3.6 Carrier Delay

Working Principle

The carrier delay refers to the delay after which the data carrier detect (DCD) signal changes from Down to Up or from Up to Down. If the DCD status changes during the delay, the system will ignore this change to avoid negotiation at the upper data link layer. If this parameter is set to a great value, nearly every DCD change is not detected. On the contrary, if the parameter is set to 0, every DCD signal change will be detected, resulting in poor stability.

If the DCD carrier is interrupted for a long time, the carrier delay should be set to a smaller value to accelerate convergence of the topology or route. On the contrary, if the DCD carrier interruption time is shorter than the topology or route convergence time, the

carrier delay should be set to a greater value to avoid topology or route flapping.

1.3.7 Link Trap Policy

You can enable or disable the link trap function on an interface.

Working Principle

When the link trap function on an interface is enabled, the Simple Network Management Protocol (SNMP) sends link traps when the link status changes on the interface.

1.3.8 Interface Index Persistence

Like the interface name, the interface index also identifies an interface. When an interface is created, the system automatically assigns a unique index to the interface. The index of an interface may change after the device is restarted. You can enable the interface index persistence function so that the interface index remains unchanged after the device is restarted.

Working Principle

After interface index persistence is enabled, the interface index remains unchanged after the device is restarted.

1.3.9 Routed Port

Working Principle

A physical port on a L3 device can be configured as a routed port, which functions as the gateway interface for L3 switching. The routed port cannot be used for L2 switching. You can run the **no switchport** command to change a switch port to a routed port and assign an IP address to this port to set up a route. Note that you must delete all L2 features of a switch port before running the **no switchport** command.

1.3.10 L3 AP Port

Working Principle

Like a L3 routed port, you can run the **no switchport** command to change a L2 AP port into a L3 AP port on a L3 device, and then assign an IP address to this AP port to set up a route. Note that you must delete all L2 features of the AP port before running the **no switchport** command.

A L2 AP port with one or more member ports cannot be configured as a L3 AP port. Similarly, a L3 AP port with one or more member ports cannot be changed to a L2 AP port.

1.3.11 Selection of Interface Medium Type

You can select the medium type (fiber or copper) of a combo port as required.

Working Principle

You can choose either fiber or copper as the medium, but the two media cannot take effect at the same time. Once you select the medium, attributes, including the connection status, speed, duplex mode, and flow control mode, are attributes of the selected medium. If you change the medium, the interface will adopt the default settings, and you must re-configure these attributes according to requirements.

The Combo Port Supports Automatic Selection of the Medium Type

- If you enable automatic selection of the medium type, the device uses the current medium if only one medium is available.
- If both media are available, the device uses the preferred medium that is configured. By default, the preferred medium is copper. You can run the **medium-type auto-select prefer fiber** command to configure fiber as the preferred media. In automatic medium selection mode, the interface adopts the default settings of attributes, such as the speed, duplex mode, and flow control mode.

If an interface is enabled with automatic selection, its peer interface must be enabled with auto negotiation; otherwise, an error will occur.

The command takes effect only on a physical port. An AP port or SVI does not support configuration of the medium type.

The command takes effect only on a port that supports medium selection.

All ports that are configured as member ports of an AP port must have the same medium type; otherwise, they cannot be added to the AP port. The type of member ports cannot be modified. A port enabled with automatic medium selection cannot be added to an AP port.

1.3.12 Interface Speed, Duplex Mode, Flow Control Mode, and Auto Negotiation Mode

You can configure the interface speed, duplex mode, flow control mode, and auto negotiation mode of an Ethernet physical port or AP port.

Working Principle

Speed

Generally, the speed of an Ethernet physical port is determined through negotiation with the peer device. The negotiated speed can be any speed within the interface capability. You can also configure any speed within the interface capability for the Ethernet physical port.

When you configure the speed of an AP port, the configuration takes effect on all of its member ports. (All these member ports are Ethernet physical ports.)

Duplex Mode

- The duplex mode of an Ethernet physical port or AP port can be configured as follows:
- Set the duplex mode of the interface to full-duplex so that the interface can receive packets while sending packets.
- Set the duplex mode of the interface to half-duplex so that the interface can receive or send packets at a time.
- Set the duplex mode of the interface to auto-negotiation so that the duplex mode of the interface is determined through auto negotiation between the local interface and peer interface.
- When you configure the duplex mode of an AP port, the configuration takes effect on all of its member ports. (All these member ports are Ethernet physical ports.)

Flow Control

Two flow control modes are defined for an interface:

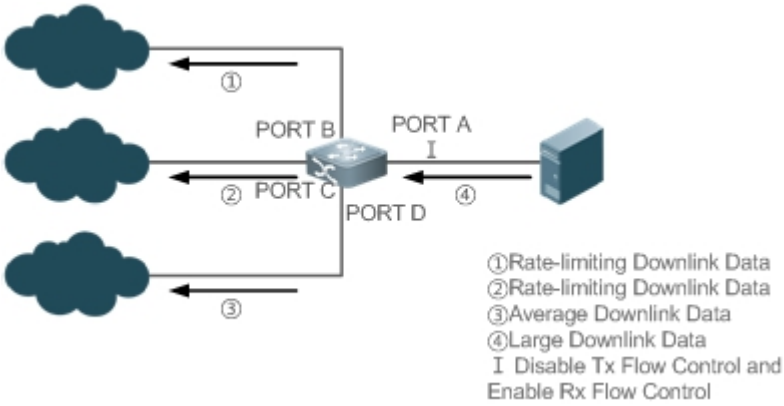
- Symmetric flow control mode: Generally, after flow control is enabled on an interface, the interface processes the received flow control frames, and sends the flow control frames when congestion occurs on the interface. The received and sent flow control frames are processed in the same way. This is called symmetric flow control mode.
- Asymmetric flow control mode: In some cases, an interface on a device is expected to process the received flow control frames to

ensure that no packet is discarded due to congestion, and not to send the flow control frames to avoid decreasing the network speed. In this case, you need to configure asymmetric flow control mode to separate the procedure for receiving flow control frames from the procedure for sending flow control frames.

- When you configure the flow control mode of an AP port, the configuration takes effect on all of its member ports. (All these member ports are Ethernet physical ports.)

As shown in Figure 1-4, Port A of the device is an uplink port, and Ports B, C and D are downlink ports. Assume that Port A is enabled with the functions of sending and receiving flow control frames. Port B and Port C are connected to different slow networks. If a large amount of data is sent on Port B and Port C, Port B and Port C will be congested, and consequently congestion occurs in the inbound direction of Port A. Therefore, Port A sends flow control frames. When the uplink device responds to the flow control frames, it reduces the data flow sent to Port A, which indirectly slows down the network speed on Port D. At this time, you can disable the function of sending flow control frames on Port A to ensure the bandwidth usage of the entire network.

Figure 1-4



Auto Negotiation Mode

- The auto negotiation mode of an interface can be On or Off. The auto negotiation state of an interface is not completely equivalent to the auto negotiation mode. The auto negotiation state of an interface is jointly determined by the interface speed, duplex mode, flow control mode, and auto negotiation mode.
- When you configure the auto negotiation mode of an AP port, the configuration takes effect on all of its member ports. (All these member ports are Ethernet physical ports.)

Generally, if one of the interface speed, duplex mode, and flow control mode is set to auto, or the auto negotiation mode of an interface is On, the auto negotiation state of the interface is On, that is, the auto negotiation function of the interface is enabled. If none of the interface speed, duplex mode, and flow control mode is set to auto, and the auto negotiation mode of an interface is Off, the auto negotiation state of the interface is Off, that is, the auto negotiation function of the interface is disabled.

For a 100M fiber port, the auto negotiation function is always disabled, that is, the auto negotiation state of a 100M fiber port is always Off. For a Gigabit copper port, the auto negotiation function is always enabled, that is, the auto negotiation state of a Gigabit copper port is always On.

1.3.13 Automatic Module Detection

If the interface speed is set to auto, the interface speed can be automatically adjusted based on the type of the inserted module.

Working Principle

Currently, the automatic module detection function can be used to detect only the SFP and SFP+ modules. The SFP is a Gigabit module, whereas SFP+ is a 10 Gigabit module. If the inserted module is SFP, the interface works in Gigabit mode. If the inserted module is SFP+, the interface works in 10 Gigabit mode.

The automatic module detection function takes effect only when the interface speed is set to auto.

1.3.14 Protected Port

In some application environments, it is required that communication be disabled between some ports. For this purpose, you can configure some ports as protected ports. You can also disable routing between protected ports.

Working Principle

Protected Port

After ports are configured as protected ports, protected ports cannot communicate with each other, but can communicate with non-protected ports.

Protected ports work in either of the two modes. In the first mode, L2 switching is blocked but routing is allowed between protected ports. In the second mode, both L2 switching and routing are blocked between protected ports. If a protected port supports both modes, the first mode is used by default.

When two protected port are configured as a pair of mirroring ports, frames sent or received by the source port can be mirrored to the destination port.

Currently, only an Ethernet physical port or AP port can be configured as a protected port. When an AP port is configured as a protected port, all of its member ports are configured as protected ports.

Blocking L3 Routing Between Protected Ports

By default, L3 routing between protected ports is not blocked. In this case, you can run the **protected-ports route-deny** command to block routing between protected ports.

1.3.15 Port Errdisable Recovery

Some protocols support the port errdisable recovery function to ensure security and stability of the network. For example, in the port security protocol, when you enable port security and configure the maximum number of security addresses on the port, a port violation event is generated if the number of addresses learned on this port exceeds the maximum number of security addresses. Other protocols, such as the Spanning Tree Protocol (STP), DOT1X, and REUP, support the similar functions, and a violating port will be automatically shut down to ensure security.

Working Principle

After a port is shut down due to a violation, you can run the **errdisable recovery** command in global configuration mode to recovery all the ports in errdisable state and enable these ports. You can manually recover a port, or automatically recover a port at a scheduled time.

1.3.16 EEE

Energy Efficient Ethernet (EEE) is an energy efficient Ethernet solution. When EEE is enabled, the port enters low power consumption mode when the Ethernet connection is idle, thus saving the energy.

Low Power Idle (LPI) is the low power consumption mode. After a port enters LPI mode, it reduces signals significantly, and only sends

signals that are sufficient to maintain the connection on the port to save the energy.

Working Principle

According to the Ethernet standards or specifications, interfaces with a bandwidth of 100M or above have the idle state. An interface will consume much power if it maintains connection without being affected by data transmission. Therefore, the power consumption is high no matter whether any data is transmitted on the link. Even if no data is transmitted, the port will always send the idle signals to retain the connection state of the link.

EEE enables a port to enter LPI mode for the purpose of saving energy. In LPI mode, the power consumption is low when the link is idle. The EEE technology can also quickly change the LPI state of a port to the normal state, providing high-performance data transmission.

After enabled with EEE, the port automatically enters LPI mode if the port is always Up without sending or receiving any packet in a period of time. The port recovers the working mode when it needs to send or receive packets, thus saving the energy. To make the EEE function take effect, the peer port must also support the EEE function.

Only a copper port working in 100M or 1000M speed mode supports the EEE function.

The EEE function takes effect only on the port enabled with auto negotiation.

1.3.17 Port Flapping Protection

When flapping occurs on a port, a lot of hardware interruptions occur, consuming a lot of CPU resources. On the other hand, frequent port flapping damages the port. You can configure the flapping protection function to protect ports.

Working Principle

By default, the port flapping protection function is enabled. You can disable this function as required.

There are two kinds of port oscillation protection mechanism:

- When flapping occurs on a port, the port detects flapping every 2s or 10s. If flapping occurs six times within 2s on a port, the device displays a prompt. If 10 prompts are displayed continuously, that is, port flapping is detected continuously within 20s, the port is disabled. If flapping occurs 10 times within 10s on a port, the device displays a prompt without disabling the port.
- Flapping detection is enabled every 30s. A section of flapping includes at least 60 flappings within 30s. Then, a syslog will be printed after three consecutive sections of flapping occur. If flapping protection is enabled, the port will be shut down. If not, the port will not be shut down.

1.3.18 Syslog

You can enable or disable the syslog function to determine whether to display information about the interface changes or exceptions.

Working Principle

You can enable or disable the syslog function as required. By default, this function is enabled. When an interface becomes abnormal, for example, the interface status changes, or the interface receives error frames, or flapping occurs, the system displays prompts to notify users.

1.3.19 Global MTU

Users can set the global MTU to control the maximum length of frames that can be sent and received over all ports.

Working Principle

When large-throughput data exchange is performed over a port, frames whose length is longer than that of a standard Ethernet frame may exist, and these frames are called jumbo frames. The MTU indicates the length of valid data fields in a frame, excluding the Ethernet encapsulation overhead.

If the length of a frame received or forwarded by a port exceeds the MTU value, the frame will be discarded.

The MTU value ranges from 64 to 9216 bytes. The granularity is four bytes. The default value is 1500 bytes.

- The IP MTU automatically changes to the value of the link MTU of an interface when the globally set link MTU changes.
- The MTU of an interface takes precedence over the global MTU. After the global MTU is configured, the MTU of an interface cannot be set to the default value.

1.4 Configuration

Configuration	Description and Command	
Performing Basic Configurations	(Optional) It is used to manage interface configurations, for example, creating/deleting an interface, or configuring the interface description.	
	interface	Creates an interface and enters configuration mode of the created interface or a specified interface.
	interface range	Enters an interface range, creates these interfaces (if not created), and enters interface configuration mode.
	define interface-range	Creates a macro to specify an interface range.
	snmp-server if-index persist	Enables the interface index persistence function so that the interface index remains unchanged after the device is restarted.
	description	Configures the interface description of up to 80 characters in interface configuration mode.
	snmp trap link-status	Configures whether to send the link traps of the interface.
	shutdown	Shuts down an interface in interface configuration mode.
	physical-port dither protect	Configures the port flapping protection function in global configuration mode.
	logging [link-updown | error-frame | link-dither]	Configures the syslog function on an interface in global configuration mode.
	port-dither protect	Enables port oscillation protection globally.
Configuring Interface Attributes	(Optional) It is used to configure interface attributes.	
	bandwidth	Configures the bandwidth of an interface in interface configuration mode.
	carrier-delay	Configures the carrier delay of an interface in interface configuration mode.
	load-interval	Configures the interval for load calculation of an interface.
	duplex	Configures the duplex mode of an interface.
	flowcontrol	Enables or disables flow control of an interface.
	medium-type	Selects the medium type of an interface.
	mtu	Configures the MTU of an interface.

Configuration	Description and Command	
	negotiation mode	Configures the auto negotiation mode of an interface.
	speed	Configures the speed of an interface.
	switchport	Configures an interface as a L2 interface in interface configuration mode. (Run the **no switchport** command to configure an interface as a L3 interface.)
	switchport protected	Configures a port as a protected port.
	protected-ports route-deny	Blocks L3 routing between protected ports in global configuration mode.
	errdisable recovery	Recovers a port in errdisable state in global configuration mode.
	fiber antifake ignore	Disables the optical module antifake detection function in global configuration mode.
	fiber antifake enable	Enables the optical module antifake detection function in global configuration mode.
	eee enable	Enables EEE in interface configuration mode.

1.4.1 Performing Basic Configurations

Configuration Effect

- Create a specified logical interface and enter configuration mode of this interface, or enter configuration mode of an existing physical or logical interface.
- Create multiple specified logical interfaces and enter interface configuration mode, or enter configuration mode of multiple existing physical or logical interfaces.
- The interface indexes remain unchanged after the device is restarted.
- Configure the interface description so that users can directly learn information about the interface.
- Enable or disable the link trap function of an interface.
- Enable or disable an interface.
- Enable or disable port flapping protection.

Notes

- The **no** form of the command can be used to delete a specified logical interface or logical interfaces in a specified range, but cannot be used to delete a physical port or physical ports in a specified range.
- The **default** form of the command can be used in interface configuration mode to restore default settings of a specified physical or logical interface, or interfaces in a specified range.

Configuration Steps

Configuring a Specified Interface

- Optional.
- Run this command to create a logical interface or enter configuration mode of a physical port or an existing logical interface.

Command	**interface** *interface-type interface-number*

Parameter Description	*interface-type interface-number*: Indicates the type and number of the interface. The interface can be an Ethernet physical port, AP port, SVI, or loopback interface.
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	● If a logical interface is not created yet, run this command to create this interface and enter configuration mode of this interface. ● For a physical port or an existing logical interface, run this command to enter configuration mode of this interface. ● Use the **no** form of the command to delete a specified logical interface. ● Use the **default** form of the command to restore default settings of the interface in interface configuration mode.

Configuring Interfaces Within a Range

- Optional.
- Run this command to create multiple logical interfaces or enter configuration mode of multiple physical port or existing logical interfaces.

Command	**interface range** { *port-range* | **macro** *macro_name* }
Parameter Description	*port-range*: Indicates the type and ID range of interfaces. These interfaces can be Ethernet physical ports, AP ports, SVIs, or loopback interfaces. *macro_name*: Indicates the name of the interface range macro.
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	● If logical interfaces are not created yet, run this command to create these interfaces and enter interface configuration mode. ● For multiple physical ports or existing logical interfaces, run this command to enter interface configuration mode. ● Use the **default** form of the command to restore default settings of these interfaces in interface configuration mode. ● Before using a macro, run the **define interface-range** command to define the interface range as a macro name in global configuration mode, and then run the **interface range macro** *macro_name* command to apply the macro.

Configuring Interface Index Persistence

- Optional.
- Run this command when the interface indexes must remain unchanged after the device is restarted.

Command	**snmp-server if-index persist**
Parameter Description	N/A
Defaults	By default, interface index persistence is disabled.
Command	Global configuration mode

Mode	
Usage Guide	After this command is executed, current indexes of all interfaces will be saved, and the indexes remain unchanged after the device is restarted. You can use the **no** or **default** form of the command to disable the interface index persistence function.

Configuring the Description of an Interface

- Optional.
- Run this command to configure the description of an interface.

Command	**description** *string*
Parameter Description	*string:* Indicates a string of up to 80 characters.
Defaults	By default, no description is configured.
Command Mode	Interface configuration mode
Usage Guide	This command is used to configure the description of an interface. You can use the **no** or **default** form of the command to delete the description of an interface.-

Configuring the Link Trap Function of an Interface

- Optional.
- Run this command to obtain the link traps through SNMP.

Command	**snmp trap link-status**
Parameter Description	N/A
Defaults	By default, the link trap function is enabled.
Command Mode	Interface configuration mode
Usage Guide	This command is used to configure the link trap function on an interface. When this function is enabled, the SNMP sends link traps when the link status changes on the interface. You can use the **no** or **default** form of the command to disable the link trap function.

Configuring the Administrative Status of an Interface

- Optional.
- Run this command to enable or disable an interface.
- An interface cannot send or receive packets after it is disabled.

Command	**shutdown**
Parameter Description	N/A
Defaults	By default, the administrative status of an interface is Up.
Command Mode	Interface configuration mode
Usage Guide	You can run the **shutdown** command to disable an interface, or the **no shutdown** command to enable an interface.

	In some cases, for example, when an interface is in errdisable state, you cannot run the **no shutdown** command on an interface. You can use the **no** or **default** form of the command to enable the interface.

Configuring Port Flapping Protection

- Optional.
- Run this command to protect the port against flapping.

Command	**physical-port dither protect**
Parameter Description	N/A
Defaults	By default, port flapping protection is enabled.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring Port Flapping Protection

- Optional.
- Run this command to protect the port against flapping.

Command	**port dither protect**
Parameter Description	N/A
Defaults	By default, port flapping protection is enabled.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Syslog Function

- Optional.
- Run this command to enable or disable the syslog function on an interface.

Command	**[no] logging [link-updown | error-frame | link-dither]**
Parameter Description	N/A
Defaults	By default, the syslog function is enabled on an interface.
Command Mode	Global configuration mode
Usage Guide	N/A

Verification

Configuring a Specified Interface

- Run the **interface** command. If you can enter interface configuration mode, the configuration is successful.
- For a logical interface, after the **no interface** command is executed, run the **show running** or **show interfaces** command to check whether the logical interface exists. If not, the logical interface is deleted.

- After the **default interface** command is executed, run the **show running** command to check whether the default settings of the corresponding interface are restored. If yes, the operation is successful.

Configuring Interfaces Within a Range

- Run the **interface range** command. If you can enter interface configuration mode, the configuration is successful.
- After the **default interface range** command is executed, run the **show running** command to check whether the default settings of the corresponding interfaces are restored. If yes, the operation is successful.

Configuring Interface Index Persistence

- After the **snmp-server if-index persist** command is executed, run the **write** command to save the configuration, restart the device, and run the **show interface** command to check the interface index. If the index of an interface remains the same after the restart, interface index persistence is enabled.

Configuring the Link Trap Function of an Interface

- Remove and then insert the network cable on a physical port, and enable the SNMP server. If the SNMP server receives link traps, the link trap function is enabled.
- Run the **no** form of the **snmp trap link-status** command. Remove and then insert the network cable on a physical port. If the SNMP server does not receive link traps, the link trap function is disabled.

Configuring the Administrative Status of an Interface

- Insert the network cable on a physical port, enable the port, and run the **shutdown** command on this port. If the syslog is displayed on the Console indicating that the state of the port changes to Down, and the indicator on the port is off, the port is disabled. Run the **show interfaces** command, and verify that the interface state changes to Administratively Down. Then, run the **no shutdown** command to enable the port. If the syslog is displayed on the Console indicating that the state of the port changes to Up, and the indicator on the port is on, the port is enabled.

Configuring Port Flapping Protection

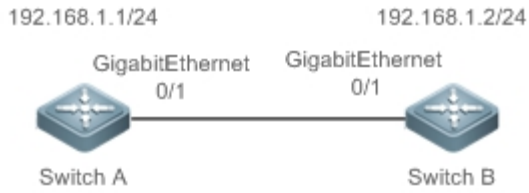
- Run the **physical-port dither protect** command in global configuration mode. Frequently remove and insert the network cable on a physical port to simulate port flapping. Verify that a syslog indicating port flapping is displayed on the Console. After such a syslog is displayed for several times, the system prompts that the port will be shut down.

Configuring the Syslog Function

- Run the **logging link-updown** command in global configuration mode to display the interface status information. Remove and then insert the network cable on a physical port. The interface state will change twice. Verify that the information is displayed on the Console, indicating that the interface state changes from Up to Down, and then from Down to Up. Run the **no logging link-updown** command. Remove and then insert the network cable. Verify that the related information is no longer displayed on the Console. This indicates that the syslog function is normal.

Configuration Example

Configuring Basic Attributes of Interfaces

Scenario **Figure 1- 5**	
Configuration Steps	● Connect two devices through the switch ports. ● Configure an SVI respectively on two devices, and assign IP addresses from a network segment to the two SVIs. ● Enable interface index persistence on the two devices. ● Enable the link trap function on the two devices. ● Configure the interface administrative status on the two devices.
A	A# configure terminal A(config)# snmp-server if-index persist A(config)# interface vlan 1 A(config-if-VLAN 1)# ip address 192.168.1.1 255.255.255.0 A(config-if-VLAN 1)# exit A(config)# interface gigabitethernet 0/1 A(config-if-GigabitEthernet 0/1)# snmp trap link-status A(config-if-GigabitEthernet 0/1)# shutdown A(config-if-GigabitEthernet 0/1)# end A# write
B	B# configure terminal B(config)# snmp-server if-index persist B(config)# interface vlan 1 B(config-if-VLAN 1)# ip address 192.168.1.2 255.255.255.0 B(config-if-VLAN 1)# exit B(config)# interface gigabitethernet 0/1 B(config-if-GigabitEthernet 0/1)# snmp trap link-status B(config-if-GigabitEthernet 0/1)# shutdown B(config-if-GigabitEthernet 0/1)# end B# write
Verification	Perform verification on Switch A and Switch B as follows: ● Run the shutdown command on port GigabitEthern 0/1, and check whether GigabitEthern 0/1 and SVI 1 are Down. ● Run the shutdown command on port GigabitEthern 0/1, and check whether a trap indicating that this interface

	is Down is sent. ● Restart the device, and check whether the index of GigabitEthern 0/1 is the same as that before the restart.
A	A# show interfaces gigabitEthernet 0/1 Index(dec):1 (hex):1 **GigabitEthernet 0/1 is administratively down**, line protocol is DOWN Hardware is GigabitEthernet, address is 00d0.f865.de9b (bia 00d0.f865.de9b) Interface address is: no ip address MTU 1500 bytes, BW 1000000 Kbit Encapsulation protocol is Bridge, loopback not set Keepalive interval is 10 sec, set Carrier delay is 2 sec Rxload is 1/255, Txload is 1/255 Queue Transmitted packets Transmitted bytes Dropped packets Dropped bytes 0 0 0 0 0 1 0 0 0 0 2 0 0 0 0 3 0 0 0 0 4 0 0 0 0 5 0 0 0 0 6 0 0 0 0 7 4 440 0 0 Switchport attributes: interface's description:"" lastchange time:0 Day:20 Hour:15 Minute:22 Second Priority is 0 admin medium-type is Copper, oper medium-type is Copper admin duplex mode is AUTO, oper duplex is Unknown admin speed is AUTO, oper speed is Unknown flow control admin status is OFF, flow control oper status is Unknown admin negotiation mode is OFF, oper negotiation state is ON Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF Port-type: access Vlan id: 1 10 seconds input rate 0 bits/sec, 0 packets/sec

	10 seconds output rate 0 bits/sec, 0 packets/sec 4 packets input, 408 bytes, 0 no buffer, 0 dropped Received 0 broadcasts, 0 runts, 0 giants 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort 4 packets output, 408 bytes, 0 underruns, 0 dropped 0 output errors, 0 collisions, 0 interface resets A# show interfaces vlan 1 Index(dec):4097 (hex):1001 VLAN 1 is UP, line protocol is DOWN Hardware is VLAN, address is 00d0.f822.33af (bia 00d0.f822.33af) **Interface address is: 192.168.1.1/24** ARP type: ARPA, ARP Timeout: 3600 seconds MTU 1500 bytes, BW 1000000 Kbit Encapsulation protocol is Ethernet-II, loopback not set Keepalive interval is 10 sec, set Carrier delay is 2 sec Rxload is 0/255, Txload is 0/255
B	B# show interfaces gigabitEthernet 0/1 Index(dec):1 (hex):1 **GigabitEthernet 0/1 is administratively down**, line protocol is DOWN Hardware is GigabitEthernet Interface address is: no ip address, address is 00d0.f865.de9b (bia 00d0.f865.de9b) MTU 1500 bytes, BW 1000000 Kbit Encapsulation protocol is Bridge, loopback not set Keepalive interval is 10 sec, set Carrier delay is 2 sec Rxload is 1/255, Txload is 1/255 Queue Transmitted packets Transmitted bytes Dropped packets Dropped bytes 0 0 0 0 0 1 0 0 0 0 2 0 0 0 0 3 0 0 0 0 4 0 0 0 0

```
        5                    0                    0                    0                    0
        6                    0                    0                    0                    0
        7                    4                  440                    0                    0
  Switchport attributes:
    interface's description:""
    lastchange time:0 Day:20 Hour:15 Minute:22 Second
    Priority is 0
    admin medium-type is Copper, oper medium-type is Copper
    admin duplex mode is AUTO, oper duplex is Unknown
    admin speed is AUTO, oper speed is Unknown
    flow control admin status is OFF, flow control oper status is Unknown
    admin negotiation mode is OFF, oper negotiation state is ON
    Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
  Port-type: access
    Vlan id: 1
   10 seconds input rate 0 bits/sec, 0 packets/sec
   10 seconds output rate 0 bits/sec, 0 packets/sec
    4 packets input, 408 bytes, 0 no buffer, 0 dropped
    Received 0 broadcasts, 0 runts, 0 giants
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
    4 packets output, 408 bytes, 0 underruns, 0 dropped
    0 output errors, 0 collisions, 0 interface resets
B# show interfaces vlan 1
Index(dec):4097 (hex):1001
VLAN 1 is UP, line protocol is DOWN
Hardware is    VLAN, address is 00d0.f822.33af (bia 00d0.f822.33af)
Interface address is: 192.168.1.2/24
ARP type: ARPA, ARP Timeout: 3600 seconds
  MTU 1500 bytes, BW 1000000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec, set
  Carrier delay is 2 sec
  Rxload is 0/255, Txload is 0/255
```

1.4.2 Configuring Interface Attributes

Configuration Effect

- Enable the device to connect and communicate with other devices through the switch port or routed port.
- Adjust various interface attributes on the device.

Configuration Steps

Configuring a Routed Port

- Optional.
- Run this command to configure a port as a L3 routed port.
- After a port is configured as a L3 routed port, L2 protocols running on the port do not take effect.
- This command is applicable to a L2 switch port.

Command	**no switchport**
Parameter Description	N/A
Defaults	By default, an Ethernet physical port is a L2 switch port.
Command Mode	Interface configuration mode
Usage Guide	On a L3 device, you can run this command to configure a L2 switch port as a L3 routed port. You can run the **switchport** command to change a L3 routed port into a L2 switch port.

Configuring a L3 AP Port

- Optional.
- Run the **no switchport** command in interface configuration mode to configure a L2 AP port as a L3 AP port. Run the **switchport** command to configure a L3 AP port as a L2 AP port.
- After a port is configured as a L3 routed port, L2 protocols running on the port do not take effect.
- This command is applicable to a L2 AP port.

Command	**no switchport**
Parameter Description	N/A
Defaults	By default, an AP port is a L2 AP port.
Command Mode	Interface configuration mode
Usage Guide	After entering configuration mode of a L2 AP port on a L3 device, you can run this command to configure a L2 AP port as a L3 AP port. After entering configuration mode of a L3 AP port, you can run the **switchport** command to change a L3 AP port into a L2 AP port.

Configuring the Medium Type of an Interface

- Optional.
- By default, the medium type of a combo port is copper.
- Port flapping may occur if the configured medium type of a port changes.
- This command is applicable to an Ethernet physical port or AP port.

Command	**medium-type** { **auto-select** [**prefer** [**fiber | copper**]] **| fiber | copper** }
Parameter Description	**auto-select**: Indicates that the medium type is selected automatically. **prefer** [**fiber** | **copper**]: Indicates the medium type that will be preferentially selected. **fiber**: Indicates that fiber is forcibly selected as the medium type. **copper**: Indicates that copper is forcibly selected as the medium type.
Defaults	By default, the medium type of an interface is copper.
Command Mode	Interface configuration mode
Usage Guide	Select either fiber or copper as the medium type of a port when both medium types are available. Once the medium type is selected, all interface attributes, including the status, duplex mode, and speed, are configured for the interface of the selected medium type. If the interface type is changed, the attributes of the new interface type are the default attributes. You can reconfigure these attributes as required. If you enable automatic selection of the medium type, the device uses the current medium if only one medium is available. If both media are available, the device uses the preferred medium as configured. By default, the preferred medium is copper. You can run the **medium-type auto-select prefer fiber** command to configure fiber as the preferred media. In automatic medium selection mode, the interface adopts the default settings of attributes, such as the speed, duplex mode, and flow control mode.

Configuring the Speed of an Interface

- Optional.
- Port flapping may occur if the configured speed of a port changes.
- This command is applicable to an Ethernet physical port or AP port.

Command	**speed** [**10** | **100** | **1000** | **auto**]
Parameter Description	**10**: Indicates that the speed of the interface is 10 Mbps. **100**: Indicates that the speed of the interface is 100 Mbps. **1000**: Indicates that the speed of the interface is 1000 Mbps. **auto**: Indicates that the speed of the interface automatically adapts to the actual condition.
Defaults	By default, the speed of an interface is auto.
Command Mode	Interface configuration mode
Usage Guide	If an interface is an AP member port, the speed of this interface is determined by the speed of the AP port. When the interface exits the AP port, it uses its own speed configuration. You can run **show interfaces** to display the speed configurations. The speed options available to an interface vary with the type of the interface. For example, you cannot set the speed of an SFP interface to 10 Mbps.

Configuring the Duplex Mode of an Interface

- Optional.
- Port flapping may occur if the configured duplex mode of a port changes.
- This command is applicable to an Ethernet physical port or AP port.

Command	**duplex** { **auto** | **full** | **half** }
Parameter Description	**auto**: Indicates automatic switching between full duplex and half duplex. **full**: Indicates full duplex. **half**: Indicates half duplex.

Defaults	By default, the duplex mode of an interface is auto.
Command Mode	Interface configuration mode
Usage Guide	The duplex mode of an interface is related to the interface type. You can run **show interfaces** to display the configurations of the duplex mode.

Configuring the Flow Control Mode of an Interface

- Optional.
- Generally, the flow control mode of an interface is off by default. For some products, the flow control mode is on by default.
- After flow control is enabled on an interface, the flow control frames will be sent or received to adjust the data volume when congestion occurs on the interface.
- Port flapping may occur if the configured flow control mode of a port changes.
- This command is applicable to an Ethernet physical port or AP port.

Command	**flowcontrol** { **auto** | **off** | **on** }
Parameter Description	**auto**: Indicates automatic flow control. **off**: Indicates that flow control is disabled. **on**: Indicates that flow control is enabled.
Defaults	By default, flow control is disabled on an interface.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Auto Negotiation Mode of an Interface

- Optional.
- Port flapping may occur if the configured auto negotiation mode of a port changes.
- This command is applicable to an Ethernet physical port or AP port.

Command	**negotiation mode** { **on** | **off** }
Parameter Description	**on**: Indicates that the auto negotiation mode is on. **off**: Indicates that the auto negotiation mode is off.
Defaults	By default, the auto negotiation mode is off.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the MTU of an Interface

- Optional.
- You can configure the MTU of a port to limit the length of a frame that can be received or sent over this port.
- This command is applicable to an Ethernet physical port or SVI.

Command	**mtu** *num*
Parameter	*num*: *64–9216*

Description	
Defaults	By default, the MTU of an interface is 1500 bytes.
Command Mode	Interface configuration mode
Usage Guide	This command is used to configure the interface MTU, that is, the maximum length of a data frame at the link layer. Currently, you can configure MTU for only a physical port or an AP port that contains one or more member ports.

Configuring the Bandwidth of an Interface

- Optional.
- Generally, the bandwidth of an interface is the same as the speed of the interface.

Command	**bandwidth** *kilobits*
Parameter Description	*kilobits*: The value ranges from 1 to 2,147,483,647. The unit is kilo bits.
Defaults	Generally, the bandwidth of an interface matches the type of the interface. For example, the default bandwidth of a gigabit Ethernet physical port is 1,000,000, and that of a 10G Ethernet physical port is 10,000,000.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Carrier Delay of an Interface

- Optional.
- If the configured carrier delay is long, it takes a long time to change the protocol status when the physical status of an interface changes. If the carrier delay is set to 0, the protocol status changes immediately after the physical status of an interface changes.

Command	**carrier-delay {[milliseconds]** *num* **| up [milliseconds]** *num* **down [milliseconds]** *num***}**
Parameter Description	*num*: The value ranges from 0 to 60. The unit is second. ***milliseconds***: Indicates the carrier delay. The value ranges from 0 to 60,000. The unit is millisecond. ***Up***: Indicates the delay after which the state of the DCD changes from Down to Up. ***Down***: Indicates the delay after which the state of the DCD changes from Up to Down.
Defaults	By default, the carrier delay of an interface is 2s.
Command Mode	Interface configuration mode
Usage Guide	If millisecond is used as the unit, the configured carrier delay must be an integer multiple of 100 milliseconds.

Configuring the Load Interval of an Interface

- Optional.
- The configured load interval affects computation of the average packet rate on an interface. If the configured load interval is short, the average packet rate can accurately reflect the changes of the real-time traffic.

Command	**load-interval** *seconds*
Parameter Description	*seconds*: The value ranges from 5 to 600. The unit is second.
Defaults	By default, the load interval of an interface is 10s.

Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring a Protected Port

- Optional.
- L2 packets cannot be forwarded between protected ports.
- This command is applicable to an Ethernet physical port or AP port.

Command	**switchport protected**
Parameter Description	N/A
Defaults	By default, no protected port is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Blocking L3 Routing Between Protected Ports

- Optional.
- After this command is configured, L3 routing between protected ports are blocked.

Command	**protected-ports route-deny**
Parameter Description	N/A
Defaults	By default, the function of blocking L3 routing between protected ports is disabled.
Command Mode	Global configuration mode
Usage Guide	By default, L3 routing between protected ports is not blocked. In this case, you can run this command to block routing between protected ports.

Configuring Port Errdisable Recovery

- Optional.
- By default, a port will be disabled and will not be recovered after a violation occurs. After port errdisable recovery is configured, a port in errdisable state will be recovered and enabled.

Command	**errdisable recovery [interval** *time* **]**
Parameter Description	*time*: Indicates the automatic recovery time. The value ranges from 30 to 86,400. The unit is second.
Defaults	By default, port errdisable recovery is disabled.
Command Mode	Global configuration mode
Usage Guide	By default, a port in errdisable state is not recovered. You can recover the port manually or run this command to automatically recover the port.

Optical Module Antifake Detection

- (Optional) Run this command to enable optical module antifake detection.
- Optical module antifake detection is disabled by default, and the system does not display any alarm if a non-FS optical module is inserted. After this function is enabled, the system will display alarms for several times if a non-FS optical module is inserted.

Command	**fiber antifake { ignore | enable }**
Parameter Description	ignore: Disables the optical module antifake detection function in global configuration mode. enable: Enables the optical module antifake detection function in global configuration mode.
Defaults	By default, optical module antifake detection is disabled.
Command Mode	Global configuration mode
Usage Guide	You can run the **fiber antifake enable** command to enable optical module antifake detection.

Configuring EEE

- Optional.
- The EEE mode of a port is enabled after this command is configured.

Command	**eee enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	By default, the EEE mode of a port is disabled. You can run this command to enable EEE, and use the no or default form of the command to disable EEE.

Verification

- Run the **show interfaces** command to display the attribute configurations of interfaces.

Command	**show interfaces** [*interface-type interface-number*] [**description | switchport | trunk**]
Parameter Description	*interface-type interface-number*: Indicates the type and number of the interface. **description**: Indicates the interface description, including the link status. **switchport**: Indicates the L2 interface information. This parameter is effective only for a L2 interface. **trunk**: Indicates the Trunk port information. This parameter is effective for a physical port or an AP port.
Command Mode	Privileged EXEC mode
Usage Guide	Use this command without any parameter to display the basic interface information.
	SwitchA#show interfaces GigabitEthernet 0/1 Index(dec):1 (hex):1 GigabitEthernet 0/1 is DOWN, line protocol is DOWN Hardware is Broadcom 5464 GigabitEthernet, address is 00d0.f865.de9b (bia 00d0.f865.de9b) Interface address is: no ip address Interface IPv6 address is: No IPv6 address

```
MTU 1500 bytes, BW 1000000 Kbit
Encapsulation protocol is Ethernet-II, loopback not set
Keepalive interval is 10 sec, set
Carrier delay is 2 sec
Ethernet attributes:
    Last link state change time: 2012-12-22 14:00:48
    Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds
    Priority is 0
    Medium-type is Copper
    Admin duplex mode is AUTO, oper duplex is Unknown
    Admin speed is AUTO, oper speed is Unknown
    Flow receive control admin status is OFF,flow send control admin status is OFF
    Flow receive control oper status is Unknown,flow send control oper status is Unknown
    Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
Bridge attributes:
    Port-type: trunk
    Native vlan:1
    Allowed vlan lists:1-4094    //Allowed VLAN list of the Trunk port
    Active vlan lists:1, 3-4    //Active VLAN list (indicating that only VLAN 1, VLAN 3, and VLAN 4 are created on the device)
Rxload is 1/255,Txload is 1/255
5 minutes input rate 0 bits/sec, 0 packets/sec
5 minutes output rate 0 bits/sec, 0 packets/sec
    0 packets input, 0 bytes, 0 no buffer, 0 dropped
    Received 0 broadcasts, 0 runts, 0 giants
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
    0 packets output, 0 bytes, 0 underruns, 0 dropped
    0 output errors, 0 collisions, 0 interface resets
```

- Run the **show eee interfaces status** command to display the EEE status of an interface.

Command	**show eee interfaces** { *interface-type interface-number* | *status* }
Parameter Description	*interface-type interface-number*: Indicates the type and number of an interface. ***status*:** Indicates the EEE status of all interfaces.
Command Mode	Privileged EXEC mode
Usage Guide	If the interface is specified, the EEE status of the specified interface is displayed; otherwise, the EEE status of all

interfaces is displayed.

1. Display the EEE status of GigabitEthernet 0/1.

```
FS#show eee interface gigabitEthernet 0/1
Interface                    : Gi0/1
EEE Support                  : Yes
Admin Status                 : Enable
Oper Status                  : Disable
Remote Status                : Disable
Trouble Cause                : Remote Disable
```

Interface	Indicates the interface information.
EEE Support	Indicates whether EEE is supported.
Admin Status	Indicates the administrative status.
Oper Status	Indicates the operational status.
Trouble Cause	Indicates the reason why the EEE status of an interface abnormal.

2. Display the EEE status of all interfaces.

```
FS#show eee interface status
Interface EEE      Admin      Oper       Remote     Trouble
          Support Status     Status     Status     Cause
--------- ------- -------- -------- -------- --------------------
Gi0/1     Yes     Enable     Disable    Disable    Remote Disable
Gi0/2     Yes     Enable     Disable    Unknown    None
Gi0/3     Yes     Enable     Enable     Enable     None
Gi0/4     Yes     Enable     Enable     Enable     None
Gi0/5     Yes     Enable     Enable     Enable     None
Gi0/6     Yes     Enable     Enable     Enable     None
Gi0/7     Yes     Enable     Enable     Enable     None
Gi0/8     Yes     Enable     Enable     Enable     None
Gi0/9     Yes     Enable     Enable     Enable     None
Gi0/10    Yes     Enable     Enable     Enable     None
```

Interface	Indicates the interface information.
EEE Support	Indicates whether EEE is supported.
Admin Status	Indicates the administrative status.
Oper Status	Indicates the operational status.
Trouble Cause	Indicates the reason why the EEE status of an interface abnormal.

- Run the **show interfaces brief** command to display transmission rates of ports.

Command	**show interfaces** [*interface-type interface-number*] **brief [up | down]**
Parameter Description	*interface-type interface-number*: Specifies an interface. If this field is not specified, information of all ports are displayed. **up**: Indicates connected ports. **down**: indicates disconnected ports.
Command Mode	Privileged EXEC mode
Usage Guide	By running this command, brief information of ports is displayed, including port status, interface status, output and input bandwidth usage, and the number of output and input packet errors.

1. Display the brief information of the TenGigabitEthernet 0/1 port.

```
FS#show interfaces TenGigabitEthernet 0/1 brief
down: link down
*down: administratively down
disabled: err-disabled(Please reference to command [show interface status err-disabled] for detail.)
Interface    Link Stat    Protocol Stat    Output Usage    Input Usage    inErrors    outErrors
--------    ---------    --------------    --------------    -------------    ----------    ----------
Te0/1    disabled    down    0.00%    0.00%    0    0
```

2. Display the brief information of all connected ports.

```
FS#show interfaces brief up
down: link down
*down: administratively down
disabled: err-disabled(Please reference to command [show interface status err-disabled] for detail.)
Interface    Link Stat    Protocol Stat    Output Usage    Input Usage    inErrors    outErrors
--------    ---------    --------------    --------------    -------------    ----------    ----------
Te0/1    up    up    83.40%    78.20%    0    0
Te0/2    up    up    82.00%    73.40%    0    0
```

3. Displays the brief information of all the ports.

```
FS#show interfaces brief
down: link down
*down: administratively down
disabled: err-disabled(Please reference to command [show interface status err-disabled] for detail.)
Interface    Link Stat    Protocol Stat    Output Usage    Input Usage    inErrors    outErrors
--------    ---------    --------------    --------------    -------------    ----------    ----------
Te0/1    down    down    0.00%    0.00%    0    0
Te0/2    down    down    0.00%    0.00%    0    0
```

```
Te0/3      down      down      0.00%      0.00%      0      0
Te0/4      down      down      0.00%      0.00%      0      0
Te0/5      down      down      0.00%      0.00%      0      0
Te0/6      down      down      0.00%      0.00%      0      0
Te0/7      down      down      0.00%      0.00%      0      0
Te0/8      down      down      0.00%      0.00%      0      0
Te0/9      down      down      0.00%      0.00%      0      0
Te0/10     disabled  down      0.00%      0.00%      0      0
```

- Run the **show interfaces ethernet brief** command to display transmission rates of ports.

Command	**show interfaces** [*interface-type interface-number*] **ethernet brief [up | down]**
Parameter Description	*interface-type interface-number*: Specifies an interface. If this field is not specified, information of all ports are displayed. **up**: Indicates connected ports. **down**: indicates disconnected ports.
Command Mode	Privileged EXEC mode
Usage Guide	By running this command, brief information of all physical, aggregation and management ports is displayed, including status, VLANs, auto negotiation, duplex mode, speed, bandwidth usage, and description.

1. Display the brief information of the GigabitEthernet 0/1 port.

```
FS#show interfaces TenGigabitEthernet 0/1 ethernet brief
down: link down
*down: administratively down
disabled: err-disabled(Please reference to command [show interface status err-disabled] for detail.)
Interface    Link Stat    Vlan    Auto-Neg    Duplex    Speed    Input Usage    Output Usage    Description
---------    ---------    ----    --------    -------   -------  -----------    ------------    -----------
Gi0/1        down         1       OFF         Unknown   Unknown  0.00%          0.00%           10G port
```

2. Display the brief information of all connected ports.

```
FS#show interfaces ethernet brief up
down: link down
*down: administratively down
disabled: err-disabled(Please reference to command [show interface status err-disabled] for detail.)
Interface    Link Stat    Vlan    Auto-Neg    Duplex    Speed    Input Usage    Output Usage    Description
---------    ---------    ----    --------    -------   -------  -----------    ------------    -----------
Gi0/1        UP           1       OFF         Full      1000M    79.77%         79.77%          10G port
```

3. Displays the brief information of all the ports.

```
FS#show interfaces ethernet brief
```

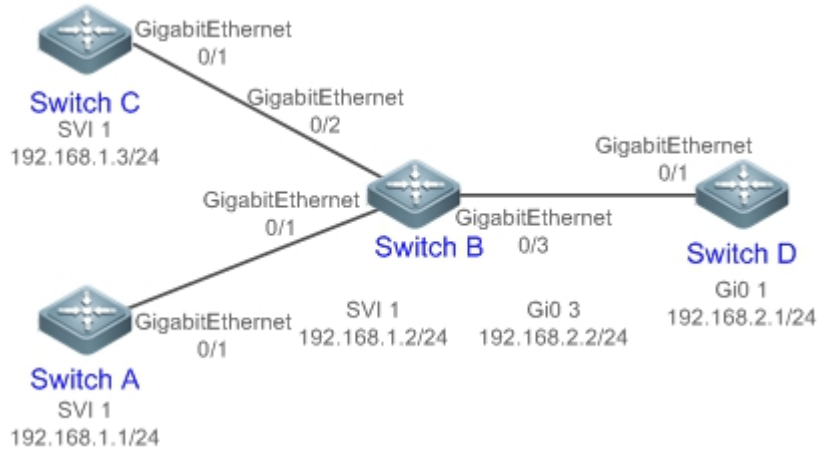
```
down: link down
*down: administratively down
disabled: err-disabled(Please reference to command [show interface status err-disabled] for detail.)
Interface    Link Stat    Vlan    Auto-Neg    Duplex    Speed    Input Usage    Output Usage    Description
---------    ---------    ----    --------    -------   -------  -----------    ------------    -----------
Gi0/1        *down        1       OFF         Unknown   Unknown  0.00%          0.00%           10G port
Gi0/2        down         1       OFF         Unknown   Unknown  0.00%          0.00%
Gi0/3        down         1       OFF         Unknown   Unknown  0.00%          0.00%
Ag1          up           1       OFF         Full      1000M    46.78%         46.77%
Mg0          up           routed  --          Full      1000M    --             --              IP management
Console
```

Configuration Example

Configuring Interface Attributes

Scenario

Figure 1-6



Configuration Steps

- On Switch A, configure GigabitEthernet 0/1 as an access mode, and the default VLAN ID is 1. Configure SVI 1, assign an IP address to SVI 1, and set up a route to Switch D.
- On Switch B, configure GigabitEthernet 0/1 and GigabitEthernet 0/2 as Trunk ports, and the default VLAN ID is 1. Configure SVI 1, and assign an IP address to SVI 1. Configure GigabitEthernet 0/3 as a routed port, and assign an IP address from another network segment to this port.
- On Switch C, configure GigabitEthernet 0/1 as an Access port, and the default VLAN ID is 1. Configure SVI 1, and assign an IP address to SVI 1.
- On Switch D, configure GigabitEthernet 0/1 as a routed port, assign an IP address to this port, and set up a route to Switch A.

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# switchport mode access
```

	A(config-if-GigabitEthernet 0/1)# switchport access vlan 1 A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip address 192.168.1.1 255.255.255.0 A(config-if-VLAN 1)# exit A(config)# ip route 192.168.2.0 255.255.255.0 VLAN 1 192.168.1.2
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# switchport mode trunk B(config-if-GigabitEthernet 0/1)# exit B(config)# interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)# switchport mode trunk B(config-if-GigabitEthernet 0/2)# exit B(config)# interface vlan 1 B(config-if-VLAN 1)# ip address 192.168.1.2 255.255.255.0 B(config-if-VLAN 1)# exit B(config)# interface GigabitEthernet 0/3 B(config-if-GigabitEthernet 0/3)# no switchport B(config-if-GigabitEthernet 0/3)# ip address 192.168.2.2 255.255.255.0 B(config-if-GigabitEthernet 0/3)# exit
C	C# configure terminal C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# port-group 1 C(config-if-GigabitEthernet 0/1)# exit C(config)# interface aggregateport 1 C(config-if-AggregatePort 1)# switchport mode access C(config-if-AggregatePort 1)# switchport access vlan 1 C(config-if-AggregatePort 1)# exit C(config)# interface vlan 1 C(config-if-VLAN 1)# ip address 192.168.1.3 255.255.255.0 C(config-if-VLAN 1)# exit
D	D# configure terminal D(config)# interface GigabitEthernet 0/1

	D(config-if-GigabitEthernet 0/1)# no switchport D(config-if-GigabitEthernet 0/1)# ip address 192.168.2.1 255.255.255.0 D(config-if-GigabitEthernet 0/1)# exit A(config)# ip route 192.168.1.0 255.255.255.0 GigabitEthernet 0/1 192.168.2.2
Verification	Perform verification on Switch A, Switch B, Switch C, and Switch D as follows: ● On Switch A, ping the IP addresses of interfaces of the other three switches. Verify that you can access the other three switches on Switch A.. ● Verify that switch B and Switch D can be pinged mutually. ● Verify that the interface status is correct.
A	A# show interfaces gigabitEthernet 0/1 Index(dec):1 (hex):1 GigabitEthernet 0/1 is UP, line protocol is UP Hardware is GigabitEthernet, address is 00d0.f865.de90 (bia 00d0.f865.de90) Interface address is: no ip address MTU 1500 bytes, BW 100000 Kbit Encapsulation protocol is Ethernet-II, loopback not set Keepalive interval is 10 sec, set Carrier delay is 2 sec Ethernet attributes: Last link state change time: 2012-12-22 14:00:48 Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds Priority is 0 Admin medium-type is Copper, oper medium-type is Copper Admin duplex mode is AUTO, oper duplex is Full Admin speed is AUTO, oper speed is 100M Flow control admin status is OFF, flow control oper status is OFF Admin negotiation mode is OFF, oper negotiation state is ON Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF **Bridge attributes:** **Port-type: access** **Vlan id: 1** Rxload is 1/255, Txload is 1/255 10 seconds input rate 0 bits/sec, 0 packets/sec 10 seconds output rate 67 bits/sec, 0 packets/sec

	362 packets input, 87760 bytes, 0 no buffer, 0 dropped Received 0 broadcasts, 0 runts, 0 giants 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort 363 packets output, 82260 bytes, 0 underruns, 0 dropped 0 output errors, 0 collisions, 0 interface resets
B	B# show interfaces gigabitEthernet 0/1 Index(dec):1 (hex):1 GigabitEthernet 0/1 is UP, line protocol is UP Hardware is GigabitEthernet, address is 00d0.f865.de91 (bia 00d0.f865.de91) Interface address is: no ip address MTU 1500 bytes, BW 100000 Kbit Encapsulation protocol is Ethernet-II, loopback not set Keepalive interval is 10 sec, set Carrier delay is 2 sec Ethernet attributes: Last link state change time: 2012-12-22 14:00:48 Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds Priority is 0 Admin medium-type is Copper, oper medium-type is Copper Admin duplex mode is AUTO, oper duplex is Full Admin speed is AUTO, oper speed is 100M Flow control admin status is OFF, flow control oper status is OFF Admin negotiation mode is OFF, oper negotiation state is ON Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF Bridge attributes: **Port-type: trunk** **Native vlan: 1** **Allowed vlan lists: 1-4094** **Active vlan lists: 1** Rxload is 1/255, Txload is 1/255 10 seconds input rate 0 bits/sec, 0 packets/sec 10 seconds output rate 67 bits/sec, 0 packets/sec 362 packets input, 87760 bytes, 0 no buffer, 0 dropped

	Received 0 broadcasts, 0 runts, 0 giants 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort 363 packets output, 82260 bytes, 0 underruns, 0 dropped 0 output errors, 0 collisions, 0 interface resets
C	C# show interfaces gigabitEthernet 0/1 Index(dec):1 (hex):1 GigabitEthernet 0/1 is UP, line protocol is UP Hardware is GigabitEthernet, address is 00d0.f865.de92 (bia 00d0.f865.de92) Interface address is: no ip address MTU 1500 bytes, BW 100000 Kbit Encapsulation protocol is Ethernet-II, loopback not set Keepalive interval is 10 sec, set Carrier delay is 2 sec Ethernet attributes: Last link state change time: 2012-12-22 14:00:48 Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds Priority is 0 Admin medium-type is Copper, oper medium-type is Copper Admin duplex mode is AUTO, oper duplex is Full Admin speed is AUTO, oper speed is 100M Flow control admin status is OFF, flow control oper status is OFF Admin negotiation mode is OFF, oper negotiation state is ON Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF Rxload is 1/255, Txload is 1/255 10 seconds input rate 0 bits/sec, 0 packets/sec 10 seconds output rate 67 bits/sec, 0 packets/sec 362 packets input, 87760 bytes, 0 no buffer, 0 dropped Received 0 broadcasts, 0 runts, 0 giants 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort 363 packets output, 82260 bytes, 0 underruns, 0 dropped 0 output errors, 0 collisions, 0 interface resets
D	D# show interfaces gigabitEthernet 0/1 Index(dec):1 (hex):1

```
GigabitEthernet 0/1 is UP, line protocol is UP
Hardware is GigabitEthernet, address is 00d0.f865.de93 (bia 00d0.f865.de93)
Interface address is: 192.168.2.1/24
  MTU 1500 bytes, BW 100000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec, set
  Carrier delay is 2 sec
  Ethernet attributes:
    Last link state change time: 2012-12-22 14:00:48
    Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds
    Priority is 0
    Admin medium-type is Copper, oper medium-type is Copper
    Admin duplex mode is AUTO, oper duplex is Full
    Admin speed is AUTO, oper speed is 100M
    Flow control admin status is OFF, flow control oper status is OFF
    Admin negotiation mode is OFF, oper negotiation state is ON
    Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
  Rxload is 1/255, Txload is 1/255
  10 seconds input rate 0 bits/sec, 0 packets/sec
  10 seconds output rate 67 bits/sec, 0 packets/sec
    362 packets input, 87760 bytes, 0 no buffer, 0 dropped
    Received 0 broadcasts, 0 runts, 0 giants
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
    363 packets output, 82260 bytes, 0 underruns, 0 dropped
    0 output errors, 0 collisions, 0 interface resets
```

1.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the counters of a specified interface.	**clear counters** [*interface-type interface-number*]
Resets the interface hardware.	**clear interface** *interface-type interface-number*

Displaying

Displaying Interface Configurations and Status

Description	Command
Displays all the status and configuration information of a specified interface.	**show interfaces** [*interface-type interface-number*]
Displays the interface status.	**show interfaces** [*interface-type interface-number*] **status**
Displays the interface errdisable status.	**show interfaces** [*interface-type interface-number*] **status err-disable**
Displays the link status change time and count of a specified port.	**show interfaces** [*interface-type interface-number*] **link-state-change statistics**
Displays the administrative and operational states of switch ports (non-routed ports).	**show interfaces** [*interface-type interface-number*] **switchport**
Displays the description and status of a specified interface.	**show interfaces** [*interface-type interface-number*] **description**
Displays the counters of a specified port, among which the displayed speed may have an error of ±0.5%.	**show interfaces** [*interface-type interface-number*] **counters**
Displays the number of packets increased in a load interval.	**show interfaces** [*interface-type interface-number*] **counters increment**
Displays statistics about error packets.	**show interfaces** [*interface-type interface-number*] **counters error**
Displays the packet sending/receiving rate of an interface.	**show interfaces** [*interface-type interface-number*] **counters rate**
Displays a summary of interface information.	**show interfaces** [*interface-type interface-number*] **counters summary**
Displays the line detection status. When a cable is short-circuited or disconnected, line detection helps you correctly determine the working status of the cable.	**show interfaces** [*interface-type interface-number*] **line-detect**
Displays the bandwidth usage of an interface.	**show interfaces** [*interface-type interface-number*] **usage**
Displays brief information of ports.	**show interfaces** [*interface-type interface-number*] **brief [up | down]**
Displays the EEE status of an interface.	**show eee interfaces** { *interface-type interface-number* | *status* }
Displays configurations.	**show running-config**

Displaying Optical Module Information

Description	Command
Displays basic information about the optical module of a specified interface.	**show interfaces** [*interface-type interface-number*] **transceiver**
Displays the fault alarms of the optical module on a specified interface. If no fault occurs, "None" is displayed.	**show interfaces** [*interface-type interface-number*] **transceiver alarm**
Displays the optical module diagnosis values of a specified interface.	**show interfaces** [*interface-type interface-number*] **transceiver diagnosis**

Line Detection

The administrator can run the **line-detect** command to check the working status of a cable. When a cable is short-circuited or disconnected, line detection helps you determine the working status of the cable.

Only a physical port using copper as the medium supports line detection. A physical port using fiber as the medium or an AP port does not support line detection.

When line detection is performed on an operational interface, the interface will be temporarily disconnected, and then re-connected.

Description	Command
Performs line detection in interface configuration mode. When a cable is short-circuited or disconnected, line detection helps you determine the working status of the cable.	**line-detect**

2 Configuring MAC Address

2.1 Overview

A MAC address table contains the MAC addresses, interface numbers and VLAN IDs of the devices connected to the local device.

When a device forwards a packet, it finds an output port from its MAC address table according to the destination MAC address and the VLAN ID of the packet.

After that, the packet is unicast, multicast or broadcast.

This document covers dynamic MAC addresses, static MAC addresses and filtered MAC addresses. For the management of multicast MAC addresses, please see *Configuring IGMP Snooping Configuration*.

Protocols and Standards

- IEEE 802.3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications
- IEEE 802.1Q: Virtual Bridged Local Area Networks

2.2 Applications

Application	Description
MAC Address Learning	Forward unicast packets through MAC addresses learning.
MAC Address Change Notification	Monitor change of the devices connected to a network device through MAC address change notification.

2.2.1 MAC Address Learning

Scenario

Usually a device maintains a MAC address table by learning MAC addresses dynamically. The operating principle is described as follows:

As shown in the following figure, the MAC address table of the switch is empty. When User A communicates with User B, it sends a packet to the port GigabitEthernet 0/2 of the switch, and the switch learns the MAC address of User A and stores it in the table.
As the table does not contain the MAC address of User B, the switch broadcasts the packet to the ports of all connected devices except User A, including User B and User C.

Figure 2- 1 Step 1 of MAC Address Learning

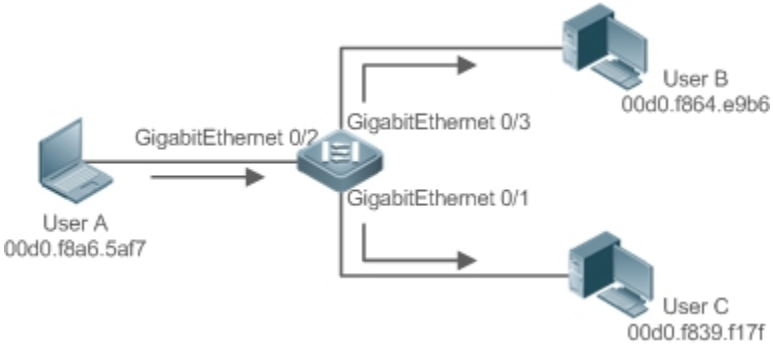


Figure 2- 2 MAC Address Table 1

Status	VLAN	MAC address	Interface
Dynamic	1	00d0.f8a6.5af7	GigabitEthernet 0/2

When User B receives the packet, it sends a reply packet to User A through port GigabitEthernet 0/3 on the switch. As the MAC address of User A is already in the MAC address table, the switch send the reply unicast packet to port GigabitEthernet 0/2 port and learns the MAC address of User B. User C does not receive the reply packet from User B to User A.

Figure 2- 3 Step 2 of MAC Address Learning

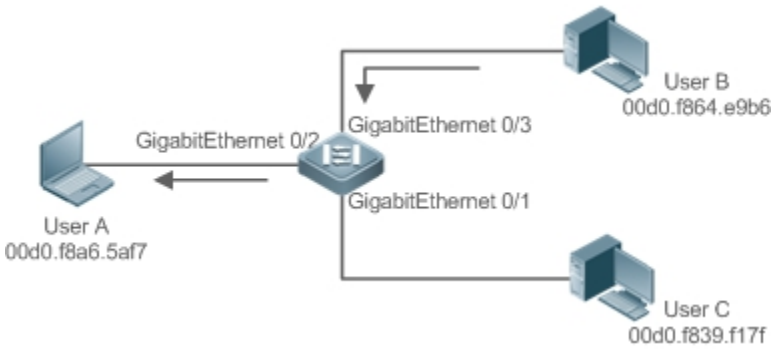


Figure 2- 4 MAC Address Table 2

Status	VLAN	MAC address	Interface
Dynamic	1	00d0.f8a6.5af7	GigabitEthernet 0/2
Dynamic	1	00d0.f8a4.e9b6	GigabitEthernet 0/3

Through the interaction between User A and User B, the switch learns the MAC addresses of User A and User B. After that, packets between User A and User B will be exchanged via unicast without being received by User C.

Deployment

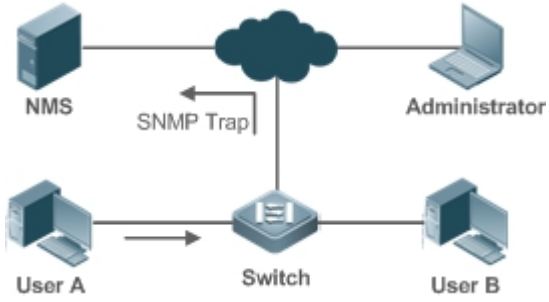
- With MAC address learning, a layer-2 switch forwards packets through unicast, reducing broadcast packets and network load.

2.2.2 MAC Address Change Notification

MAC address change notification provides a mechanism for the network management system (NMS) to monitor the change of devices connected to a network device.

Scenario

Figure 2- 5 MAC Address Change Notification



After MAC address change notification is enabled on a device, the device generates a notification message when the device learns a new MAC address or finishes aging a learned MAC address, and sends the message in an SNMP Trap message to a specified NMS.

A notification of adding a MAC address indicates that a new user accesses the network, and that of deleting a MAC address indicates that a user sends no packets within an aging time and usually the user exits the network.

When a network device is connected to a number of devices, a lot of MAC address changes may occur in a short time, resulting in an increase in traffic. To reduce traffic, you may configure an interval for sending MAC address change notifications. When the interval expires, all notifications generated during the interval are encapsulated into a message.

±When a notification is generated, it is stored in the table of historical MAC address change notifications. The administrator may know recent MAC address changes by checking the table of notification history even without NMS.

A MAC address change notification is generated only for a dynamic MAC address.

Deployment

- Enable MAC address change notification on a layer-2 switch to monitor the change of devices connected to a network device.

2.3 Features

Basic Concepts

Dynamic MAC Address

A dynamic MAC address is a MAC address entry generated through the process of MAC address learning by a device.

Address Aging

A device only learns a limited number of MAC addresses, and inactive entries are deleted through address aging.

A device starts aging a MAC address when it learns it. If the device receives no packet containing the source MAC address, it will delete the MAC address from the MAC address table when the time expires.

Forwarding via Unicast

If a device finds in its MAC address table an entry containing the MAC address and the VLAN ID of a packet and the output port is unique, it will send the packet through the port directly.

Forwarding via Broadcast

If a device receives a packet containing the destination address ffff.ffff.ffff or an unidentified destination address, it will send the packet through all the ports in the VLAN where the packet is from, except the input port.

Overview

Feature	Description
Dynamic Address Limit for VLAN	Limit the number of dynamic MAC addresses in a VLAN.
Dynamic Address Limit for Interface	Limit the number of dynamic MAC addresses on an interface.

2.3.1 Dynamic Address Limit for VLAN

Working Principle

The MAC address table with a limited capacity is shared by all VLANs. Configure the maximum number of dynamic MAC addresses for each VLAN to prevent one single VLAN from exhausting the MAC address table space.

A VLAN can only learn a limited number of dynamic MAC addresses after the limit is configured. The packets exceeding the limit are

forwarded.User can configure the maximum MAC addresses learned by a VLAN. After the maximum number exceeds the limit, the VLAN will stop learning MAC address, and packets will be discarded.

If the number of learned MAC addresses is greater than the limit, a device will stop learning the MAC addresses from the VLAN and will not start learning again until the number drops below the limit after address aging.

The MAC addresses copied to a specific VLAN are not subject to the limit.

2.3.2 Dynamic Address Limit for Interface

Working Principle

An interface can only learn a limited number of dynamic MAC addresses after the limit is configured. The packets exceeding the limit are forwarded.

User can configure the maximum MAC addresses learned by a VLAN. After the maximum number exceeds the limit, the VLAN will stop learning MAC address, and packets will be discarded.

If the number of learned MAC addresses is greater than the limit, a device will stop learning the MAC addresses from the interface and will not start learning again until the number drops below the limit after address aging.

2.4 Configuration

Configuration	Description and Command	
Configuring Dynamic MAC Address	(Optional) It is used to enable MAC address learning.	
	mac-address-learning	Configures MAC address learning globally or on an interface.
	mac-address-table aging-time	Configures an aging time for a dynamic MAC address.
Configuring a Static MAC Address	(Optional) It is used to bind the MAC address of a device with a port of a switch.	
	mac-address-table static	Configures a static MAC address.
Configuring a MAC Address for Packet Filtering	(Optional) It is used to filter packets.	
	mac-address-table filtering	Configures a MAC address for packet filtering.
Configuring MAC Address Change Notification	(Optional) It is used to monitor change of devices connected to a network device.	
	mac-address-table notification	Configures MAC address change notification globally.
	snmp trap mac-notification	Configures MAC address change notification on an interface.
Configuring Maximum Number of MAC Addresses Learned by a VLAN	(Optional) It is used to configure the maximum number of MAC addresses learned by a VLAN/port.	
	max-dynamic-mac-count *count*	Configures the maximum number of MAC addresses learned by a VLAN/port.

2.4.1 Configuring Dynamic MAC Address

Configuration Effect

Learn MAC addresses dynamically and forward packets via unicast.

Configuration Steps

Configuring Global MAC Address Learning

- Optional.
- You can perform this configuration to disable global MAC address learning.
- Configuration:

Command	**mac-address-learning** { **enable** | **disable** }
Parameter Description	**enable**: Enables global MAC address learning. **disable**: Disable global MAC address learning.
Defaults	Global MAC address learning is enabled by default.
Command Mode	Global configuration mode
Usage Guide	N/A

By default, global MAC address learning is enabled. When global MAC address learning is enabled, the MAC address learning configuration on an interface takes effect; when the function is disabled, MAC addresses cannot be learned globally.

Configuring MAC Address Learning on Interface

- Optional.
- You can perform this configuration to disable MAC address learning on an interface.
- Configuration:

Command	**mac-address-learning**
Parameter Description	**N/A**
Defaults	MAC address learning is enabled by default.
Command Mode	Interface configuration mode
Usage Guide	Perform this configuration on a layer-2 interface, for example, a switch port or an AP port.

By default, MAC address learning is enabled. If DOT1X, IP SOURCE GUARD, or a port security function is configured on a port, MAC address learning cannot be enabled. Access control cannot be enabled on a port with MAC address learning disabled.

Configuring an Aging Time for a Dynamic MAC Address

- Optional.
- Configure an aging time for dynamic MAC addresses.
- Configuration:

Command	**mac-address-table aging-time** *value*

Parameter Description	*value*: Indicates the aging time. The value is either 0 or in the range from 10 to 1000,000.
Defaults	The default is 300s.
Command Mode	Global configuration mode
Usage Guide	If the value is set to 0, MAC address aging is disabled and learned MAC addresses will not be aged.

The actual aging time may be different from the configured value, but it is not more than two times of the configured value.

Verification

- Check whether a device learns dynamic MAC addresses.
- Run the **show mac-address-table dynami**c command to display dynamic MAC addresses.
- Run the **show mac-address-table aging-time** command to display the aging time for dynamic MAC addresses.

Command	**show mac-address-table dynamic** [**address** *mac-address*] [**interface** *interface-id*] [**vlan** *vlan-id*]
Parameter Description	**address** *mac-address*: Displays the information of a specific dynamic MAC address. **interface** *interface-id*: Specifies a physical interface or an AP port. **vlan** *vlan-id*: Displays the dynamic MAC addresses in a specific VLAN.
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A

```
FS# show mac-address-table dynamic

Vlan        MAC Address          Type          Interface

----        ------------         ------        ------------------

1           0000.0000.0001       DYNAMIC    GigabitEthernet 1/1

1           0001.960c.a740       DYNAMIC    GigabitEthernet 1/1

1           0007.95c7.dff9       DYNAMIC    GigabitEthernet 1/1

1           0007.95cf.eee0       DYNAMIC    GigabitEthernet 1/1

1           0007.95cf.f41f       DYNAMIC    GigabitEthernet 1/1

1           0009.b715.d400       DYNAMIC    GigabitEthernet 1/1

1           0050.bade.63c4       DYNAMIC    GigabitEthernet 1/1
```

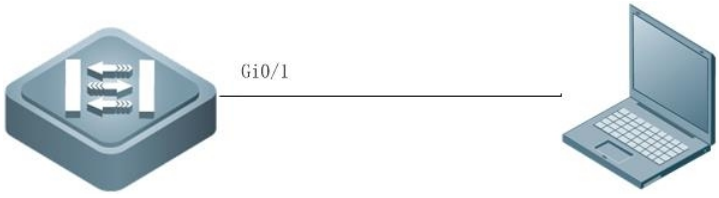
Field	Description
Vlan	Indicates the VLAN where the MAC address resides.
MAC Address	Indicates a MAC Address.
Type	Indicates a MAC address type.
Interface	Indicates the interface where the MAC address resides.

Command	**show mac-address-table aging-time**
Parameter Description	N/A

Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A
	FS# show mac-address-table aging-time **Aging time: 300**

Configuration Example

Configuring Dynamic MAC Address

Scenario **Figure 2- 6**	
Configuration Steps	● Enable MAC address learning on an interface. ● Configure the aging time for dynamic MAC addresses to 180s. ● Delete all dynamic MAC addresses in VLAN 1 on port GigabitEthernet 0/1.
	FS# configure terminal FS(config-if-GigabitEthernet 0/1)# mac-address-learning FS(config-if-GigabitEthernet 0/1)# exit FS(config)# mac aging-time 180 FS# clear mac-address-table dynamic interface GigabitEthernet 0/1 vlan 1
Verification	● Check MAC address learning on an interface. ● Display the aging time for dynamic MAC addresses. ● Display all dynamic MAC addresses in VLAN 1 on port GigabitEthernet 0/1.
	FS# show mac-address-learning **GigabitEthernet 0/1 learning ability: enable** FS# show mac aging-time Aging time : **180 seconds** FS# show mac-address-table dynamic interface GigabitEthernet 0/1 vlan 1 Vlan MAC Address Type Interface ---------- ---------- -------- ---------- **1 00d0.f800.1001 STATIC GigabitEthernet 1/1**

Common Errors

Configure MAC address learning on an interface before configuring the interface as a layer-2 interface, for example, a switch port or an AP port.

2.4.2 Configuring a Static MAC Address

Configuration Effect

- Bind the MAC address of a network device with a port of a switch.

Configuration Steps

Configuring a Static MAC address

- Optional.
- Bind the MAC address of a network device with a port of a switch.
- Configuration:

Command	**mac-address-table static** *mac-address* **vlan** *vlan-id* **interface** *interface-id*
Parameter Description	**address** *mac-address*: Specifies a MAC address. **vlan** *vlan-id*: Specifies a VLAN where the MAC address resides. **interface** *interface-id*: Specifies a physical interface or an AP port.
Defaults	By default, no static MAC address is configured.
Command Mode	Global configuration mode
Usage Guide	When the switch receives a packet containing the specified MAC address on the specified VLAN, the packet is forwarded to the bound interface.

Verification

- Run the **show mac-address-table static** command to check whether the configuration takes effect.

Command	**show mac-address-table static** [**address** *mac-address*] [**interface** *interface-id*] [**vlan** *vlan-id*]
Parameter Description	**address** *mac-address*: Specifies a MAC address. **interface** *interface-id*: Specifies a physical interface or an AP port. **vlan** *vlan-id*: Specifies a VLAN where the MAC address resides.
Command Mode	Privileged EXEC mode/Global configuration mode /Interface configuration mode
Usage Guide	N/A
	FS# show mac-address-table static Vlan MAC Address Type Interface ----- ---------- -------- ---------- **1 00d0.f800.1001 STATIC GigabitEthernet 1/1** **1 00d0.f800.1002 STATIC GigabitEthernet 1/1** **1 00d0.f800.1003 STATIC GigabitEthernet 1/1**

Configuration Example

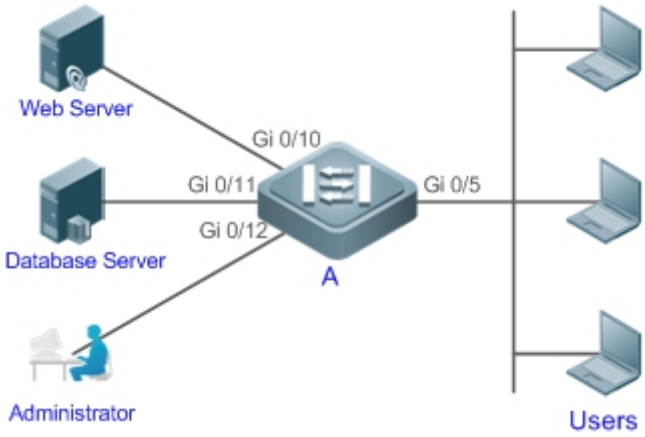
Configuring a Static MAC address

In the above example, the relationship of MAC addresses, VLAN and interfaces is shown in the following table.

Role	MAC Address	VLAN ID	Interface ID
Web Server	00d0.3232.0001	VLAN2	Gi0/10
Database Server	00d0.3232.0002	VLAN2	Gi0/11
Administrator	00d0.3232.1000	VLAN2	Gi0/12

Scenario

Figure 2- 7



Configuration Steps

- Specify destination MAC addresses (*mac-address*).
- Specify the VLAN (*vlan-id*) where the MAC addresses reside.
- Specify interface IDs (*interface-id*).

A

```
A# configure terminal
A(config)# mac-address-table static 00d0.f800.3232.0001 vlan 2 interface gigabitEthernet 0/10
A(config)# mac-address-table static 00d0.f800.3232.0002 vlan 2 interface gigabitEthernet 0/11
A(config)# mac-address-table static 00d0.f800.3232.1000 vlan 2 interface gigabitEthernet 0/12
```

Verification

Display the static MAC address configuration on a switch.

A

```
A# show mac-address-table static
Vlan        MAC Address          Type      Interface
----------  -------------------- -------- -------------------
   2        00d0.f800.3232.0001  STATIC   GigabitEthernet 0/10
   2        00d0.f800.3232.0002  STATIC   GigabitEthernet 0/11
   2        00d0.f800.3232.1000  STATIC   GigabitEthernet 0/12
```

Common Errors

- Configure a static MAC address before configuring the specific port as a layer-2 interface, for example, a switch port or an AP port.

2.4.3 Configuring a MAC Address for Packet Filtering

Configuration Effect

- If a device receives packets containing a source MAC address or destination MAC address specified as the filtered MAC address, the packets are discarded.

Configuration Steps

Configuring a MAC Address for Packet Filtering

- Optional.
- Perform this configuration to filter packets.
- Configuration:

Command	**mac-address-table filtering** *mac-address* **vlan** *vlan-id*
Parameter Description	**address** *mac-address*: Specifies a MAC address. **vlan** *vlan-id*: Specifies a VLAN where the MAC address resides.
Defaults	By default, no filtered MAC address is configured.
Command Mode	Global configuration mode
Usage Guide	If a device receives packets containing a source MAC address or destination MAC address specified as the filtered MAC address, the packets are discarded.

Verification

- Run the **show mac-address-table filter** command to display the filtered MAC address.

Command	**show mac-address-table filter** [**address** *mac-address*] [**vlan** *vlan-id*]
Parameter Description	**address** *mac-address*: Specifies a MAC address. **vlan** *vlan-id*: Specifies a VLAN where the MAC address resides.
Command Mode	Privileged EXEC mode/Global configuration mode /Interface configuration mode
Usage Guide	N/A
	FS# show mac-address-table filtering Vlan MAC Address Type Interface ------ ---------- -------- ---------- **1 0000.2222.2222 FILTER**

Configuration Example

Configuring a MAC Address for Packet Filtering

Configuration Steps	• Specify a destination MAC address (*mac-address*) for filtering. • Specify a VLAN where the MAC addresses resides.
	FS# configure terminal FS(config)# mac-address-table static 00d0.f800.3232.0001 vlan 1
Verification	Display the filtered MAC address configuration.
	FS# show mac-address-table filter Vlan MAC Address Type Interface

	---------- ---------- -------- ---------- **1 00d0.f800.3232.0001 FILTER**

2.4.4 Configuring MAC Address Change Notification

Configuration Effect

- Monitor change of devices connected to a network device.

Configuration Steps

Configuring NMS

- Optional.
- Perform this configuration to enable an NMS to receive MAC address change notifications.
- Configuration:

Command	**snmp-server host** *host-addr* **traps** [**version** { **1** | **2c** | **3** [**auth** | **noauth** | **priv**] }] *community-string*
Parameter Description	**host** *host-addr*: Specifies the IP address of a receiver. ● **version** { **1** | **2c** | **3** [**auth** | **noauth** | **priv**] }: Specifies the version of SNMP TRAP messages. You can also specify authentication and a security level for packets of Version 3. *community-string*: Indicates an authentication name.
Defaults	By default, the function is disabled.
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling SNMP Trap

- Optional.
- Perform this configuration to send SNMP Trap messages.
- Configuration:

Command	**snmp-server enable traps**
Parameter Description	N/A
Defaults	By default, the function is disabled.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring Global MAC Address Change Notification

- Optional.
- If MAC address change notification is disabled globally, it is disabled on all interfaces.
- Configuration:

Command	**mac-address-table notification**

Parameter Description	N/A
Defaults	By default, MAC address change notification is disabled globally.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring MAC Address Change Notification On Interface

- Optional.
- Perform this configuration to enable MAC address change notification on an interface.
- Configuration:

Command	**snmp trap mac-notification { added | removed }**
Parameter Description	**added**: Generates a notification when an MAC address is added. **removed**: Generates a notification when an MAC address is deleted.
Defaults	By default, MAC address change notification is disabled on an interface.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring Interval for Generating MAC Address Change Notifications and Volume of Notification History

- Optional.
- Perform this configuration to modify the interval for generating MAC address change notifications and the volume of notification history.
- Configuration:

Command	**mac-address-table notification** { **interval** *value* | **history-size** *value* }
Parameter Description	**interval** *value*: (Optional) Indicates the interval for generating MAC address change notifications. The value ranges from 1 to 3600 seconds,. **history-size** *value*: Indicates the maximum number of entries in the table of notification history. The value ranges from 1 to 200.
Defaults	The default interval is 1 second. The default maximum amount of notifications is 50.
Command Mode	Global configuration mode
Usage Guide	N/A

Verification

- Run the **show mac-address-table notification** command to check whether the NMS receives MAC address change notifications.

Command	**show mac-address-table notification [interface [*interface-id*] | history]**
Parameter Description	**Interface**:Displays the configuration of MAC address change notification on all interfaces. ***interface-id***: Displays the configuration of MAC address change notification on a specified interface. **history**: Displays the history of MAC address change notifications.
Command	Privileged EXEC mode/Global configuration mode /Interface configuration mode

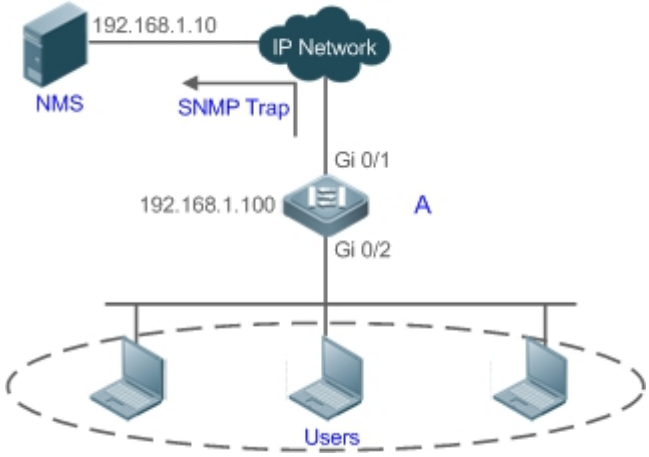
Mode	
Usage Guide	N/A
Usage Guide	Display the configuration of global MAC address change notification. FS#show mac-address-table notification MAC Notification Feature : **Enabled** Interval(Sec): **300** Maximum History Size : **50** Current History Size : **0**

Field	Description
Interval(Sec)	Indicates the interval for generating MAC address change notifications.
Maximum History Size	Indicates the maximum number of entries in the table of notification history.
Current History Size	Indicates the current notification entry number.

Configuration Example

Scenario

Figure 2- 8



The figure shows an intranet of an enterprise. Users are connected to A via port Gi0/2.

The Perform the configuration to achieve the following effects:

- When port Gi0/2 learns a new MAC address or finishes aging a learned MAC address, a MAC address change notification is generated.
- Meanwhile, A sends the MAC address change notification in an SNMP Trap message to a specified NMS.
- In a scenario where A is connected to a number of Users, the configuration can prevent MAC address change notification burst in a short time so as to reduce the network flow.

Configuration Steps

- Enable global MAC address change notification on A, and configure MAC address change notification on port Gi0/2.
- Configure the IP address of the NMS host, and enable A with SNMP Trap. A communicates with the NMS via routing.
- Configure the interval for sending MAC address change notifications to 300 seconds (1 second by default).

A	FS# configure terminal FS(config)# mac-address-table notification FS(config)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)# snmp trap mac-notification added FS(config-if-GigabitEthernet 0/2)# snmp trap mac-notification removed FS(config-if-GigabitEthernet 0/2)# exit FS(config)# snmp-server host 192.168.1.10 traps version 2c comefrom2 FS(config)# snmp-server enable traps FS(config)# mac-address-table notification interval 300
Verification	● Check t whether MAC address change notification is enabled globally . ● Check whether MAC address change notification is enabled on the interface. ● Display the MAC addresses of interfaces, and run the **clear mac-address-table dynamic** command to simulate aging dynamic MAC addresses. ● Check whether global MAC address change notification is enabled globally. ● Display the history of MAC address change notifications.
A	FS# show mac-address-table notification MAC Notification Feature : **Enabled** Interval(Sec): **300** Maximum History Size : **50** Current History Size : **0** FS# show mac-address-table notification interface GigabitEthernet 0/2 Interface MAC Added Trap MAC Removed Trap ---------- ---------- ---------- GigabitEthernet 0/2 Enabled Enabled FS# show mac-address-table interface GigabitEthernet 0/2 Vlan MAC Address Type Interface ---------- ---------- -------- ---------- 1 00d0.3232.0001 DYNAMIC GigabitEthernet 0/2 FS# show mac-address-table notification MAC Notification Feature : Enabled Interval(Sec): 300 Maximum History Size : 50 Current History Size : 1 FS# show mac-address-table notification history

	History Index : 0 Entry Timestamp: 221683 MAC Changed Message : Operation:DEL Vlan:1 MAC Addr: 00d0.3232.0003 GigabitEthernet 0/2

2.4.5 Configuring the Maximum Number of MAC Addresses Learned by a Port

Configuration Effect

- Only a limited number of dynamic MAC addresses can be learned by a port.

Notes

None

Configuration Steps

Configuring the Maximum Number of MAC Addresses Learned by a Port

- Optional
- Perform this operation on the switch.

Command	**max-dynamic-mac-count** *count*
Parameter Description	**count**: Indicates the maximum number of MAC addresses learned by a port.
Defaults	By default, the number of MAC addresses learned by a port is not limited. After the number of MAC addresses learned by a port is limited and after the maximum number of MAC addresses exceeds the limit, packets from source MAC addresses are forwarded by default.
Command Mode	Interface configuration mode
Usage Guide	

Verification

- Run **show run** to query the configuration result.

Configuration Example

Configuring the Maximum Number of MAC Addresses Learned by a Port

Configuration Steps	• Configure the maximum number of MAC addresses learned by a port.
	• Configure the maximum number of MAC addresses learned by a port and the countermeasure for the case that the number of MAC addresses exceeds the limit. FS(config)# interface GigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# max-dynamic-mac-count 100
Verification	Run **show running** on the switch to query the configuration.

Common Errors

None

2.5 Monitoring

Clearing

Running the clear commands may lose vital information and interrupt services.

Description	Command
Clears dynamic MAC addresses.	**clear mac-address-table dynamic** [**address** *mac-address*] [**interface** *interface-id*] [**vlan** *vlan-id*]

Displaying

Description	Command
Displays the MAC address table.	**show mac-address-table** { **dynamic** | **static** | **filter** } [**address** *mac-address*] [**interface** *interface-id*] [**vlan** *vlan-id*]
Displays the aging time for dynamic MAC addresses.	**show mac-address-table aging-time**
Displays the maximum number of dynamic MAC addresses.	**show mac-address-table max-dynamic-mac-count**
Displays the configuration and history of MAC address change notifications.	**show mac-address-table notification [interface [*interface-id*] | history]**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs MAC address operation.	**debug bridge mac**

3 Configuring Aggregated Port

3.1 Overview

An aggregated port (AP) is used to bundle multiple physical links into one logical link to increase the link bandwidth and improve connection reliability.

An AP port supports load balancing, namely, distributes load evenly among member links. Besides, an AP port realizes link backup. When a member link of the AP port is disconnected, the load carried by the link is automatically allocated to other functional member links. A member link does not forward broadcast or multicast packets to other member links.

For example, the link between two devices supports a maximum bandwidth of 1,000 Mbps. When the service traffic carried by the link exceeds 1,000 Mbps, the traffic in excess will be discarded. Port aggregation can be used to solve the problem. For example, you can connect the two devices with network cables and combine multiple links to form a logical link capable of multiples of 1,000 Mbps.

For example, there are two devices connected by a network cable. When the link between the two ports of the devices is disconnected, the services carried by the link will be interrupted. After the connected ports are aggregated, the services will not be affected as long as one link remains connected.

Protocols and Standards

- IEEE 802.3ad

3.2 Applications

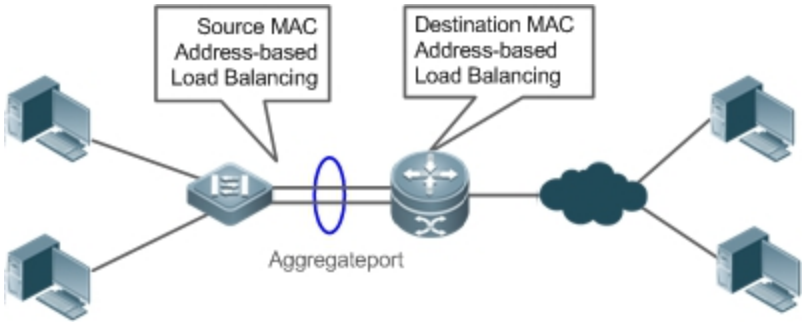
Applications	Description
AP Link Aggregation and Load Balancing	A large number of packets are transmitted between an aggregation device and a core device, which requires a greater bandwidth. To meet this requirement, you can bundle the physical links between the devices into one logical link to increase the link bandwidth, and configure a proper load balancing algorithm to distribute the work load evenly to each physical link, thus improving bandwidth utilization.

3.2.1 AP Link Aggregation and Load Balancing

Scenario

In Figure 3-1, the switch communicates with the router through an AP port. All the devices on the intranet (such as the two PCs on the left) use the router as a gateway. All the devices on the extranet (such as the two PCs on the right) send packets to the internet devices through the router, with the gateway's MAC address as its source MAC address. To distribute the load between the router and other hosts to other links, configure destination MAC address-based load balancing. On the switch, configure source MAC address-based load balancing.

Figure 3-1 AP Link Aggregation and Load Balancing



Deployment

- Configure the directly connected ports between the switch and router as a static AP port or a Link Aggregation Control Protocol (LACP) AP port.
- On the switch, configure a source MAC address-based load balancing algorithm.
- On the router, configure a destination MAC address-based load balancing algorithm.

3.3 Features

Basic Concepts

Static AP

The static AP mode is an aggregation mode in which physical ports are directly added to an AP aggregation group through manual configuration to allow the physical ports to forward packets when the ports are proper in link state and protocol state.

An AP port in static AP mode is called a static AP, and its member ports are called static AP member ports.

LACP

LACP is a protocol about dynamic link aggregation. It exchanges information with the connected device through LACP data units (LACPDUs).

An AP port in LACP mode is called an LACP AP port, and its member ports are called LACP AP member ports.

AP Member Port Mode

There are three aggregation modes available, namely, active, passive, and static.

AP member ports in active mode initiate LACP negotiation. AP member ports in passive mode only respond to received LACPDUs. AP member ports in static mode do not send LACPDUs for negotiation. The following table lists the requirements for peer port mode.

Port Mode	Peer Port Mode
Active mode	Active or passive mode
Passive mode	Active mode
Static Mode	Static Mode

AP Member Port State

There are two kinds of AP member port state available:

- When a member port is Down, the port cannot forward packets. The Down state is displayed.
- When a member port is Up and the link protocol is ready, the port can forward packets. The Up state is displayed.

There are three kinds of LACP member port state:

- When the link of a port is Down, the port cannot forward packets. The Down state is displayed.
- When the link of a port is Up and the port is added to an aggregation group, the bndl state is displayed.
- When the link of a port is Up but the port is suspended because the peer end is not enabled with LACP or the attributes of the ports are inconsistent with those of the master port, the susp state is displayed. (The port in susp state does not forward packets.)

Only full-duplex ports are capable of LACP aggregation.

LACP aggregation can be implemented only when the rates, flow control approaches, medium types, and Layer-2/3 attributes of member ports are consistent.

If you modify the preceding attributes of a member port in the aggregation group, LACP aggregation will fail.

The ports which are prohibited from joining or exiting an AP port cannot be added to or removed from a static AP port or an LACP AP port.

LACP System ID

One device can be configured with only one LACP aggregation system. The system is identified by a system ID and each system has a priority, which is a configurable value. The system ID consists of the LACP system priority and MAC address of the device. A lower system priority indicates a higher priority of the system ID. If the system priorities are the same, a smaller MAC address of the device indicates a higher priority of the system ID. The system with an ID of a higher priority determines the port state. The port state of a system with an ID of a lower priority keeps consistent with that of a higher priority.

LACP Port ID

Each port has an independent LACP port priority, which is a configurable value. The port ID consists of the LACP port priority and port number. A smaller port priority indicates a higher priority of the port ID. If the port priorities are the same, a smaller port number indicates a higher priority of the port ID.

LACP Master Port

When dynamic member ports are Up, LACP selects one of those ports to be the master port based on the rates and duplex modes, ID priorities of the ports in the aggregation group, and the bundling state of the member ports in the Up state. Only the ports that have the same attributes as the master port are in Bundle state and participate in data forwarding. When the attributes of ports are changed, LACP reselects a master port. When the new master port is not in Bundle state, LACP disaggregates the member ports and performs aggregation again.

Overview

Overview	Description
Link Aggregation	Aggregates physical links statically or dynamically to realize bandwidth extension and link backup.
Load Balancing	Balances the load within an aggregation group flexibly by using different load balancing methods.

3.3.1 Link Aggregation

Working Principle

There are two kinds of AP link aggregation. One is static AP, and the other is dynamic aggregation through LACP.

- Static AP

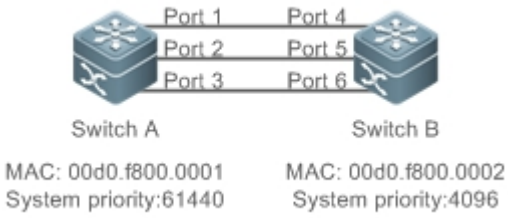
The static AP configuration is simple. Run a command to add the specified physical port to the AP port. After joining the aggregation group, a member port can receive and transmit data and participate in load balancing within the group.

- Dynamic AP (LACP)

An LACP-enabled port sends LACPDUs to advertise its system priority, system MAC address, port priority, port number, and operation key. When receiving the LACPDU from the peer end, the device compares the system priorities of both ends based on the system ID in the packet. The end with a higher system ID priority sets the ports in the aggregation group to Bundle state based on the port ID priorities in a descending order, and sends an updated LACPDU. When receiving the LACPDU, the peer end sets corresponding ports to Bundle state so that both ends maintain consistency when a port exits or joins the aggregation group. The physical link can forward packets only after the ports at both ends are bundled dynamically.

After link aggregation, the LACP member ports periodically exchange LACPDUs. When a port does not receive an LACPDU in the specified time, a timeout occurs and the links are unbundled. In this case, the member ports cannot forward packets. There are two timeout modes: long timeout and short timeout. In long timeout mode, a port sends a packet every 30s. If it does not receive a packet from the peer end in 90s, a timeout occurs. In short timeout mode, a port sends a packet every 1s. If it does not receive a packet from the peer end in 3s, a timeout occurs.

Figure 3-2 LACP Negotiation



In Figure 3-2, Switch A is connected to Switch B through three ports. Set the system priorities of Switch A and Switch B to 61440 and 4096 respectively. Enable LACP on the Ports 1–6, set the aggregation mode to the active mode, and set the port priority to the default value 32768.

When receiving an LACPDU from Switch A, Switch B finds that it has a higher system ID priority than Switch A (the system priority of Switch B is higher than that of Switch A). Switch B sets Port 4, Port 5, and Port 6 to Bundle state based on the order of port ID priorities (or in an ascending order of port numbers if the port priorities are the same). When receiving an updated LACPDU from Switch B, Switch A finds that Switch B has a higher system ID priority and has set Port 4, Port 5, and Port 6 to Bundle state. Then Switch A also sets Port 1, Port 2, and Port 3 to Bundle state.

3.3.2 Load Balancing

Working Principle

AP ports segregate packet flows by using load balancing algorithms based on packet features, such as the source and destination MAC addresses, source and destination IP addresses, and Layer-4 source and destination port numbers. The packet flow with the consistent feature is transmitted by one member link, and different packet flows are evenly distributed to member links. For example, in source MAC address-based load balancing, packets are distributed to the member links based on the source MAC addresses of the packets. Packets with different source MAC addresses are evenly distributed to member links. Packets with the identical source MAC address are forwarded by one member link.

Currently, there are several AP load balancing modes as follows:

- Source MAC address or destination MAC address
- Source MAC address + destination MAC address
- Source IP address or destination IP address
- Source IP address + destination IP address

Load balancing based on IP addresses or port numbers is applicable only to Layer-3 packets. When a device enabled with this load balancing method receives Layer-2 packets, it automatically switches to the default load balancing method.

All the load balancing methods use a load algorithm to calculate the member links based on the input parameters of the methods. The input parameters include the source MAC address, destination MAC address, source MAC address + destination MAC address, source IP address, destination IP address, source IP address + destination IP addresses, source IP address + destination IP address + Layer-4 port number and so on. The algorithm ensures that packets with different input parameters are evenly distributed to member links. It does not indicate that these packets are always distributed to different member links. For example, in IP address-based load balancing, two packets with different source and destination IP addresses may be distributed to the same member link through calculation.

Different products may support different load balancing algorithms.

3.4 Configuration

Configuration	Description and Command	
Configuring Static AP Ports	(Mandatory) It is used to configure link aggregation manually.	
	interface aggregateport	Creates an Ethernet AP port.
	port-group	Configures static AP member ports.
Configuring LACP AP Ports	(Mandatory) It is used to configure link aggregation dynamically.	
	port-group mode	Configures LACP member ports.
	lacp system-priority	Configures the LACP system priority.
	lacp port-priority	Configures the port priority.
	lacp short-timeout	Configures the short timeout mode on a port.
Enabling LinkTrap	(Optional) It is used to enable LinkTrap.	
	snmp trap link-status	Enables LinkTrap advertisement for an AP port.
	aggregateport member linktrap	Enables LinkTrap t for AP member ports.
Configuring a Load Balancing Mode	(Optional) It is used to configure a load balancing mode for an aggregated link.	
	aggregateport load-balance	Configures a load balancing algorithm for an AP port or AP member ports.

3.4.1 Configuring Static AP Ports

Configuration Effect

- Configure multiple physical ports as AP member ports to realize link aggregation.
- The bandwidth of the aggregation link is equal to the sum of the member link bandwidths.
- When a member link of the AP port is disconnected, the load carried by the link is automatically allocated to other functional

member links.

Notes

- Only physical ports can be added to an AP port.
- The ports of different media types or port modes cannot be added to the same AP port.
- Layer-2 ports can be added to only a Layer-2 AP port, and Layer-3 ports can be added to only a Layer-3 AP port. The Layer-2/3 attributes of an AP port that contains member ports cannot be modified.
- After a port is added to an AP port, the attributes of the port are replaced by those of the AP port.
- After a port is removed from an AP port, the attributes of the port are restored.

After a port is added to an AP port, the attributes of the port are consistent with those of the AP port. Therefore, do not perform configuration on the AP member ports or apply configuration to a specific AP member port. However, some configurations (the **shutdown** and **no shutdown** commands) can be configured on AP member ports. When you use AP member ports, check whether the function that you want to configure can take effect on a specific AP member port, and perform this configuration properly.

Configuration Steps

Creating an Ethernet AP Port

- Mandatory.
- Perform this configuration on an AP-enabled device.

Command	**interface aggregateport** *ap-number*
Parameter Description	*ap-number*: Indicates the number of an AP port.
Defaults	By default, no AP port is created.
Command Mode	Global configuration mode
Usage Guide	To create an Ethernet AP port, run **interfaces aggregateport** in global configuration mode. To delete the specified Ethernet AP port, run **no interfaces aggregateport** *ap-number* in global configuration mode.

Run **port-group** to add a physical port to a static AP port in interface configuration mode. If the AP port does not exist, it will be created automatically.

Run port-group mode to add a physical port to an LACP AP port in interface configuration mode. If the AP port does not exist, it will be created automatically.

The AP feature must be configured on the devices at both ends of a link and the AP mode must be the same (static AP or LACP AP).

Configuring Static AP Member Ports

- Mandatory.
- Perform this configuration on AP-enabled devices.

Command	**port-group** *ap-number*
Parameter Description	**port-group** *ap-number*: Indicates the number of an AP port.

Defaults	By default, no ports are added to any static AP port.
Command Mode	Interface configuration mode of the specified Ethernet port
Usage Guide	To add member ports to an AP port, run **port-group** in interface configuration mode. To remove member ports from an AP port, run **no port-group** in interface configuration mode.

- The static AP member ports configured on the devices at both ends of a link must be consistent.
- After a member port exits the AP port, the default settings of the member port are restored. Different functions deal with the default settings of the member ports differently. It is recommended that you check and confirm the port settings after a member port exits an AP port.
- After a member port exits an AP port, the port is disabled by using the **shutdown** command to avoid loops. After you confirm that the topology is normal, run **no shutdown** in interface configuration mode to enable the port again.

Converting Layer-2 APs to Layer-3 APs

- Optional.
- When you need to enable Layer-3 routing on an AP port, for example, to configure IP addresses or static route entries, convert the Layer-2 AP port to a Layer-3 AP port and enable routing on the Layer-3 AP port.
- Perform this configuration on AP-enabled devices that support Layer-2 and Layer-3 features, such as Layer-3 switches.

Command	**no switchport**
Parameter Description	N/A
Defaults	By default, the AP ports are Layer-2 AP ports.
Command Mode	Interface configuration mode of the specified AP port
Usage Guide	The Layer-3 AP feature is supported by only Layer-3 devices.

- The AP port created on a Layer-3 device that does not support Layer-2 feature is a Layer-3 AP port. Otherwise, the AP port is a Layer-2 AP port.

Verification

- Run **show running** to display the configuration.
- Run **show aggregateport summary** to display the AP configuration.

Command	**show aggregateport** *aggregate-port-number* [**load-balance | summary**]
Parameter Description	*aggregate-port-number*: Indicates the number of an AP port. **load-balance**: Displays the load balancing algorithm. **summary**: Displays the summary of each link.
Command Mode	Any mode
Usage Guide	The information on all AP ports is displayed if you do not specify the AP port number.
	FS# show aggregateport 1 summary AggregatePort MaxPorts SwitchPort Mode Load balance Ports ---------- ---------- ---------- ------ ---------- ---------- **Ag1** 8 Enabled ACCESS **dst-mac** **Gi0/2**

Configuration Example

Configuring an Ethernet Static AP Port

Scenario **Figure 3-3**	
Configuration Steps	● Add the GigabitEthernet 1/1 and GigabitEthernet 1/2 ports on Switch A to static AP port 3. ● Add the GigabitEthernet 2/1 and GigabitEthernet 2/2 ports on Switch B to static AP port 3.
Switch A	SwitchA# configure terminal SwitchA(config)# interface range GigabitEthernet 1/1-2 SwitchA(config-if-range)# port-group 3
Switch B	SwitchB# configure terminal SwitchB(config)# interface range GigabitEthernet 2/1-2 SwitchB(config-if-range)# port-group 3
Verification	● Run **show aggregateport summary** to check whether AP port 3 contains member ports GigabitEthernet 1/1 and GigabitEthernet 1/2.
Switch A	SwitchA# show aggregateport summary AggregatePort MaxPorts SwitchPort Mode Ports ---------- -------- ---------- ------ ---------- **Ag3** 8 Enabled ACCESS **Gi1/1,Gi1/2**
Switch B	SwitchB# show aggregateport summary AggregatePort MaxPorts SwitchPort Mode Ports ---------- -------- ---------- ------ ---------- **Ag3** 8 Enabled ACCESS **Gi2/1,Gi2/2**

3.4.2 Configuring LACP AP Ports

Configuration Effect

- Connected devices perform autonegotiation through LACP to realize dynamic link aggregation.
- The bandwidth of the aggregation link is equal to the sum of the member link bandwidths.
- When a member link of the AP port is disconnected, the load carried by the link is automatically allocated to other functional member links.
- It takes LACP 90s to detect a link failure in long timeout mode and 3s in short timeout mode.

Notes

- After a port exits an LACP AP port, the default settings of the port may be restored. Different functions deal with the default settings of the member ports differently. It is recommended that you check and confirm the port settings after a member port exits an LACP AP port.
- Changing the LACP system priority may cause LACP member ports to be disaggregated and aggregated again.
- Changing the priority of an LACP member port may cause the other member ports to be disaggregated and aggregated again.

Configuration Steps

Configuring LACP Member Ports

- Mandatory.
- Perform this configuration on LACP-enabled devices.

Command	**port-group** *key-number* **mode** { **active** | **passive** }
Parameter Description	*Key-number*: Indicates the management key of an AP port. In other words, it is the LACP AP port number. The maximum value is subject to the number of AP ports supported by the device. **active**: Indicates that ports are added to a dynamic AP port actively. **passive**: Indicates that ports are added to a dynamic AP port passively.
Defaults	By default, no physical ports are added to any LACP AP port.
Command Mode	Interface configuration mode of the specified physical port
Usage Guide	Use this command in interface configuration mode to add member ports to an LACP AP port.

The LACP member port configuration at both ends of a link must be consistent.

Configuring the LACP System Priority

- Optional.
- Perform this configuration when you need to adjust the system ID priority. A smaller value indicates a higher system ID priority. The device with a higher system ID priority selects an AP port.
- Perform this configuration on LACP-enabled devices.

Command	**lacp system-priority** *system-priority*
Parameter Description	*system-priority*: Indicates the LACP system priority. The value ranges from 0 to 65535.
Defaults	By default, the LACP system priority is 32768.
Command Mode	Global configuration mode
Usage Guide	Use this command in global configuration mode to configure the LACP system priority. All the dynamic member links share one LACP system priority. Changing the LACP system priority will affect all member links. To restore the default settings, run **no lacp system-priority** in interface configuration mode.

Configuring the Priority of an LACP Member Port

- Optional.
- Perform this configuration when you need to specify the port ID priority. A smaller value indicates a higher port ID priority. The port with the highest port ID priority will be selected as the master port.

- Perform this configuration on LACP-enabled devices.

Command	**lacp port-priority** *port-priority*
Parameter Description	*port-priority*: Indicates the priority of an LACP member port. The value ranges from 0 to 65535.
Defaults	By default, the priority of an LACP member port is 32768.
Command Mode	Interface configuration mode of the specified physical port
Usage Guide	Use this command in global configuration mode to configure the priority of an LACP member port. To restore the settings, run **no lacp port-priority** in interface configuration mode.

Configuring the Timeout Mode of LACP Member Ports

- Optional.
- When you need to implement real-time link failure detection, configure the short timeout mode. It takes LACP 90s to detect a link failure in long timeout mode and 3s in short timeout mode.
- Perform this configuration on LACP-enabled devices, such as switches.

Command	**lacp short-timeout**
Parameter Description	N/A
Defaults	By default, the timeout mode of LACP member ports is long timeout.
Command Mode	Interface configuration mode
Usage Guide	The timeout mode is supported only by physical ports. To restore the default settings, run **no lacp short-timeout** in interface configuration mode.

Verification

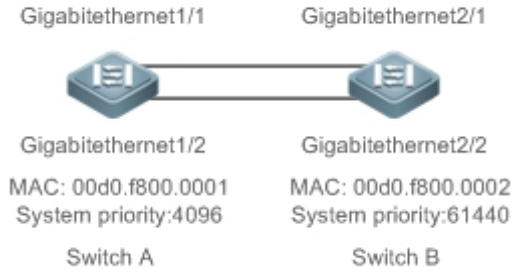
- Run **show running** to display the configuration.
- Run **show lacp summary** to display LACP link state.

Command	**show lacp summary** [*key-number*]
Parameter Description	*key-name*: Indicates the number of an LACP AP port.
Command Mode	Any mode
Usage Guide	The information on all LACP AP ports is displayed if you do not specify *key-name*.
	FS(config)# show lacp summary 3 System Id:32768, 00d0.f8fb.0002 Flags: S - Device is requesting Slow LACPDUs F - Device is requesting Fast LACPDUs. A - Device is in active mode. P - Device is in passive mode. Aggregated port 3:

	Local information: LACP port Oper Port Port Port Flags State Priority Key Number State ---------- Gi0/1 SA bndl 4096 0x3 0x1 0x3d Gi0/2 SA bndl 4096 0x3 0x2 0x3d Gi0/3 SA bndl 4096 0x3 0x3 0x3d Partner information: LACP port Oper Port Port Port Flags Priority Dev ID Key Number State ---------- Gi0/1 SA 61440 00d0.f800.0001 0x3 0x1 0x3d Gi0/2 SA 61440 00d0.f800.0001 0x3 0x2 0x3d Gi0/3 SA 61440 00d0.f800.0001 0x3 0x3 0x3d

Configuration Example

↘ Configuring LACP

Scenario **Figure 3-4**	
Configuration Steps	• On Switch A, set the LACP system priority to 4096. • Enable dynamic link aggregation on the GigabitEthernet1/1 and GigabitEthernet1/2 ports on Switch A and add the ports to LACP AP port 3. • On Switch B, set the LACP system priority to 61440. • Enable dynamic link aggregation on the GigabitEthernet2/1 and GigabitEthernet2/2 ports on Switch B and add the ports to LACP AP port 3.
Switch A	SwitchA# configure terminal SwitchA(config)# lacp system-priority 4096 SwitchA(config)# interface range GigabitEthernet 1/1-2 SwitchA(config-if-range)# port-group 3 mode active SwitchA(config-if-range)# end

Switch B	SwitchB# configure terminal SwitchB(config)# lacp system-priority 61440 SwitchB(config)# interface range GigabitEthernet 2/1-2 SwitchB(config-if-range)# port-group 3 mode active SwitchB(config-if-range)# end
Verification	● Run **show lacp summary 3** to check whether LACP AP port 3 contains member ports GigabitEthernet2/1 and GigabitEthernet2/2.
Switch A	SwitchA# show LACP summary 3 System Id:32768, 00d0.f8fb.0001 Flags: S - Device is requesting Slow LACPDUs F - Device is requesting Fast LACPDUs. A - Device is in active mode. P - Device is in passive mode. Aggregated port 3: Local information: LACP port Oper Port Port Port Flags State Priority Key Number State ---------- Gi1/1 SA bndl 32768 0x3 0x1 0x3d Gi1/2 SA bndl 32768 0x3 0x2 0x3d Partner information: LACP port Oper Port Port Port Flags Priority Dev ID Key Number State ---------- Gi2/1 SA 32768 00d0.f800.0002 0x3 0x1 0x3d Gi2/2 SA 32768 00d0.f800.0002 0x3 0x2 0x3d
Switch B	SwitchB# show LACP summary 3 System Id:32768, 00d0.f8fb.0002 Flags: S - Device is requesting Slow LACPDUs F - Device is requesting Fast LACPDUs. A - Device is in active mode. P - Device is in passive mode. Aggregated port 3: Local information: LACP port Oper Port Port

	Port Flags State Priority Key Number State ---------- Gi2/1 SA bndl 32768 0x3 0x1 0x3d Gi2/2 SA bndl 32768 0x3 0x2 0x3d Partner information: LACP port Oper Port Port Port Flags Priority Dev ID Key Number State ---------- Gi1/1 SA 32768 00d0.f800.0001 0x3 0x1 0x3d Gi1/2 SA 32768 00d0.f800.0001 0x3 0x2 0x3d

3.4.3 Enabling LinkTrap

Configuration Effect

Enable the system with LinkTrap to send LinkTrap messages when aggregation links are changed.

Configuration Steps

Enabling LinkTrap for an AP Port

- Optional.
- Enable LinkTrap in interface configuration mode. By default, LinkTrap is enabled. LinkTrap messages are sent when the link state or protocol state of the AP port is changed.
- Perform this configuration on AP-enabled devices.

Command	**snmp trap link-status**
Parameter Description	N/A
Defaults	By default, LinkTrap is enabled.
Command Mode	Interface configuration mode of the specified AP port
Usage Guide	Use this command in interface configuration mode to enable LinkTrap for the specified AP port. After LinkTrap is enabled, LinkTrap messages are sent when the link state of the AP port is changed. Otherwise, LinkTrap messages are not sent. By default, LinkTrap is enabled. To disable LinkTrap for an AP port, run **no snmp trap link-status** in interface configuration mode. LinkTrap cannot be enabled for a specific AP member port. To enable LinkTrap for all AP member ports, run **aggregateport member linktrap** in global configuration mode.

Enabling LinkTrap for AP Member Ports

- Optional.
- By default, LinkTrap is disabled for AP member ports.
- Perform this configuration on AP-enabled devices.

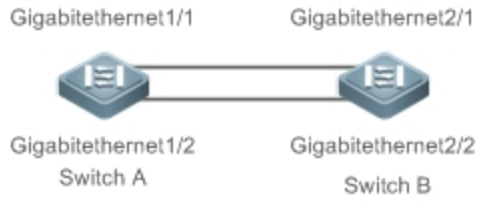
Command	**aggregateport member linktrap**
Parameter Description	N/A
Defaults	By default, LinkTrap is disabled for AP member ports.
Command Mode	Global configuration mode
Usage Guide	Use this command in global configuration mode to enable LinkTrap for all AP member ports. By default, LinkTrap messages are not sent when the link state of AP member ports is changed. To disable LinkTrap for all AP member ports, run **no aggregateport member linktrap** in global configuration mode.

Verification

- Run **show running** to display the configuration.
- After LinkTrap is enabled, you can monitor this feature on AP ports or their member ports by using the MIB software.

Configuration Example

Enabling LinkTrap for AP Member Ports

Scenario **Figure 3-5**	
Configuration Steps	• Add the GigabitEthernet 1/1 and GigabitEthernet 1/2 ports on Switch A to static AP port 3. • Add the GigabitEthernet 2/1 and GigabitEthernet 2/2 ports on Switch B to static AP port 3. • On Switch A, disable LinkTrap for AP port 3 and enable LinkTrap for its member ports. • On Switch B, disable LinkTrap for AP port 3 and enable LinkTrap its AP member ports.
Switch A	SwitchA# configure terminal SwitchA(config)# interface range GigabitEthernet 1/1-2 SwitchA(config-if-range)# port-group 3 SwitchA(config-if-range)# exit SwitchA(config)# aggregateport member linktrap SwitchA(config)# interface Aggregateport 3 SwitchA(config-if-AggregatePort 3)# no snmp trap link-status
Switch B	SwitchB# configure terminal SwitchB(config)# interface range GigabitEthernet 2/1-2 SwitchB(config-if-range)# port-group 3 SwitchB(config-if-range)# exit SwitchB(config)# aggregateport member linktrap

	SwitchB(config)# interface Aggregateport 3 SwitchB(config-if-AggregatePort 3)# no snmp trap link-status
Verification	● Run **show running** to check whether LinkTrap is enabled for AP port 3 and its member ports.
Switch A	SwitchA# show run | include AggregatePort 3 Building configuration... Current configuration: 54 bytes interface AggregatePort 3 no snmp trap link-status SwitchA# show run | include AggregatePort aggregateport member linktrap
Switch B	SwitchB# show run | include AggregatePort 3 Building configuration... Current configuration: 54 bytes interface AggregatePort 3 no snmp trap link-status SwitchB# show run | include AggregatePort aggregateport member linktrap

3.4.4 Configuring a Load Balancing Mode

Configuration Effect

The system distributes incoming packets among member links by using the specified load balancing algorithm. The packet flow with the consistent feature is transmitted by one member link, whereas different packet flows are evenly distributed to various links. A device enabled with enhanced load balancing first determines the type of packets to be transmitted and performs load balancing based on the specified fields in the packets. For example, the AP port performs source IP-based load balancing on the packets containing an ever-changing source IPv4 address.

Notes

N/A

Configuration Steps

Configuring the Global Load Balancing Algorithm of an AP port

- (Optional) Perform this configuration when you need to optimize load balancing.
- Perform this configuration on AP-enabled devices.

Command	**aggregateport load-balance** { **dst-mac** | **src-mac** | **src-dst-mac** | **dst-ip** | **src-ip** | **src-dst-ip** }
Parameter	**dst-mac**: Indicates that load is distributed based on the destination MAC addresses of incoming packets.

Description	**src-mac**: Indicates that load is distributed based on the source MAC addresses of incoming packets. **src-dst-ip**: Indicates that load is distributed based on source and destination IP addresses of incoming packets. **dst-ip**: Indicates that load is distributed based on the destination IP addresses of incoming packets. **src-ip**: Indicates that load is distributed based on the source IP addresses of incoming packets. **src-dst-mac**: Indicates that load is distributed based on source and destination MAC addresses of incoming packets.
Defaults	Load balancing can be based on source and destination MAC addresses (applicable to switches), source and destination IP addresses, or the profile of enhanced load balancing (applicable to switches with CB line cards).
Command Mode	Global configuration mode
Usage Guide	To restore the default settings, run **no aggregateport load-balance** in global configuration mode. You can run **aggregateport load-balance** in interface configuration mode of an AP port on devices that support load balancing configuration on a specific AP port. The configuration in interface configuration mode prevails. To disable the load balancing algorithm, run **no aggregateport load-balance** in interface configuration mode of the AP port. After that, the load balancing algorithm configured in global configuration mode takes effect. You can run **aggregateport load-balance** in interface configuration mode of an AP port on devices that support load balancing configuration on a specific AP port.

Verification

- Run **show running** to display the configuration.
- Run **show aggregateport load-balance** to display the load balancing configuration. If a device supports load balancing configuration on a specific AP port, run **show aggregateport summary** to display the configuration.

Command	**show aggregateport** *aggregate-port-number* [**load-balance | summary**]
Parameter Description	*aggregate-port-number*: Indicates the number of an AP port. **load-balance**: Displays the load balancing algorithm. **summary**: Displays the summary of each link.
Command Mode	Any mode
Usage Guide	The information on All AP ports is displayed if you do not specify the AP port number.
	FS# show aggregateport 1 summary AggregatePort MaxPorts SwitchPort Mode Load balance Ports ---------- ---------- ---------- ------ ---------- ---------- **Ag1** 8 Enabled ACCESS **dst-mac** **Gi0/2**

Configuration Example

Configuring a Load Balancing Mode

Scenario **Figure 3-6**	
Configuration Steps	● Add the GigabitEthernet 1/1 and GigabitEthernet 1/2 ports on Switch A to static AP port 3. ● Add the GigabitEthernet 2/1 and GigabitEthernet 2/2 ports on Switch B to static AP port 3. ● On Switch A, configure source MAC address-based load balancing for AP port 3 in global configuration mode. ● On Switch B, configure destination MAC address-based load balancing for AP port 3 in global configuration mode.
Switch A	SwitchA# configure terminal SwitchA(config)# interface range GigabitEthernet 1/1-2 SwitchA(config-if-range)# port-group 3 SwitchA(config-if-range)# exit SwitchA(config)# aggregateport load-balance src-mac
Switch B	SwitchB# configure terminal SwitchB(config)# interface range GigabitEthernet 2/1-2 SwitchB(config-if-range)# port-group 3 SwitchB(config-if-range)# exit SwitchB(config)# aggregateport load-balance dst-mac
Verification	● Run **show aggregateport load-balance** to check the load balancing algorithm configuration.
Switch A	SwitchA# show aggregatePort load-balance Load-balance : **Source MAC**
Switch B	SwitchB# show aggregatePort load-balance Load-balance : **Destination MAC**

Common Errors

N/A

3.5 Monitoring

Displaying

Description	Command
Displays the configuration of an enhanced load balancing profile.	**show load-balance-profile** [*profile-name*]

Description	Command
Displays the LACP aggregation state. You can display the information on a specified LACP AP port by specifying *key-number*.	**show lacp summary** [*key-numebr*]
Displays the summary or load balancing algorithm of an AP port.	**show aggregateport** [*ap-number*] { **load-balance** | **summary** }

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs an AP port.	**debug lsm ap**
Debugs LACP.	**debug lacp** { **packet** | **event** | **database** | **ha** | **realtime** | **stm** | **timer** | **all**}

4 Configuring VLAN

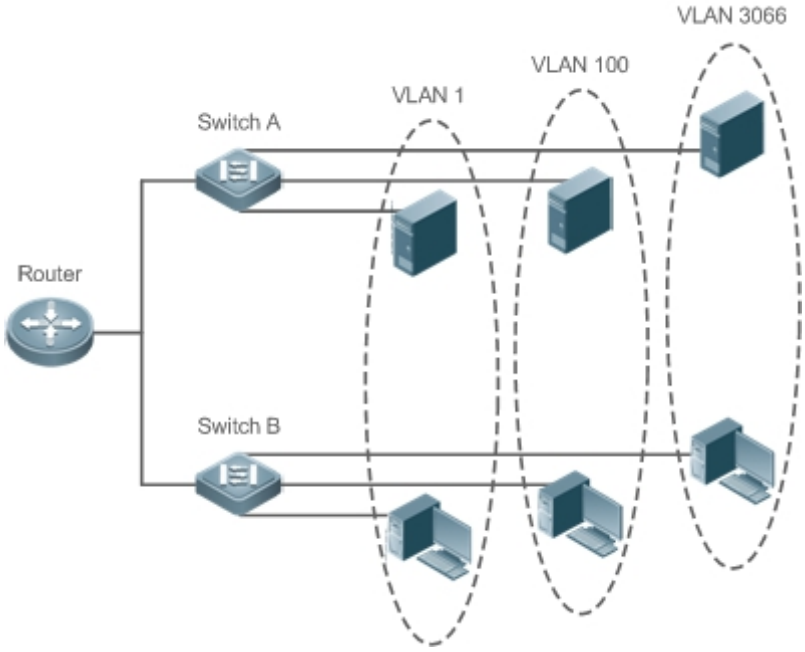
4.1 Overview

A Virtual Local Area Network (VLAN) is a logical network created based on a physical network. A VLAN can be categorized into Layer-2 networks of the OSI model.

A VLAN has the same properties as a common LAN, except for physical location limitation. Unicast, broadcast and multicast frames of Layer 2 are forwarded and transmitted within a VLAN, keeping traffic segregated.

We may define a port as a member of a VLAN, and all terminals connected to this port are parts of a virtual network that supports multiple VLANs. You do not need to adjust the network physically when adding, removing and modifying users. Communication among VLANs is realized through Layer-3 devices, as shown in the following figure.

Figure 4- 1



Protocols and Standards

- IEEE 802.1Q

4.2 Applications

Application	Description
Isolating VLANs at Layer 2 and Interconnecting VLANs at Layer 3	An intranet is divided into multiple VLANs, realizing Layer-2 isolation and Layer-3 interconnection with each other through IP forwarding by core switches.

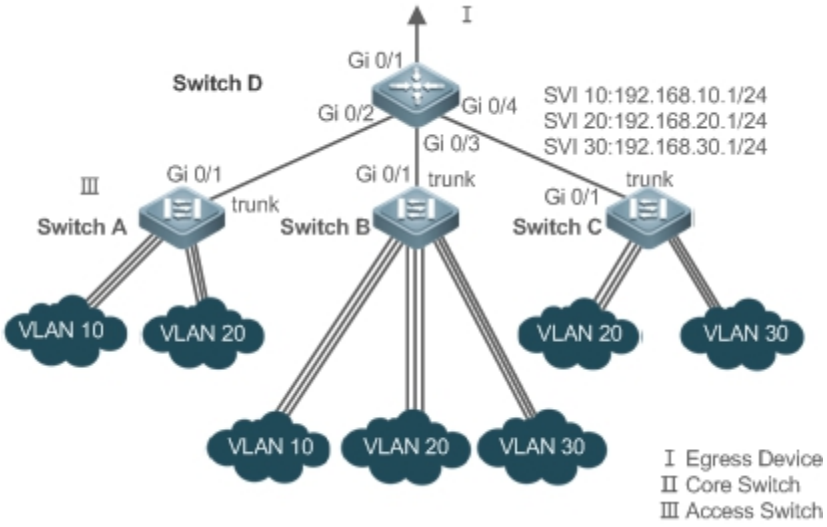
4.2.1 Isolating VLANs at Layer 2 and Interconnecting VLANs at Layer 3

Scenario

An intranet is divided into VLAN 10, VLAN 20 and VLAN 30, realizing Layer-2 isolation from each other. The three VLANs correspond respectively to the IP sub-networks 192.168.10.0/24, 192.168.20.0/24, and 192.168.30.0/24, realizing interconnection with each other

through IP forwarding by Layer-3 core switches.

Figure 4- 2



Remarks:	Switch A, Switch B and Switch C are access switches. Configure three VLANs on a core switch and the port connected to the access switches as a Trunk port, and specify a list of allowed-VLANs to realize Layer-2 isolation; Configure three SVIs on the core switch, which are the gateway interfaces of the IP sub-networks corresponding to the three VLANs, and configure the IP addresses for these interfaces. Create VLANs respectively on the three access switches, assign Access ports for the VLANs, and specify Trunk ports of the core switch.

Deployment

- Divide an intranet into multiple VLANs to realize Layer-2 isolation among them.
- Configure SVIs on a Layer-3 switch to realize Layer-3 communication among VLANs.

4.3 Features

Basic Concepts

VLAN

A VLAN is a logical network created based on a physical network. A VLAN has the same properties as a common LAN, except for physical location limitation. Unicast, broadcast and multicast frames of Layer 2 are forwarded and transmitted within a VLAN, keeping traffic segregated.

- The VLANs supported by FS products comply with the IEEE802.1Q standard. A maximum of 4094 VLANs (VLAN ID 1-4094) are supported, among which VLAN 1 cannot be deleted.
- The configurable VLAN IDs are from 1 to 4094.
- In case of insufficient hardware resources, the system returns information on VLAN creation failure.

Port Mode

You can determine the frames allowed to pass a port and the VLANs which the port belongs to by configuring the port mode. See the

following table for details.

Port Mode	Description
Access port	An Access port belongs to only one VLAN, which is specified manually.
Trunk port (802.1Q)	A Trunk port belongs to all the VLANs of an access switch by default, and it can forward the frames of all the VLANs or the frames of allowed-VLANs.
Uplink port	An Uplink port belongs to all the VLANs of an access switch by default, and it can forward the frames of all the VLANs and tag the native VLAN egress traffic.
Hybrid port	A Hybrid port belongs to all the VLANs of an access switch by default, and it can forward the frames of all the VLANs and send frames of VLANs untagged. It can also transmit frames of allowed-VLANs.

Overview

Feature	Description
VLAN	VLAN helps realize Layer-2 isolation.

4.3.1 VLAN

Every VLAN has an independent broadcast domain, and different VLANs are isolated on Layer 2.

Working Principle

Every VLAN has an independent broadcast domain, and different VLANs are isolated on Layer 2.

Layer-2 isolation: If no SVIs are configured for VLANs, VLANs are isolated on Layer 2. This means users in these VLANs cannot communicate with each other.

Layer-3 interconnection: If SVIs are configured on a Layer-3 switch for VLANs, these VLANs can communicate with each other on Layer 3.

4.4 Configuration

Configuration	Description and Command	
Configuring Basic VLAN	(Mandatory) It is used to create a VLAN.	
	vlan	Enters a VLAN ID.
	(Optional) It is used to configure an Access port to transmit the flows from a single VLAN.	
	switchport mode access	Defines a port as a Layer-2 Access port.
	switchport access vlan	Assigns a port to a VLAN.
	add interface	Adds one Access port or a group of such ports to the current VLAN.
	(Optional) It is used to rename a VLAN.	
	name	Names a VLAN.
Configuring a Trunk Port	(Mandatory) It is used to configure the port as a Trunk port.	
	switchport mode trunk	Defines a port as a Layer-2 Trunk port.

	(Optional) It is used to configure Trunk ports to transmit flows from multiple VLANs.	
	switchport trunk allowed vlan	Configures allowed-VLANs for a Trunk port.
	switchport trunk native vlan	Specifies a native VLAN for a Trunk port.
Configuring an Uplink Port	(Mandatory) It is used to configure the port as an Uplink port.	
	switchport mode uplink	Configures a port as an Uplink port.
	(Optional) It is used to restore the port mode.	
	no switchport mode	Restores the port mode.
Configuring a Hybrid Port	(Mandatory) It is used to configure a port as a Hybrid port.	
	switchport mode hybrid	Configures a port as a Hybrid port.
	(Optional) It is used to transmit the frames of multiple VLANs untagged.	
	no switchport mode	Restores the port mode.
	switchport hybrid allowed vlan	Configures allowed-VLANs for a Hybrid port.
	switchport hybrid native vlan	Configures a default VLAN for a Hybrid port.

4.4.1 Configuring Basic VLAN

Configuration Effect

- A VLAN is identified by a VLAN ID. You may add, delete, modify VLANs 2 to 4094, but VLAN 1 is created automatically and cannot be deleted. You may configure the port mode, and add or remove a VLAN.

Notes

- N/A

Configuration Steps

Creating and Modifying a VLAN

- Mandatory.
- In case of insufficient hardware resources, the system returns information on VLAN creation failure.
- Use the vlan vlan-id command to create a VLAN or enter VLAN mode.
- Configuration:

Command	**vlan** *vlan-id*
Parameter Description	*vlan-id*: indicates VLAN ID ranging from 1 to 4094.
Defaults	VLAN 1 is created automatically and is not deletable.
Command Mode	Global configuration mode
Usage Guide	If you enter a new VLAN ID, the corresponding VLAN will be created. If you enter an existing VLAN ID, the corresponding VLAN will be modified. You may use the **no vlan** *vlan-id* command to delete a VLAN. The undeletable VLANs include VLAN1, the VLANs configured with SVIs, and SubVLANs.

Renaming a VLAN

- Optional.
- You cannot rename a VLAN the same as the default name of another VLAN.
- Configuration:

Command	**name** *vlan-name*
Parameter Description	*vlan-name*: indicates a VLAN name.
Defaults	By default, the name of a VLAN is its VLAN ID. For example, the default name of the VLAN 4 is VLAN 0004.
Command Mode	VLAN configuration mode
Usage Guide	To restore the VLAN name to defaults, use the **no name** command.

Assigning Current Access port to a Specified VLAN

- Optional.
- Use the **switchport mode access** command to specify Layer-2 ports (switch ports) as Access ports.
- Use the **switchport access vlan** *vlan-id* command to add an Access port to a specific VLAN so that the flows from the VLAN can be transmitted through the port.
- Configuration:

Command	**switchport mode access**
Parameter Description	N/A
Defaults	A switch port is an Access port by default.
Command Mode	Interface configuration mode
Usage Guide	N/A

Command	**switchport access vlan** *vlan-id*
Parameter Description	*vlan-id*: indicates a VLAN ID.
Defaults	An Access port is added to VLAN 1 by default.
Command Mode	Interface configuration mode
Usage Guide	If a port is assigned to a non-existent VLAN, the VLAN will be created automatically.

Adding an Access Port to Current VLAN

- Optional.
- This command takes effect only on an Access port. After an Access port is added to a VLAN, the flows of the VLAN can be transmitted through the port.
- Configuration:

Command	**add interface** { *interface-id* | **range** *interface-range* }
Parameter Description	*interface-id*: indicates a single port. *interface-id*: indicates multiple ports.
Defaults	By default, all Layer-2 Ethernet ports belong to VLAN 1.
Command Mode	VLAN configuration mode
Usage Guide	In VLAN configuration mode, add a specific Access port to a VLAN. This command takes the same effect as command **switchport access vlan** *vlan-id*.

For the two commands of adding a port to a VLAN, the command configured later will overwrite the other one.

Verification

- Send untagged packets to an Access port, and they are broadcast within the VLAN.
- Use commands **show vlan** and **show interface switchport** to check whether the configuration takes effect.

Command	**show vlan** [**id** *vlan-id*]
Parameter Description	*vlan-id* **:** indicates a VLAN ID.
Command Mode	Any mode
Usage Guide	N/A
Command Display	FS(config-vlan)#show vlan id 20 VLAN Name Status Ports ---- --- --- --- 20 VLAN0020 STATIC Gi0/1

Configuration Example

Configuring Basic VLAN and Access Port

Configuration Steps	• Create a VLAN and rename it. • Add an Access port to the VLAN. There are two approaches. One is:
	FS# configure terminal FS(config)# vlan 888 FS(config-vlan)# name test888 FS(config-vlan)# exit FS(config)# interface GigabitEthernet 0/3 FS(config-if-GigabitEthernet 0/3)# switchport mode access FS(config-if-GigabitEthernet 0/3)# switchport access vlan 20

	The other approach is adding an Access port (GigabitEthernet 0/3) to VLAN20: FS# configure terminal SwitchA(config)#vlan 20 SwitchA(config-vlan)#add interface GigabitEthernet 0/3
Verification	Check whether the configuration is correct.
	FS(config-vlan)#show vlan VLAN Name Status Ports ---- ---------- --------- ---------- 1 VLAN0001 STATIC **20 VLAN0020 STATIC Gi0/3** **888 test888 STATIC** FS(config-vlan)# FS# show interface GigabitEthernet 0/3 switchport Interface Switchport Mode Access Native Protected VLAN lists ---------- ---------- --------- ------ ------ --------- ---------- GigabitEthernet 0/3 **enabled ACCESS 20 1 Disabled ALL** FS# show run !

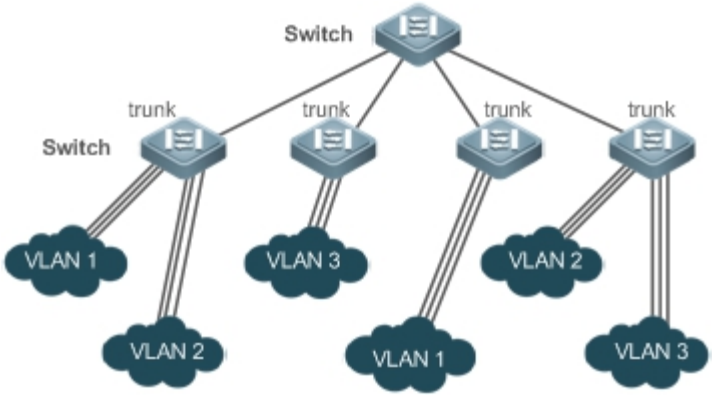
4.4.2 Configuring a Trunk Port

Configuration Effect

A Trunk is a point-to-point link connecting one Ethernet interface or multiple ones to other network devices (for example, a router or switch) and it may transmit the flows from multiple VLANs.

The Trunk of Ruije devices adopts the 802.1Q encapsulation standard. The following figure displays a network adopting a Trunk connection.

Figure 4- 3



You may configure an Ethernet port or Aggregate Port (See *Configuring Aggregate Port* for details) as a Trunk port.

You should specify a native VLAN for a Trunk port. The untagged packets received by and sent from the Trunk port are considered to belong to the native VLAN. The default VLAN ID (PVID in the IEEE 802.1Q) of this Trunk port is the native VLAN ID. Meanwhile, frames of

the native VLAN sent via the Trunk are untagged. The default native VLAN of a Trunk port is VLAN 1.

When configuring a Trunk link, make sure the Trunk ports at the two ends of the link adopt the same native VLAN.

Configuration Steps

Configuring a Trunk Port

- Mandatory.
- Configure a Trunk port to transmit the flows from multiple VLANs.
- Configuration:

Command	**switchport mode trunk**
Parameter Description	N/A
Defaults	The default mode is Access, which can be modified to Trunk.
Command Mode	Interface configuration mode
Usage Guide	To restore all properties of a Trunk port to defaults, use the **no switchport mode** command.

Defining Allowed-VLANs for a Trunk Port

- Optional.
- By default, a trunk port transmits the flows from all the VLANs (1 to 4094). You may configure a list of allowed-VLANs to prohibit flows of some VLANs from passing through a Trunk port.

- Configuration:

Command	**switchport trunk allowed vlan** { **all** | [**add** | **remove** | **except** | **only**] } *vlan-list*
Parameter Description	The parameter vlan-list can be a VLAN or some VLANs, and the VLAN IDs are connected by "-" in order. For example: 10–20. **all** indicates allowed-VLANs include all VLANs; **add** indicates adding a specific VLAN to the list of allowed-VLANs; **remove** indicates removing a specific VLAN from the list of allowed-VLANs; **except** indicates adding all VLANs except those in the listed VLAN to the list of allowed-VLANs. **only** indicates adding the listed VLANs to the list of allowed-VLANs, and removing the other VLANs from the list.
Defaults	The Trunk port and the Uplink port belong to all VLANs.
Command Mode	Interface configuration mode
Usage Guide	To restore the configuration on a Trunk port to defaults (**all**), use the **no switchport trunk allowed vlan** command.

Configuring a Native VLAN

- Optional.
- A Trunk port receives and sends tagged or untagged 802.1Q frames. Untagged frames transmit the flows from the native VLAN. The default native VLAN is VLAN 1.
- If a frame carries the VLAN ID of a native VLAN, its tag will be stripped automatically when it passes a Trunk port.
- Configuration:

Command	**switchport trunk native vlan** *vlan-id*
Parameter Description	*vlan-id*: indicates a VLAN ID.
Defaults	The default VALN for a Trunk/Uplink port is VLAN 1.
Command Mode	Interface configuration mode
Usage Guide	To restore the native VLAN of a Trunk port back to defaults, use the **no switchport trunk native vlan** command.

When you set the native VLAN of a port to a non-existent VLAN, this VLAN will not be created automatically. Besides, the native VLAN can be out of the list of allowed-VLANs for this port. In this case, the flows from the native VLAN cannot pass through the port.

Verification

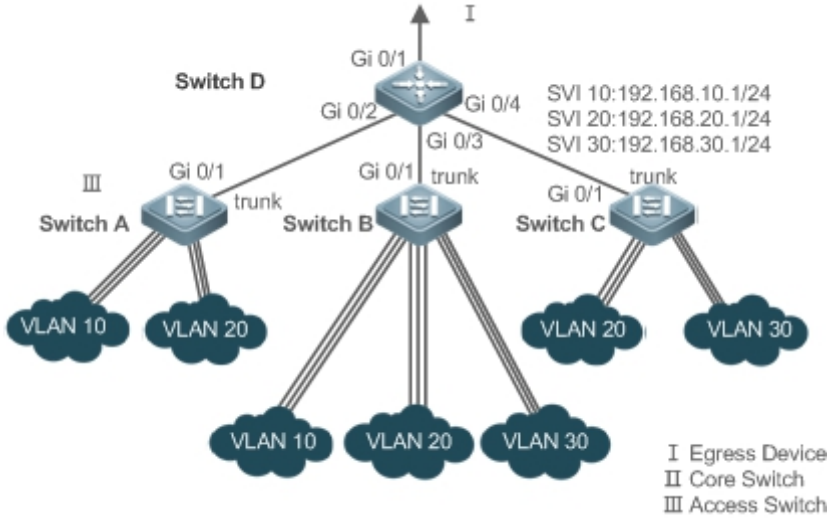
- Send tag packets to a Trunk port, and they are broadcast within the specified VLANs.

- Use commands **show vlan** and **show interface switchport** to check whether the configuration takes effect.

Command	**show vlan** [**id** *vlan-id*]
Parameter Description	*vlan-id* **:** indicates a VLAN ID.
Command Mode	Any mode
Usage Guide	N/A
Command Display	FS(config-vlan)#show vlan id 20 VLAN Name Status Ports ---- ---------- --------- ---------- **20 VLAN0020 STATIC Gi0/1**

Configuration Example

Configuring Basic VLAN to Realize Layer-2 Isolation and Layer-3 Interconnection

Scenario **Figure 4- 4**	
Configuration Steps	Networking Requirements: As shown in the figure above, an intranet is divided into VLAN 10, VLAN 20 and VLAN 30, realizing Layer-2 isolation from each other. The three VLANs correspond respectively to the IP sub-networks 192.168.10.0/24, 192.168.20.0/24, and 192.168.30.0/24, realizing interconnection with each other through IP forwarding by Layer-3 core switches. Key Points: The following example describes the configuration steps on a core switch and an access switch. ● Configure three VLANs on a core switch and the port connected to the access switches as a Trunk port, and specify a list of allowed-VLANs to realize Layer-2 isolation. ● Configure three SVIs on the core switch, which are the gateway interfaces of the IP sub-networks corresponding to the three VLANs, and configure the IP addresses for these interfaces. ● Create VLANs respectively on the three access switches, assign Access ports for the VLANs, and specify Trunk ports of the core switch. The following example describes the configuration steps on Switch A.
D	D#configure terminal D(config)#vlan 10 D(config-vlan)#vlan 20 D(config-vlan)#vlan 30

	D(config-vlan)#exit D(config)#interface range GigabitEthernet 0/2-4 D(config-if-range)#switchport mode trunk D(config-if-range)#exit D(config)#interface GigabitEthernet 0/2 D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan remove 1-4094 D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan add 10,20 D(config-if-GigabitEthernet 0/2)#interface GigabitEthernet 0/3 D(config-if-GigabitEthernet 0/3)#switchport trunk allowed vlan remove 1-4094 D(config-if-GigabitEthernet 0/3)#switchport trunk allowed vlan add 10,20,30 D(config-if-GigabitEthernet 0/3)#interface GigabitEthernet 0/4 D(config-if-GigabitEthernet 0/4)#switchport trunk allowed vlan remove 1-4094 D(config-if-GigabitEthernet 0/4)#switchport trunk allowed vlan add 20,30 D#configure terminal D(config)#interface vlan 10 D(config-if-VLAN 10)#ip address 192.168.10.1 255.255.255.0 D(config-if-VLAN 10)#interface vlan 20 D(config-if-VLAN 20)#ip address 192.168.20.1 255.255.255.0 D(config-if-VLAN 20)#interface vlan 30 D(config-if-VLAN 30)#ip address 192.168.30.1 255.255.255.0 D(config-if-VLAN 30)#exit
A	A#configure terminal A(config)#vlan 10 A(config-vlan)#vlan 20 A(config-vlan)#exit A(config)#interface range GigabitEthernet 0/2-12 A(config-if-range)#switchport mode access A(config-if-range)#switchport access vlan 10 A(config-if-range)#interface range GigabitEthernet 0/13-24 A(config-if-range)#switchport mode access A(config-if-range)#switchport access vlan 20 A(config-if-range)#exit A(config)#interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#switchport mode trunk
Verification	Display the VLAN configuration on the core switch. ● Display VLAN information including VLAN IDs, VLAN names, status and involved ports. ● Display the status of ports Gi 0/2, Gi 0/3 and Gi 0/4.
D	D#show vlan VLAN Name Status Ports ---- -------- -------- ---------- **1 VLAN0001 STATIC Gi0/1, Gi0/5, Gi0/6, Gi0/7** **Gi0/8, Gi0/9, Gi0/10, Gi0/11** **Gi0/12, Gi0/13, Gi0/14, Gi0/15**

```
                                Gi0/16, Gi0/17, Gi0/18, Gi0/19
                                Gi0/20, Gi0/21, Gi0/22, Gi0/23
                                Gi0/24
  10 VLAN0010   STATIC          Gi0/2, Gi0/3
  20 VLAN0020   STATIC          Gi0/2, Gi0/3, Gi0/4
  30 VLAN0030   STATIC          Gi0/3, Gi0/4
D#show interface GigabitEthernet 0/2 switchport
Interface                       Switchport Mode       Access Native Protected VLAN lists
-------------------------------- ---------- --------- ------ ------ --------- ---------------
GigabitEthernet 0/2             enabled    TRUNK     1      1      Disabled  10,20
D#show interface GigabitEthernet 0/3 switchport
Interface                       Switchport Mode       Access Native Protected VLAN lists
-------------------------------- ---------- --------- ------ ------ --------- ---------------
GigabitEthernet 0/3             enabled    TRUNK     1      1      Disabled  10,20,30
D#show interface GigabitEthernet 0/4 switchport
Interface                       Switchport Mode       Access Native Protected VLAN lists
-------------------------------- ---------- --------- ------ ------ --------- ---------------
GigabitEthernet 0/4             enabled    TRUNK     1      1      Disabled  20,30
```

Common Errors

- N/A

4.4.3 Configuring an Uplink Port

Configuration Effect

- An Uplink port is usually used in QinQ (the IEEE 802.1ad standard) environment, and is similar to a Trunk port. Their difference is that an Uplink port only transmits tagged frames while a Trunk port sends untagged frames of the native VLAN.

Configuration Steps

Configuring an Uplink Port

- Mandatory.
- Configure an Uplink port to transmit the flows from multiple VLANS, but only tagged frames can be transmitted.
- Configuration:

Command	**switchport mode uplink**
Parameter Description	N/A
Defaults	The default mode is Access, which can be modified to Uplink.
Command Mode	Interface configuration mode
Usage Guide	To restore all properties of an Uplink port to defaults, use the **no switchport mode** command.

Defining Allowed-VLANs for a Trunk Port

- Optional.
- You may configure a list of allowed-VLANs to prohibit flows of some VLANs from passing through an Uplink port.
- Configuration:

Command	**switchport trunk allowed vlan** { **all** | [**add** | **remove** | **except** | **only**] } *vlan-list*
Parameter Description	The parameter *vlan-list* can be a VLAN or some VLANs, and the VLAN IDs are connected by "-" in order. For example: 10–20. **all** indicates allowed-VLANs include all VLANs; **add** indicates adding a specific VLAN to the list of allowed-VLANs; **remove** indicates removing a specific VLAN from the list of allowed-VLANs; **except** indicates adding all VLANs except those in the listed VLAN to the list of allowed-VLANs; and **only** indicates adding the listed VLANs to the list of allowed-VLANs, and removing the other VLANs from the list.
Command Mode	Interface configuration mode
Usage Guide	To restore the allowed-VLANs to defaults (**all**), use the **no switchport trunk allowed vlan** command.

Configuring a Native VLAN

- Optional.
- If a frame carries the VLAN ID of a native VLAN, its tag will not be stripped when it passes an Uplink port. This is contrary to a Trunk port.
- Configuration:

Command	**switchport trunk native vlan** *vlan-id*
Parameter Description	*vlan-id*: indicates a VLAN ID.
Command Mode	Interface configuration mode
Usage Guide	To restore the native VLAN of an Uplink to defaults, use the **no switchport trunk native vlan** command.

Verification

- Send tag packets to an Uplink port, and they are broadcast within the specified VLANs.
- Use commands **show vlan** and **show interface switchport** to check whether the configuration takes effect.

Command	**show vlan** [**id** *vlan-id*]
Parameter Description	*vlan-id* **:** indicates a VLAN ID.
Command Mode	Any mode
Usage Guide	N/A
Command Display	FS(config-vlan)#show vlan id 20 VLAN Name Status Ports ---- ---------- --------- ---------- **20 VLAN0020 STATIC Gi0/1**

Configuration Example

Configuring an Uplink Port

Configuration Steps	The following is an example of configuring Gi0/1 as an Uplink port.
	FS# configure terminal FS(config)# interface gi 0/1 FS(config-if-GigabitEthernet 0/1)# switchport mode uplink FS(config-if-GigabitEthernet 0/1)# end
Verification	Check whether the configuration is correct.
	FS# show interfaces GigabitEthernet 0/1 switchport Interface Switchport Mode Access Native Protected VLAN lists ---------- --------- --------- ------ ------ --------- ---------- GigabitEthernet 0/1 enabled **UPLINK** 1 1 disabled ALL

4.4.4 Configuring a Hybrid Port

Configuration Effect

- A Hybrid port is usually used in SHARE VLAN environment. By default, a Hybrid port is the same as a Trunk port. Their difference is that a Hybrid port can send the frames from the VLANs except the default VLAN in the untagged format.

Configuration Steps

Configuring a Hybrid Port

- Mandatory.
- Configure a Hybrid port to transmit the flows from multiple VLANs.
- Configuration:

Command	**switchport mode hybrid**
Parameter Description	N/A
Defaults	The default mode is Access, which can be modified to Hybrid.
Command Mode	Interface configuration mode
Usage Guide	To restore all properties of a Hybrid port to defaults, use the **no switchport mode** command.

Defining Allowed-VLANs for a Hybrid Port

- Optional.
- By default, a Hybrid port transmits the flows from all the VLANs (1 to 4094). You may configure a list of allowed-VLANs to prohibit flows of some VLANs from passing through a Hybrid port.
- Configuration:

Command	**switchport hybrid allowed vlan** [[**add** | **only**] **tagged** | [**add**] **untaged** | **remove**] *vlan_list*
Parameter Description	*vlan-id*: indicates a VLAN ID.
Defaults	By default a Hybrid port belongs to all VLANs. The port is added to the default VLAN in untagged form and to the other VLANs in the tagged form.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring a Native VLAN

- Optional.
- If a frame carries the VLAN ID of a native VLAN, its tag will be stripped automatically when it passes a Hybrid port.
- Configuration:

Command	**switchport hybrid native vlan** *vlan_id*
Parameter Description	*vlan-id*: indicates a VLAN ID.
Defaults	The default native VLAN is VLAN 1.
Command Mode	Interface configuration mode
Usage Guide	To restore the native VLAN of a Hybrid port to defaults, use the **no switchport hybrid native vlan** command.

Verification

- Send tagged packets to an Hybrid port, and they are broadcast within the specified VLANs.
- Use commands **show vlan** and **show interface switchport** to check whether the configuration takes effect.

Command	**show vlan** [**id** *vlan-id*]
Parameter Description	*vlan-id* **:** indicates a VLAN ID.
Command Mode	Any mode
Usage Guide	N/A
Command Display	FS(config-vlan)#show vlan id 20 VLAN Name Status Ports ---- ---------- --------- ---------- **20 VLAN0020 STATIC Gi0/1**

Configuration Example

Configuring a Hybrid Port

Configuration Steps	The following is an example of configuring Gi0/1 as a Hybrid port.
	FS# configure terminal FS(config)# interface gigabitEthernet 0/1

	FS(config-if-GigabitEthernet 0/1)# switchport mode hybrid FS(config-if-GigabitEthernet 0/1)# switchport hybrid native vlan 3 FS(config-if-GigabitEthernet 0/1)# switchport hybrid allowed vlan untagged 20-30 FS(config-if-GigabitEthernet 0/1)# end
Verification	Check whether the configuration is correct.

	FS(config-if-GigabitEthernet 0/1)#show run interface gigabitEthernet 0/1
	
	Building configuration...
	Current configuration : 166 bytes
	
	interface GigabitEthernet 0/1
	  switchport
	switchport mode **hybrid**
	  switchport hybrid **native vlan 3**
	  switchport hybrid allowed vlan add **untagged 20-30**

4.5 Monitoring

Displaying

Description	Command
Displays VLAN configuration.	**show vlan**
Displays configuration of switch ports.	**show interface switchport**

Debugging

System resources are occupied when debugging information is output. Disable the debugging switch immediately after use.

Description	Command
Debugs VLANs.	**debug bridge vlan**

5 Configuring Super VLAN

5.1 Overview

Super virtual local area network (VLAN) is an approach to dividing VLANs. Super VLAN is also called VLAN aggregation, and is a management technology tailored for IP address optimization.

Using super VLAN can greatly save IP addresses. Only one IP address needs to be assigned to the super VLAN that consists of multiple sub VLANs, which greatly saves IP addresses and facilitates network management.

5.2 Application

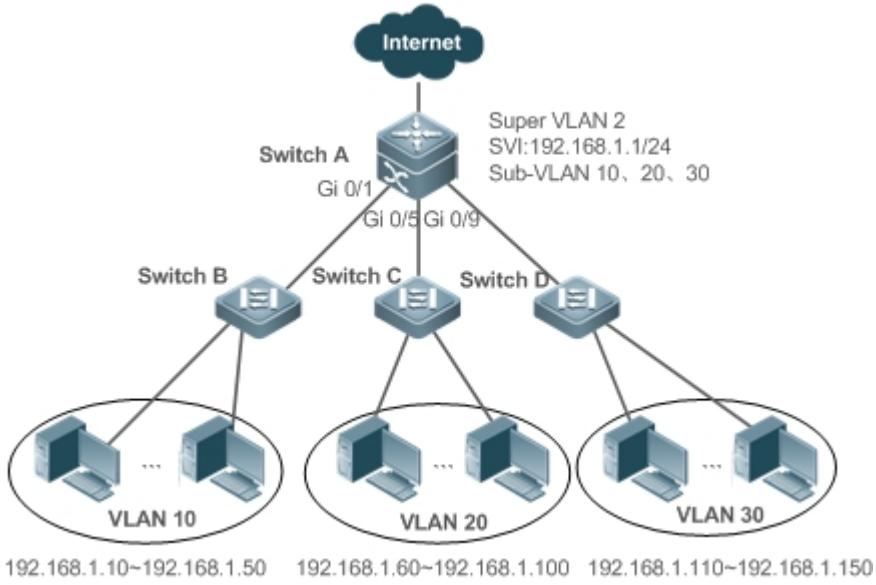
Application	Description
Sharing One IP Gateway Among Multiple VLANs	VLANs are divided to implement layer-2 (L2) isolation of access users. All VLAN users share one IP gateway to implement layer-3 (L3) communication and communication with external networks.

5.2.1 Sharing One IP Gateway Among Multiple VLANs

Scenario

Multiple VLANs are isolated at L2 on a L3 device, but users of these VLANs can perform L3 communication with each other in the same network segment.

Figure 5- 1



Remarks	Switch A is a gateway or core switch. Switch B, Switch C, and Switch D are access switches. On Switch A, a super VLAN and multiple sub VLANs are configured, and a L3 interface and the IP address of the L3 interface are configured for the super VLAN. VLAN 10 is configured on Switch B, VLAN 20 is configured on Switch C, and VLAN 30 is configured on Switch D. Different departments of the company reside in different VLANs.

Deployment

On the intranet, use the super VLAN so that multiple sub VLANs can share one IP gateway and meanwhile VLANs are mutually isolated at L2.

Users in sub VLANs can perform L3 communication through the gateway of the super VLAN.

5.3 Features

Basic Concepts

Super VLAN

Super VLAN is also called VLAN aggregation, and is a management technology tailored for IP address optimization. It aggregates multiple VLANs to one IP network segment. No physical port can be added to a super VLAN. The switch virtual interface (SVI) is used to manage the cross-VLAN communication of sub VLANs. The super VLAN cannot be used as a common 802.1Q VLAN, but can be treated as the primary VLAN of sub VLANs.

Sub VLAN

A sub VLAN is an independent broadcast domain. Sub VLANs are mutually isolated at L2. Users of sub VLANs of the same or different super VLANs communicate with each other through the L3 SVIs of their own super VLANs.

ARP Proxy

A L3 SVI can be created only for a super VLAN. Users in a sub VLAN communicates with users in other sub VLANs of the same super VLAN or users in other network segments through the ARP proxy and the L3 SVI of the super VLAN. When a user of a sub VLAN sends an ARP request to a user of another sub VLAN, the gateway of the super VLAN uses its own MAC addressto send or respond to the ARP requests. The process is called ARP proxy.

IP Address Range of the Sub VLAN

Based on the gateway IP address configured for the super VLAN, an IP address range can be configured for each sub VLAN.

Overview

Feature	Description
Super VLAN	Create a L3 interface as anSVI to allow all sub VLANs to share the same IP network segment through the ARP proxy.

5.3.1 Super VLAN

Users of all sub VLANs of a super VLAN can be allocated IP addresses in the same IP address range, and share the same IP gateway. Users can implement cross-VLAN communication through this gateway. It is unnecessary to allocate a gateway for every VLAN, which saves the IP addresses.

Working Principle

IP addresses in a network segment are allocated to different sub VLANs that belong to the same super VLAN. Each sub VLAN has an independent broadcast domain of the VLAN, and different sub VLANs are isolated from each other at L2. When users in sub VLANs need to perform L3 communication, the IP address of the SVI of the super VLAN is used as the gateway address. In this way, multiple VLANs

share the same IP gateway, and it is unnecessary to configure a gateway for every VLAN. In addition, to implement L3 communication between sub VLANs and between sub VLANs and other network segments, the ARP proxy function is used to forward and process the ARP requests and responses.

L2 communication of sub VLANs: If the SVI is not configured for the super VLAN, sub VLANs of super VLAN are mutually isolated at L2, that is, users in different sub VLANs cannot communicate with each other. If the SVI is configured for the super VLAN, and the gateway of the super VLAN can function as the ARP proxy, users in different sub VLANs of the same super VLAN can communicate with each other. This is because IP addresses of users in different sub VLANs belong to the same network segment, and communication between these users is still treated as L2 communication.

L3 communication of sub VLANs: If users in sub VLANs of a super VLAN need to perform L3 communication across network segments, the gateway of this super VLAN functions as the ARP proxy to respond to the ARP requests in place of sub VLANs.

5.4 Configuration

Configuration Item	Description and Command	
Configuring Basic Functions of the Super VLAN	Mandatory.	
	supervlan	Configures a super VLAN.
	subvlan*vlan-id-list*	Configures a sub VLAN.
	proxy-arp	Enables the ARP proxy function.
	interface vlan*vlan-id*	Creates a virtual interface for a super VLAN.
	ip address*ip mask*	Configures the IP address of the virtual interface of a super VLAN.
	Optional.	
	subvlan-address-range *start-ip end-ip*	Specifies the IP address range in a sub VLAN.

5.4.1 Configuring Basic Functions of the Super VLAN

Configuration Effect

Enable the super VLAN function and configure an SVI for the super VLAN to implement L2/L3 communication between sub VLANs across VLANs.

Users in all sub VLANs of a super VLAN share the same IP gateway. It is unnecessary to specify a network segment for every VLAN, which saves the IP addresses.

Notes

A super VLAN does not belong to any physical port. Therefore, the device configured with the super VLAN cannot process packets that contain the super VLAN tag.

Both the super VLAN function and the ARP proxy function of each sub VLAN must be enabled.

An SVI and an IP address must be configured for a super VLAN. The SVI is a virtual interface used for communication of users in all sub VLANs.

Configuration Steps

Configuring a Super VLAN

- Mandatory.
- No physical port exists in a super VLAN.
- The ARP proxy function must be enabled. This function is enabled by default.
- You can run the **supervlan** command to change a common VLAN into a super VLAN.
- After a common VLAN becomes a super VLAN, ports added to this VLAN will be deleted from this VLAN because no physical port exists in a super VLAN.

A super VLAN is valid only after you configure sub VLANs for this super VLAN.

VLAN 1 cannot be configured as a super VLAN.

A super VLAN cannot be configured as a sub VLAN of another super VLAN. A sub VLAN of a super VLAN cannot be configured as a super VLAN.

Command	**supervlan**
Parameter Description	N/A
Defaults	By default, a VLAN is a common VLAN.
Command Mode	VLAN configuration mode
Usage Guide	By default, the super VLAN function is disabled. No physical port can be added to a super VLAN. Once a VLAN is not a super VLAN, all its sub VLANs become common static VLANs.

Configuring a Virtual Interface for a Super VLAN

- Mandatory.
- No physical port can be added to a super VLAN. You can configure the L3 SVI for a VLAN.

When a super VLAN is configure with an SVI, it allocates a L3 interface i to each sub VLANs. If a sub VLAN is not allocated a L3 interfacedue to resource deficiency, the sub VLAN becomes a common VLAN again.

Command	**interface vlan***vlan-id*
Parameter Description	*vlan-id*: Indicates the ID of the super VLAN.
Defaults	By default, no super VLAN is configured.
Command Mode	Global configuration mode
Usage Guide	A L3 interface must be configured as the virtual interface of a super VLAN.

Configuring the Gateway of a Super VLAN

- Mandatory.
- The IP gateway on the L3 SVI is configured as the proxy for all users in sub VLANs to respond to ARP requests.

Command	**ip address***ip mask*
Parameter Description	*ip*: Indicates the IP address of the gateway on the virtual interface of a super VLAN. *Mask*: Indicates the mask.
Defaults	By default, no gateway is configured for a super VLAN.
Command Mode	Interface configuration mode
Usage Guide	Run this command to configure the gateway for a super VLAN. Users of all sub VLANs of the super VLAN share this gateway.

Configuring a Sub VLAN

- Mandatory.
- Physical ports can be added to sub VLANs. Sub VLANs of a super VLAN share the gateway address of the super VLAN and reside in the same network segment.
- The ARP proxy function must be enabled. This function is enabled by default.
- You can run the **subvlan***vlan-id-list* command to change a common VLAN into a sub VLAN of a super VLAN. Physical ports can be added to sub VLANs.
- Communication of users in a sub VLAN is managed by the super VLAN.

You must change a sub VLAN into a common VLAN before you can delete this sub VLAN by running the **no vlan** command.

One sub VLAN belongs to only one super VLAN.

Command	**subvlan***vlan-id-list*
Parameter Description	*vlan-id-list* : Specifies multiple VLANs as sub VLANs of a super VLAN.
Defaults	By default, a VLAN is a common VLAN.
Command Mode	VLAN configuration mode
Usage Guide	Connection interfaces can be added to a sub VLAN. You must change a sub VLAN into a common VLAN before you can delete this sub VLAN by running the **no vlan** [*id*]command. You cannot configure a L3 SVI of the VLAN for a sub VLAN. If you have configured a L3 SVI for a super VLAN, the attempt of adding more sub VLANs may fail due to resource deficiency. If you configure sub VLANs to a super VLAN, and then configure a L3 SVI of the VLAN for a super VLAN, some sub VLANs may become common VLANs again due to resource deficiency.

Configuring the ARP Proxy

- (Mandatory) The ARP proxy function is enabled by default.
- Users in sub VLANs can implement L2/L3 communication across VLANs through the gateway proxy only after the ARP proxy function is enabled on both the super VLAN and sub VLANs.

- Users in sub VLANs can communicate with users of other VLANs only after the ARP proxy function is enabled on both the super VLAN and sub VLANs.

The ARP proxy function must be enabled on both the super VLAN and sub VLANs.Otherwise, this function does not take effect.

Command	**proxy-arp**
Parameter Description	N/A
Defaults	By default, the ARP proxy function is enabled.
Command Mode	VLAN configuration mode
Usage Guide	By default, the ARP proxy function is enabled. Run this command to enable the ARP proxy function on both the super VLAN and sub VLANs. Users in sub VLANs can implement L2/L3 communication across VLANs only after the ARP proxy function is enabled on both the super VLAN and sub VLANs.

Configuring the IP Address Range of the Sub VLAN

- You can allocate an IP address range to each sub VLAN. Users in a sub VLAN can communicate with users of other VLANs only when their IP addresses are in the specified range.
- Unless otherwise specified, you do not need to configure the IP address range.

IP addresses dynamically allocated to users through DHCP may not be in the allocated IP address range. If the IP addresses allocated through DCHP are not in the specified range, users in a sub VLAN cannot communicate with users of other VLANs. Therefore, be cautious in using the **subvlan-address-range** *start-ip end-ip*command.

The IP address range of a sub VLAN must be within the IP address range of the super VLAN to which the sub VLAN belongs.Otherwise, users in sub VLANs cannot communicate with each other.

IP addresses of users in a sub VLAN must be within the IP address range of the sub VLAN.Otherwise, users in the sub VLAN cannot communicate with each other.

Command	**subvlan-address-range** *start-ip end-ip*
Parameter Description	*start-ip*: Indicates the start IP address of a sub VLAN. *end-ip*: Indicates the end IP address of a sub VLAN.
Defaults	By default, no IP address range is configured.
Command Mode	VLAN configuration mode
Usage Guide	Optional. Run this command to configure the IP address range of users in a sub VLAN. IP address ranges of different sub VLANs of a super VLAN cannot overlap with each other. The IP address range of a sub VLAN must be within the IP address range of the super VLAN to which the sub VLAN belongs. Otherwise, users in sub VLANs cannot communicate with each other. Users in a sub VLAN can communicate with users of other VLANs only when their IP addresses (either dynamically allocated through DHCP or statically configured) are in the configured IP address range. IP addresses allocated through DHCP may not be in the configured IP address range. In this case, users in a sub

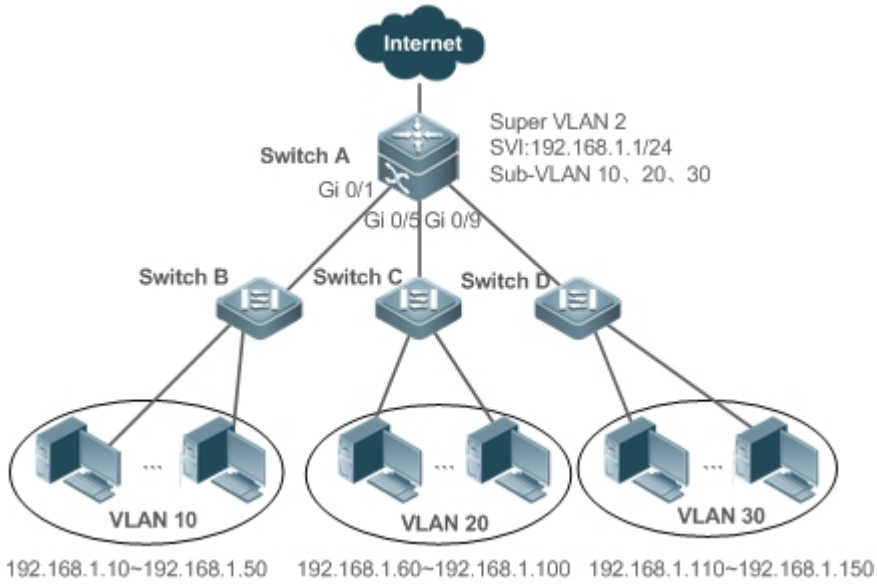
	VLAN cannot communicate with users of other VLANs. Therefore, be cautious when using this command.

Verification

After each sub VLAN is correlated with the gateway of the super VLAN, users in sub VLANs can ping each other.

Configuration Example

↘ Configuring a Super VLAN on the Network so That Users in its Sub VLANs Use the Same Network Segment and Share the Same IP Gateway to Save IP Addresses

Scenario **Figure 5- 2**	
Configuration Steps	Perform the related super VLAN configuration on the core switch. On the access switches, configure the common VLANs corresponding to the sub VLANs on the core switch.
A	SwitchA#configure terminal Enter configuration commands, one per line. End with CNTL/Z. SwitchA(config)#vlan 2 SwitchA(config-vlan)#exit SwitchA(config)#vlan 10 SwitchA(config-vlan)#exit SwitchA(config)#vlan 20 SwitchA(config-vlan)#exit SwitchA(config)#vlan 30 SwitchA(config-vlan)#exit SwitchA(config)#vlan 2 SwitchA(config-vlan)#supervlan

	SwitchA(config-vlan)#subvlan 10,20,30 SwitchA(config-vlan)#exit SwitchA(config)#interface vlan 2 SwitchA(config-if-VLAN 2)#ip address 192.168.1.1 255.255.255.0 SwitchA(config)#vlan 10 SwitchA(config-vlan)#subvlan-address-range 192.168.1.10 192.168.1.50 SwitchA(config-vlan)#exit SwitchA(config)#vlan 20 SwitchA(config-vlan)#subvlan-address-range 192.168.1.60 192.168.1.100 SwitchA(config-vlan)#exit SwitchA(config)#vlan 30 SwitchA(config-vlan)#subvlan-address-range 192.168.1.110 192.168.1.150 SwitchA(config)#interface range gigabitEthernet 0/1,0/5,0/9 SwitchA(config-if-range)#switchport mode trunk
Verification	Verify that the source host (192.168.1.10) and the destination host (192.168.1.60) can ping each other.
A	SwitchA(config-if-range)#show supervlan supervlan id supervlan arp-proxy subvlan id subvlan arp-proxy subvlan ip range ---------- ---------- ---------- ---------- --------- 2 ON 10 ON192.168.1.10 - 192.168.1.50 20 ON 192.168.1.60 - 192.168.1.100 30 ON 192.168.1.110 - 192.168.1.150

Common Errors

The SVI and IP gateway are not configured for the super VLAN. Consequently, communication fails between sub VLANs and between sub VLANs and other VLANs.

The ARP proxy function is disabled on the super VLAN or sub VLANs. Consequently, users in sub VLANs cannot communicate with users of other VLANs.

The IP address range of the sub VLAN is configured, but IP addresses allocated to users are not in this range.

5.5 Monitoring

Displaying

Description	Command
Displays the super VLAN configuration.	**show supervlan**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the super VLAN.	**debug bridge svlan**

6 Configuring MAC VLAN

6.1 Overview

The MAC VLAN function refers to assigning VLANs based on MAC addresses, which is a new method of VLAN assignment. This function is often used with 802.1Xdynamic VLAN assignment to implement secure and flexible access of 802.1Xterminals. After an 802.1Xuser passes authentication, the access switch automatically generates a MAC VLAN entry based on the VLAN and user MAC address pushed by the authentication server. A network administrator can also configure the association between a MAC address and a VLAN on the switch in advance.

Protocols

- IEEE 802.1Q: Virtual Bridged Local Area Networks and Standards

6.2 Applications

Application	Description
Configuring MAC VLAN	Configures the MAC VLAN function to assign VLANs based on users' MAC addresses. When the physical location of a user changes, i.e. switching from one switch to another, it is unnecessary to re-configure the VLAN of the port used by the user.

6.2.1 Configuring MAC VLAN

Scenario

With popularization of mobile office, terminal devices usually do not use fixed ports for network access. A terminal device may use port A to access the network this time, but use port B to access the network next time. If the VLAN configurations of ports A and B are different, the terminal device will be assigned to a different VLAN in the second access, and fail to use the resources of the previous VLAN. If the VLAN configurations of ports A and B are the same, security issues may be introduced when port B is assigned to other terminal devices. How to allow hosts of different VLANs to access the network on the same port? The MAC VLAN function is hereby introduced.

The biggest advantage of MAC VLAN lies in that when the physical location of a user changes, i.e. switching from one switch to another, it is unnecessary to re-configure the VLAN of the port used by the user. Therefore, MAC address-based VLAN assignment can be regarded as user-based.

Deployment

- Configure or push MAC VLAN entries on a layer-2 switch to assign VLANs based on users' MAC addresses.

6.3 Overview

Feature

Feature	Description
Configuring MAC VLAN	Configures the MAC VLAN function to assign VLANs based on users' MAC addresses.

6.3.1 Configuring MAC VLAN

Working Principle

When a switch receives a packet, the switch compare the source MAC address of the packet with the MAC address specified in a MAC VLAN entry. If they match, the switch forwards the packet to the VLAN specified in the MAC VLAN entry. If they don't match, the VLAN to which the data stream belongs is still determined by the VLAN assignment rule of the port.

To ensure that a PC is assigned to a specified VLAN no matter which switch it is connected to, you can perform configuration by using the following approaches:

- Static configuration by using commands. You can configure the association between a MAC address and a VLAN on a local switch by using commands.
- Automatic configuration by using an authentication server (802.1Xdynamic VLAN assignment). After a user passes authentication, a switch dynamically creates an association between the MAC address and a VLAN based on the information provided by the authentication server. When the user goes offline, the switch automatically deletes the association. This approach requires that the MAC-VLAN association be configured on the authentication server. For details about 802.1Xdynamic VLAN assignment, refer to the Configuring 802.1X.

MAC VLAN entries support both of the two approaches, that is, the entries can be configured on both a local switch and an authentication server. The configurations can take effect only if they are consistent. If the configurations are different, the configuration performed earlier takes effect.

- The MAC VLAN function can be configured on hybrid ports only.
- MAC VLAN entries are effective only for untagged packets, but not effective for tagged packets.
- For MAC VLAN entries statically configured or dynamically generated, the specified VLANs must exist.
- VLANs specified in MAC VLAN entries cannot be Super VLANs (but can be Sub VLANs), Remote VLANs, or Primary VLANs (but can be Secondary VLANs).
- MAC addresses specified in MAC VLAN entries must be unicast addresses.
- MAC VLANs are effective for all hybrid ports that are enabled with the MAC VLAN function.

6.4 Configuration

Configuration	Description and Command	
Enabling MAC VLAN on a Port	(Mandatory) It is used to enable the MAC VLAN function on a port.	
	mac-vlan enable	Enables MAC VLAN on a port.
Adding a Static MAC VLAN Entry Globally	(Optional) It is used to bind MAC addresses with VLANs.	
	mac-vlan mac-address	Configures a static MAC VLAN entry.

6.4.1 Enabling MAC VLAN on a Port

Configuration Effect

Enable the MAC VLAN function on a port so that MAC VLAN entries can take effect on the port.

Notes

N/A

Configuration Steps

Enabling MAC VLAN on a Port

- Mandatory.
- By default, the MAC VLAN function is disabled on ports and all MAC VLAN entries are ineffective on the ports.
- Enable MAC VLAN on a switch.

Command	**mac-vlan enable**
Parameter Description	N/A
Defaults	The MAC VLAN function is disabled on a port.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

- Run the **show mac-vlan interface** command to display information about the ports enabled with the MAC VLAN function.

Command	**show mac-vlan interface**
Parameter Description	N/A
Command Mode	Privileged configuration mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A
Command Display	FS# show mac-vlan interface MAC VLAN is enabled on following interface: ---------- **FastEthernet 0/1**

Configuration Example

Enabling MAC VLAN on a Port

Configuration Steps	• Enable the MAC VLAN function on the Fast Ethernet 0/10 port.
	FS# configure terminal FS(config)# interface FastEthernet0/10 FS(config-if-FastEthernet 0/10)# mac-vlan enable
Verification	• Check the information about the port enabled with the MAC VLAN function.
	FS# show mac-vlan interface MAC VLAN is enabled on following interface: ---------- **FastEthernet 0/10**

Common Errors

When the MAC VLAN function is enabled on a port, the port is not configured as a layer-2 port (such as switch port or AP port) in advance.

6.4.2 Adding a Static MAC VLAN Entry Globally

Configuration Effect

- Configure a static MAC VLAN entry to bind a MAC addresses with a VLAN. The 802.1p priority can be configured, which is 0 by default.

Notes

N/A

Configuration Steps

Adding a Static MAC VLAN Entry

- Optional.
- To bind a MAC addresses with a VLAN, you should perform this configuration. The 802.1p priority can be configured, which is 0 by default.
- Add a static MAC VLAN entry on a switch.

Command	**mac-vlan mac-address** *mac-address* [**mask** *mac-mask*] **vlan** *vlan-id* [**priority** *pri_val*]
Parameter Description	**mac-address** *mac-address*: Indicates a MAC address. **mask** *mac-mask*: Indicates a mask. **vlan** *vlan-id*: Indicates the associated VLAN. **priority** *pri_val*: Indicates the priority.
Defaults	No static MAC VLAN entry is configured by default.
Command Mode	Global configuration mode
Usage Guide	N/A

- If an untagged packet is matched with a MAC VLAN entry, the packet is modified to the VLAN specified by the MAC VLAN entry once arriving at the switch since the MAC VLAN entry has the highest priority. Subsequent functions and protocols are implemented based on the modified VLAN. Possible influences are as follows:
- If an 802.1Xuser fails to be authenticated, the hybrid port jumps to VLAN 100 specified by the FAIL VLAN function; however, the MAC VLAN entry statically configured redirects all packets of this user to VLAN 200. Consequently, the user cannot implement normal communication in FAIL VLAN 100.
- After an untagged packet is matched with a MAC VLAN entry, the VLAN that triggers MAC address learning is the VLAN redirected based on the MAC VLAN entry.
- For a port that is enabled with the MAC VLAN function, if received packets are matched with both MAC VLAN entries with full F masks and those without full F masks, the packets are processed based on the MAC VLAN entries without full F masks.
- If an untagged packet is matched with both a MAC VLAN entry and a VOICE VLAN entry, the packet priority is modified simultaneously. The priority of the VOICE VLAN entry is used as that of the packet.

If an untagged packet is matched with both a MAC VLAN entry and a PROTOCOL VLAN entry, the VLAN carried in the packet should be the MAC VLAN.

The MAC VLAN function is applied only to untagged packets, but not applied to PRIORITY packets (packets whose VLAN tag is 0 and carrying COS PRIORITY information) and the processing actions are uncertain.

The QoS packet trust model on a switch is disabled by default, which will change PRIORITY of all packets to 0 and overwrite the modification on packet priorities by the MAC VLAN function. Run the mls qos trust cos command in the interface configuration mode to enable the QoS trust model and trust packet priorities.

Deleting All Static MAC VLAN Entries

- Optional.
- To delete all static MAC VLAN entries, you should perform this configuration.
- Perform this configuration on a switch.

Command	**no mac-vlan all**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Deleting the Static MAC VLAN Entry of a Specified MAC Address

- Optional.
- To delete the MAC VLAN entry of a specified MAC address, you should perform this configuration.
- Perform this configuration on a switch.

Command	**no mac-vlan mac-address** *mac-address* [**mask** *mac-mask*]
Parameter Description	**mac-address** *mac-address*: Indicates a MAC address. **mask** *mac-mask*: Indicates a mask.
Command Mode	Global configuration mode
Usage Guide	N/A

Deleting the Static MAC VLAN Entry of a Specified VLAN

- Optional.
- To delete the MAC VLAN entry of a specified VLAN, you should perform this configuration.
- Perform this configuration on a switch.

Command	**no mac-vlan vlan** *vlan-id*
Parameter Description	**vlan** *vlan-id*: Indicates a VLAN.
Command Mode	Global configuration mode
Usage Guide	N/A

Verification

- Run the **show mac-vlan static** command to check whether all static MAC VLAN entries are correct.
- Run the **show mac-vlan vlan** *vlan-id* command to check whether the MAC VLAN entry of a specified VLAN is correct.
- Run the **show mac-vlan mac-address** *mac-address* [**mask** *mac-mask*] command to display the MAC VLAN entry of a specified MAC address.

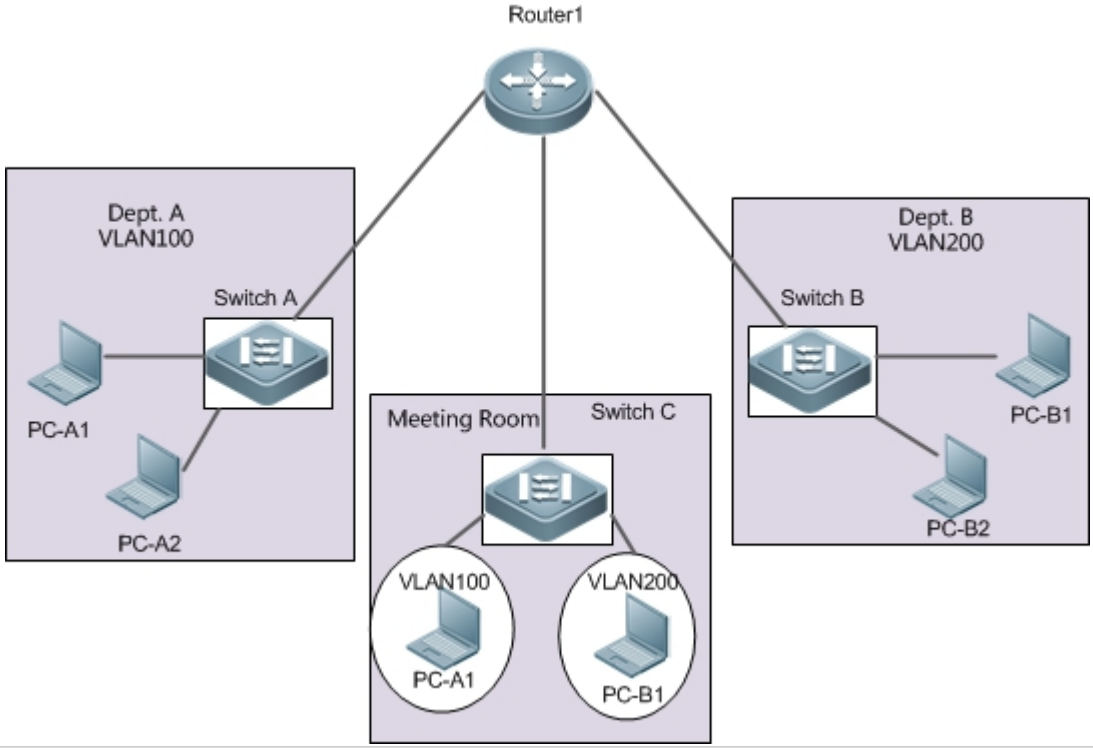
Command	**show mac-vlan static** **show mac-vlan vlan** *vlan-id* **show mac-vlan mac-address** *mac-address* [**mask** *mac-mask*]
Parameter Description	**vlan** *vlan-id*: Indicates a specified VLAN. **mac-address** *mac-address*: Indicates a specified MAC address. **mask** *mac-mask*: Indicates a specified mask.
Command Mode	Privileged configuration mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A
Command Display	FS# show mac-vlan all The following MAC VLAN address exist: S: Static D: Dynamic MAC ADDR MASK VLAN ID PRIO STATE ---------- **0000.0000.0001 ffff.ffff.ffff 2 0 D** **0000.0000.0002 ffff.ffff.ffff 3 3 S** **0000.0000.0003 ffff.ffff.ffff 3 3 S&D** **Total MAC VLAN address count: 3**

Configuration Example

Adding a Static MAC VLAN Entry Globally

As shown in Figure 5-1,PC-A1 and PC-A2 belong to department A and are assigned to VLAN 100. PC-B1 and PC-B2 belong to department B and are assigned to VLAN 200. Due to employee mobility, the company provides a temporary office at the meeting room but requires that accessed employees be assigned to the VLANs of their own departments. For example, PC-A1 must be assigned to VLAN 100 and PC-B1 must be assigned to VLAN 200 after access.

Since the access ports for PCs at the meeting room are not fixed, the MAC VLAN function can be used to associate the PC MAC addresses with the VLANs of their departments. No matter which ports the employees use for access, the MAC VLAN function automatically assigns the VLANs of their departments.

Scenario **Figure 5- 1**	
Configuration Steps	● Configure the port connecting Switch C and Router 1 as a Trunk port. ● Configure all ports connecting PCs on Switch C as hybrid ports, enable the MAC VLAN function and modify the default untagged VLAN list. ● Configure MAC VLAN entries on Switch C.
A	A# configure terminal A(config)# interface interface_name A(config-if)# switchport mode trunk A(config-if)# exit A(config)# interface interface_name A(config-if)# switchport mode hybrid A(config-if)# switchport hybrid allowed vlan add untagged 100,200 A(config-if)# mac-vlan enable A(config-if)# exit A(config)# mac-vlan mac-address PC-A1-mac vlan 100 A(config)# mac-vlan mac-address PC-B1-mac vlan 200
Verification	Check the configured static MAC VLAN entries on Switch C.
A	A# FS# show mac-vlan static The following MAC VLAN address exist: S: Static D: Dynamic

	MAC ADDR MASK VLAN ID PRIO STATE ---------- **PC-A1-macffff.ffff.ffff 100 0 S** **PC-B1-macffff.ffff.ffff 200 3 S** **Total MAC VLAN address count: 2**

6.5 Monitoring

Displaying

Description	Command
Displays all the MAC VLAN entries, including static and dynamic.	**show mac-vlan all**
Displays the dynamic MAC VLAN entries.	**show mac-vlan dynamic**
Displays the static MAC VLAN entries.	**show mac-vlan static**
Displays the MAC VLAN entries of a specified VLAN.	**show mac-vlan vlan** *vlan-id*
Displays the MAC VLAN entries of a specified MAC address.	**show mac-vlan mac-address** *mac-address* [**mask** *mac-mask*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the MAC VLAN function.	**debug bridge mvlan**

7 Configuring Protocol VLAN

7.1 Overview

The protocol VLAN technology is a VLAN distribution technology based on the packet protocol type. It can distribute packets of a certain protocol type with a null VLAN ID to the same VLAN. That is, the switch, based on the protocol type and encapsulation format of packets received by ports, matches the received untagged packets with protocol profiles. If the matching is successful, the switch automatically distributes the packets to a relevant VLAN for transmission. There are two types of protocol VLANs: IP address-based protocol VLAN and protocol VLAN based on the packet type and Ethernet type on ports. The protocol VLAN based on the packet type and Ethernet type on ports is called protocol VLAN for short and the IP address-based protocol VLAN is called subnet VLAN for short.

The protocol VLAN is applicable only to Trunk ports and Hybrid ports.

Protocols and Standards

IEEE standard 802.1Q

7.2 Applications

Application	Description
Configuration and Application of Protocol VLAN	Implements Layer-2 communication isolation of user hosts that use different protocol packets for communication to reduce the network traffic.
Configuration and Application of Subnet VLAN	Specifies the VLAN range based on the IP network segment to which user packets belong.

7.2.1 Configuration and Application of Protocol VLAN

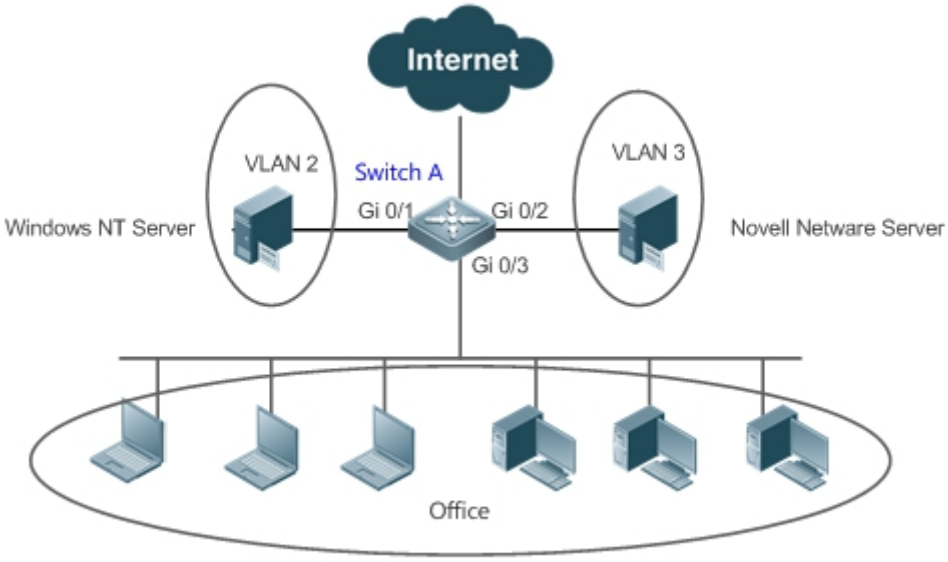
Scenario

As shown in the following figure, the network architecture is composed of the interconnected Windows NT server and Novell Netware server and the office area is connected to the Layer-3 device Switch A through a hub. There are different PCs in the office area. Some PCs use the Windows NT operating system (OS) and support the IP protocol, and some PCs use the Novell Netware OS and support the IPX protocol. PCs in the office area communicate with the external network and servers through the uplink port Gi 0/3.

The main requirements are as follows:

- The Layer-2 communication of PCs using the Windows NT OS is isolated from that of PCs using the Novell Netware OS, so as to reduce the network traffic.

Figure 6-1



Remarks	Switch A is a switch and Port Gi 0/3 is a Hybrid port. Port Gi 0/1 is an Access port and belongs to VLAN 2. Port Gi 0/2 is also an Access port and belongs to VLAN 3.

Deployment

- Configure profiles of the packet type and Ethernet type (in this example, configure Profile 1 for IP protocol packets and configure Profile 2 for IPX protocol packets).
- Apply the profiles to the uplink port (Port Gi 0/3 in this example) and associate them with VLANs (in this example, associate Profile 1 with VLAN 2 and associate Profile 2 with VLAN 3).

The configured protocol VLANs take effect only on the Trunk ports and Hybrid ports.

7.2.2 Configuration and Application of Subnet VLAN

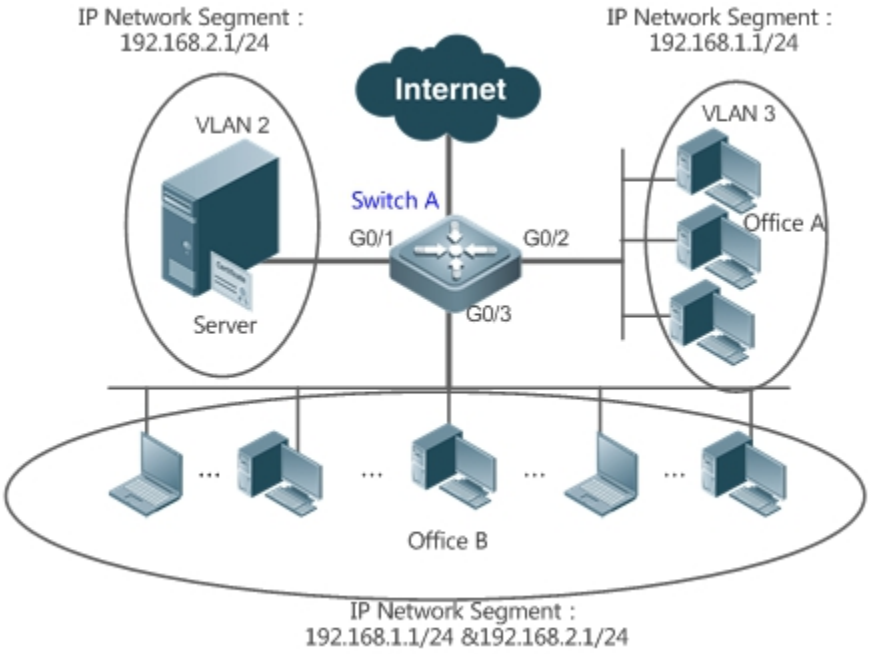
Scenario

As shown in the following figure, PCs in Office A and Office B are connected to the Layer-3 device Switch A through hubs. In Office A, the PCs belong to a fixed network segment and they are distributed to the same VLAN by port. In Office B, the PCs belong to two network segments, but they cannot be distributed to VLANs by fixed port.

The main requirements are as follows:

For PCs in Office B, Switch A can determine the VLAN range of the PCs based on the IP network segment to which their packets belong.

Figure 6-2



Remarks	Switch A is a switch. Port G0/1 is an Access port and belongs to VLAN 2. Port G0/2 is also an Access port and belongs to VLAN 3. Port G0/3 is a Hybrid port.

Deployment

- Globally configure subnet VLANs (in this example, allocate the IP network segment 192.168.1.1/24 to VLAN 3 and the IP network segment 192.168.2.1/24 to VLAN 2) and enable the subnet VLAN function on the uplink port (Port Gi 0/3 in this example).

The configured subnet VLANs take effect only on the Trunk ports and Hybrid ports.

7.3 Features

Basic Concepts

Protocol VLAN

The protocol VLAN technology is a VLAN distribution technology based on the packet protocol type. It can distribute packets of a certain protocol type with a null VLAN ID to the same VLAN.

VLANs need to be specified for packets received by device ports so that a packet belongs to a unique VLAN. There are three possible cases:

- If a packet contains a null VLAN ID (untagged or priority packet) and the device supports only port-based VLAN distribution, the VLAN ID in the tag added to the packet is the PVID of the input port.
- If a packet contains a null VLAN ID (untagged or priority packet) and the device supports VLAN distribution based on the packet protocol type, the VLAN ID in the tag added to the packet is selected from the VLAN IDs mapped to the protocol suite configuration of the input port. If the protocol type of the packet does not match all protocol suite configuration of the input port, a VLAN ID is allocated according to the port-based VLAN distribution.

- If a packet is a tagged packet, the VLAN to which the packet belongs is determined by the VLAN ID in the tag.

Subnet VLANs can be configured only globally that is, only the protocol VLAN function can be enabled or disabled on ports. The matching configuration is globally performed for the protocol VLAN, the matching configuration is selected on ports and the VLAN IDs are specified for packets that are matched successfully.

- If an input packet contains a null VLAN ID and the IP address of the input packet matches an IP address, the packet is distributed to the subnet VLAN.
- If an input packet contains a null VLAN ID and the packet type and Ethernet type of the input packet match the packet type and Ethernet type of an input port, the packet is allocated to the protocol VLAN.

Protocol VLAN Priority

The priority of a subnet VLAN is higher than that of a protocol VLAN. That is, if a subnet VLAN and protocol VLAN are configured at the same time and an input packet conforms to both the subnet VLAN and protocol VLAN, the subnet VLAN prevails.

Overview

Feature	Description
Automatic VLAN Distribution Based on Packet Type	The service types supported on a network are bound with VLANs or packets from a specified IP network segment are transmitted in a specified VLAN to facilitate management and maintenance.

7.3.1 Automatic VLAN Distribution Based on Packet Type

Working Principle

Set rules on the hardware and enable the rules on ports. The rules take effect only after they are enabled on ports. The rules include the packet type and IP address of packets. When a port receives untagged data packets that meet the rules, the port automatically distributes them to the VLAN specified in the rules for transmission. When the rules are disabled on ports, untagged data packets are distributed to the Native VLAN according to the port configuration.

7.4 Configuration

Configuration	Description and Command	
Configuring the Protocol VLAN Function	(Mandatory) It is used to enable the VLAN distribution function based on the packet type and Ethernet type of the protocol VLAN.	
	protocol-vlan profile *num* frame-type [*type*] ether-type [*type*]	Configures the profile of the packet type and Ethernet type.
	protocol-vlan profile *num* ether-type [*type*]	Configures the profile of the Ethernet type (some models do not support frame identification).
	protocol-vlan profile *num* vlan *vid*	(Interface configuration mode) Applies the protocol VLAN on a port.
Configuring the Subnet VLAN Function	(Mandatory) It is used to enable IP address-based VLAN distribution function of the protocol VLAN.	
	protocol-vlan ipv4 *address* mask *address* vlan *vid*	Configures an IP address, subnet mask, and VLAN distribution.

Configuration	Description and Command	
	protocol-vlan ipv4	(Interface configuration mode) Enables the subnet VLAN on a port.

7.4.1 Configuring the Protocol VLAN Function

Configuration Effect

Bind service types supported in a network with VLANs to facilitate management and maintenance.

Notes

- It is recommended that the protocol VLAN be configured after VLANs, and the Trunk, Hybrid, Access, and AP attributes of ports are configured.
- If protocol VLAN is configured on a Trunk port or Hybrid port, all VLANs relevant to the protocol VLAN need to be contained in the permitted VLAN list of the Trunk port or Hybrid port.

Configuration Steps

Configuring the Protocol VLAN Globally

- Mandatory.
- The protocol VLAN can be applied on an interface only in global configuration mode.

Command	**protocol-vlan profile** *num* **frame-type** *type* **ether-type** *type*
Parameter Description	*num:* Indicates the profile index. *type*: Indicates the packet type and Ethernet type.
Defaults	The protocol VLAN is disabled by default.
Command Mode	Global configuration mode
Usage Guide	The protocol VLAN can be configured on an interface only when the protocol VLAN is globally configured. When the global configuration of a protocol VLAN profile is deleted, the protocol VLAN configuration is deleted from all interfaces corresponding to the profile of the protocol VLAN.

Switching the Port Mode to Trunk/Hybrid Mode

- Mandatory. The protocol VLAN function takes effect only on ports that are in Trunk/Hybrid mode.

Enabling the Protocol VLAN on a Port

- Mandatory. The protocol VLAN is disabled by default.
- The protocol VLAN is truly enabled only when it is applied on interfaces.

Command	**protocol-vlan profile *num* vlan *vid***
Parameter Description	*num:* Indicates the profile index. *vid*: Indicates the VLAN ID. The value **1** indicates the maximum VLAN ID supported by the product.
Defaults	The protocol VLAN is disabled by default.
Command Mode	Interface configuration mode

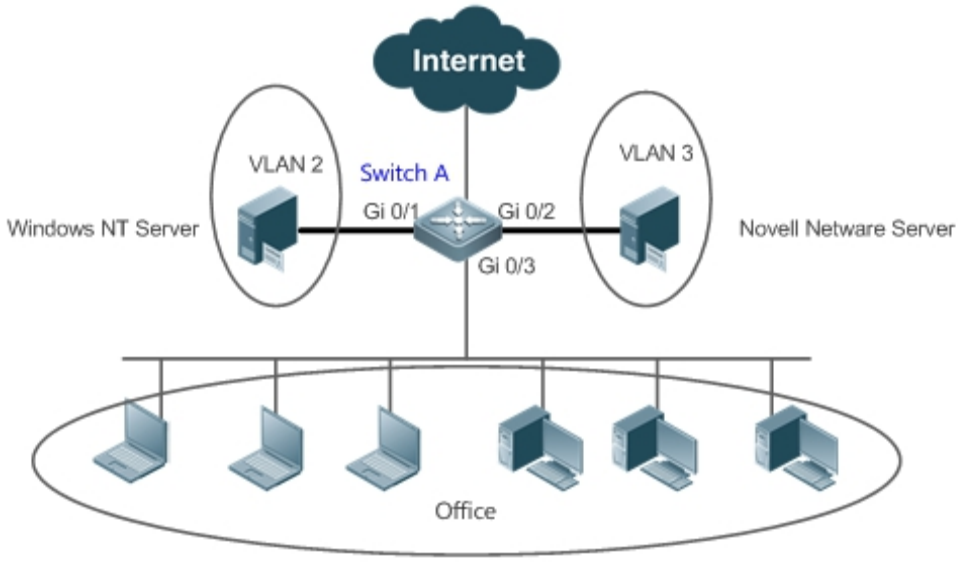
Usage Guide	An interface must work in Trunk/Hybrid mode.

Verification

Run the **show protocol-vlan profile** command to check the configuration.

Configuration Example

Enabling the Protocol VLAN Function in the Topological Environment

Scenario **Figure 6-3**	
Configuration Steps	● Configure VLAN 2 and VLAN 3 for user communication on Switch A. ● Configure the protocol VLAN globally on Switch A (in this example, configure Profile 1 for IP protocol packets and configure Profile 2 for IPX protocol packets), enable the protocol VLAN function on the uplink port (Port Gi 0/3 in this example), and complete the protocol-VLAN association (in this example, associate Profile 1 with VLAN 2 and associate Profile 2 with VLAN 3). ● Port Gi 0/1 is an Access port and belongs to VLAN 2. Port Gi 0/2 is also an Access port and belongs to VLAN 3. Port Gi 0/3 is a Hybrid port. Ensure that the user communication VLANs are contained in the permitted untagged VLAN list of the Hybrid port.
A	1. Create VLAN 2 and VLAN 3 for user network communication. # configure terminal Enter configuration commands, one per line. End with CNTL/Z. A(config)# vlan range 2-3 2. Configure the port mode. A(config)#interface gigabitEthernet 0/1

	A(config-if-GigabitEthernet 0/1)#switchport A(config-if-GigabitEthernet 0/1)#switchport access vlan 2 A(config-if-GigabitEthernet 0/1)#exit A(config)#interface gigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)#switchport A(config-if-GigabitEthernet 0/2)#switchport access vlan 3 A(config-if-GigabitEthernet 0/2)#exit A(config)# interface gigabitEthernet 0/3 A(config-if-GigabitEthernet 0/3)#switchport A(config-if-GigabitEthernet 0/3)# switchport mode hybrid A(config-if-GigabitEthernet 0/3)# switchport hybrid allowed vlan untagged 2-3 3. Configure the protocol VLAN globally. Configure Profile 1 for IP protocol packets and Profile 2 for IPX protocol packets (in this example, assume that packets are encapsulated using Ethernet II and the Ethernet types of IP protocol packets and IPX protocol packets are 0X0800 and 0X8137 respectively). A(config)#protocol-vlan profile 1 frame-type ETHERII ether-type 0x0800 A(config)#protocol-vlan profile 2 frame-type ETHERIIether-type 0x8137 4. Apply Profile 1 and Profile 2 to Port Gi 0/3 and allocate Profile 1to VLAN 2 and Profile 2 to VLAN 3. A(config)# interface gigabitEthernet 0/3 A(config-if-GigabitEthernet 0/3) #protocol-vlan profile 1 vlan 2 A(config-if-GigabitEthernet 0/3) #protocol-vlan profile 2 vlan 3
Verification	Check whether the protocol VLAN configuration on the device is correct.
A	A(config)#show protocol-vlan profile profile frame-type ether-type/DSAP+SSAP interface vlan ------- ---------- ---------- ---------- ---- 1 ETHERII 0x0800 Gi0/3 2 2 ETHERII 0x8137

	Gi0/3 3

Common Errors

- A port connected to the device is not in Trunk/Hybrid mode.
- The permitted VLAN list of the port connected to the device does not contain the user communication VLANs.
- The protocol VLAN function is disabled on a port.

7.4.2 Configuring the Subnet VLAN Function

Configuration Effect

Distribute packets from a specified network segment or IP address to a specified VLAN for transmission.

Notes

- It is recommended that the protocol VLAN be configured after VLANs, and the Trunk, Hybrid, Access, and AP attributes of ports are configured.
- If protocol VLAN is configured on a Trunk port or Hybrid port, all VLANs relevant to the protocol VLAN need to be contained in the permitted VLAN list of the Trunk port or Hybrid port.

Configuration Steps

Configuring the Subnet VLAN Globally

- Mandatory.
- The subnet VLAN can be applied on an interface only in global configuration mode.

Command	**protocol-vlan ipv4** *address* **mask** *address* **vlan** *vid*
Parameter Description	*address*: Indicates the IP address. *vid*: Indicates the VLAN ID. The value **1** indicates the maximum VLAN ID supported by the product.
Defaults	The subnet VLAN is disabled by default.
Command Mode	Global configuration mode
Usage Guide	The subnet VLAN can be enabled on an interface even if the protocol VLAN is not enabled globally. Nevertheless, the subnet VLAN takes effect only when the protocol VLAN is configured globally.

Switching the Port Mode to Trunk/Hybrid Mode

- Mandatory. The subnet VLAN function takes effect only on ports that are in Trunk/Hybrid mode.

Enabling the Subnet VLAN on a Port

- Mandatory. The subnet VLAN is disabled by default.
- The subnet VLAN is truly enabled only when it is applied on interfaces.

Command	**protocol-vlan ipv4**
Parameter Description	N/A
Defaults	The subnet VLAN is disabled by default.

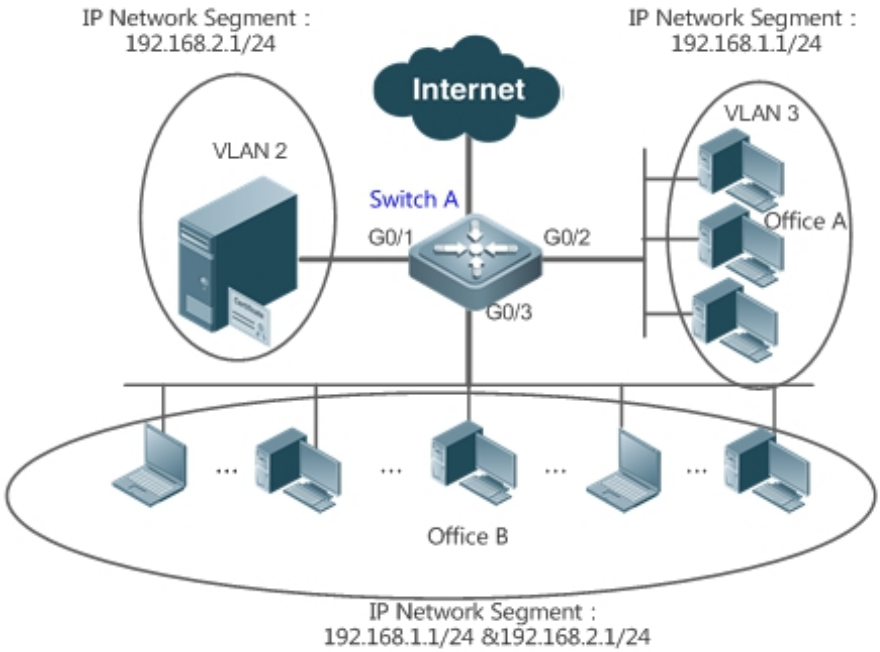
Command Mode	Interface configuration mode
Usage Guide	An interface must work in Trunk/Hybrid mode.

Verification

Run the **show protocol-vlan ipv4** command to check the configuration.

Configuration Example

Enabling the Subnet VLAN Function in the Topological Environment

Scenario **Figure 6-4**	
Configuration Steps	● Configure VLAN 2 and VLAN 3 for user communication on Switch A. ● Globally configure subnet VLANs on Switch A (in this example, allocate the IP network segment 192.168.1.1/24 to VLAN 3 and the IP network segment 192.168.2.1/24 to VLAN 2) and enable the subnet VLAN function on the uplink port (Port Gi 0/3 in this example). ● Port Gi 0/1 is an Access port and belongs to VLAN 2. Port Gi 0/2 is also an Access port and belongs to VLAN 3. Port Gi 0/3 is a Hybrid port. Ensure that the user communication VLANs are contained in the permitted untagged VLAN list of the Hybrid port.
A	1. Create VLAN 2 and VLAN 3 for user network communication. A# configure terminal Enter configuration commands, one per line. End with CNTL/Z. A(config)# vlan range 2-3 2. Configure the port mode.

	A(config)#interface gigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#switchport A(config-if-GigabitEthernet 0/1)#switchport access vlan 2 A(config-if-GigabitEthernet 0/1)#exit A(config)#interface gigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)#switchport A(config-if-GigabitEthernet 0/2)#switchport access vlan 3 A(config-if-GigabitEthernet 0/2)#exit A(config)# interface gigabitEthernet 0/3 A(config-if-GigabitEthernet 0/3)#switchport A(config-if-GigabitEthernet 0/3)# switchport mode hybrid A(config-if-GigabitEthernet 0/3)# switchport hybrid allowed vlan untagged 2-3 3. Configure the subnet VLAN globally. A(config)# protocol-vlan ipv4 192.168.1.0 mask 255.255.255.0 vlan 3 A(config)# protocol-vlan ipv4 192.168.2.0 mask 255.255.255.0 vlan 2 4. Enable the subnet VLAN on interfaces. The subnet VLAN is disabled by default. (config-if-GigabitEthernet 0/3)# protocol-vlan ipv4
Verification	Check whether the subnet VLAN configuration on the device is correct.
A	A# show protocol-vlan ipv4 ip mask vlan --- --- --- 192.168.1.0 255.255.255.0 3 192.168.2.0 255.255.255.0 2 interface ipv4 status --- --- Gi0/3 enable

Common Errors

- A port connected to the device is not in Trunk/Hybrid mode.
- The permitted VLAN list of the port connected to the device does not contain the user communication VLANs.

- The subnet VLAN is disabled on a port.

7.5 Monitoring

Displaying

Description	Command
Displays the protocol VLAN content.	**show protocol-vlan**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the protocol VLAN.	**debug bridge protvlan**

8 Configuring Private VLAN

8.1 Overview

Private VLAN divides the Layer-2 broadcast domain of a VLAN into multiple subdomains. Each subdomain is composed of one private VLAN pair: primary VLAN and secondary VLAN.

One private VLAN domain may consist of multiple private VLAN pairs and each private VLAN pair represents one subdomain. In a private VLAN domain, all private VLAN pairs share the same primary VLAN. The secondary VLAN IDs of subdomains are different.

If a service provider allocates one VLAN to each user, the number of users that can be supported by the service provider is restricted because one device supports a maximum of 4,096 VLANs. On a Layer-3 device, one subnet address or a series of addresses are allocated to each VLAN, which results in the waste of IP addresses. The private VLAN technology properly solves the preceding two problems. Private VLAN is hereinafter called PVLAN for short.

8.2 Applications

Application	Description
Cross-Device Layer-2 Application of PVLAN	Users of an enterprise can communicate with each other but the user communication between enterprises is isolated.
Layer-3 Application of PVLAN on a Single Device	All enterprise users share the same gateway address and can communicate with the external network.

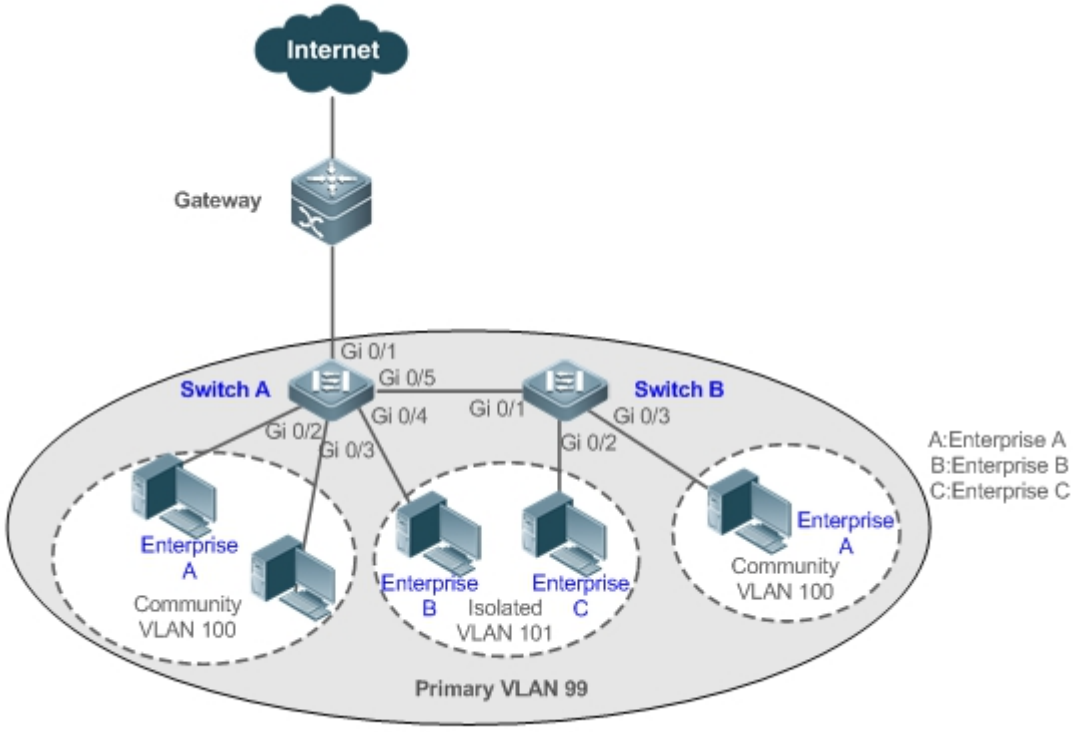
8.2.1 Cross-Device Layer-2 Application of PVLAN

Scenario

As shown in the following figure, in the hosting service operation network, enterprise user hosts are connected to the network through Switch A or Switch B. The main requirements are as follows:

- Users of an enterprise can communicate with each other but the user communication between enterprises is isolated.
- All enterprise users share the same gateway address and can communicate with the external network.

Figure 7- 1



Remarks	Switch A and Switch B are access switches. PVLAN runs across devices. The ports for connecting the devices need to be configured as Trunk ports, that is, Port Gi 0/5 of Switch A and Port Gi 0/1 of Switch B are configured as Trunk ports. Port Gi 0/1 for connecting Switch A to the gateway needs to be configured as a promiscuous port. Port Gi 0/1 of the gateway can be configured as a Trunk port or Hybrid port and the Native VLAN is the primary VLAN of PVLAN.

Deployment

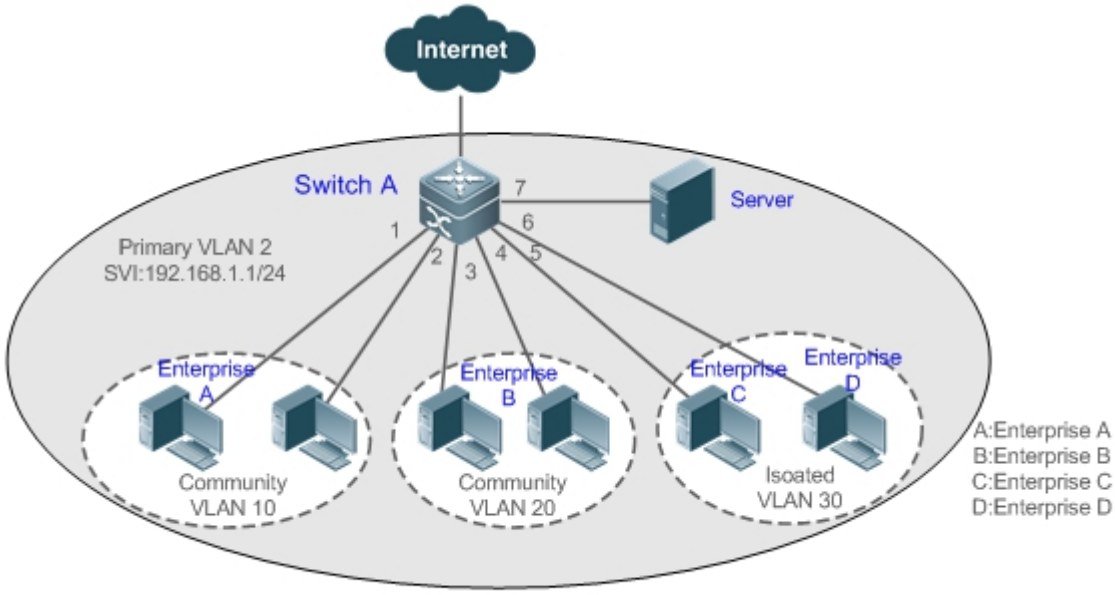
- Configure all enterprises to be in the same PVLAN (primary VLAN 99 in this example). All enterprise users share the same Layer-3 interface through this VLAN to communicate with the external network.
- If an enterprise has multiple user hosts, allocate the user hosts of different enterprises to different community VLANs. That is, configure the ports connected to the enterprise user hosts as the host ports of a community VLAN, so as to implement user communication inside an enterprise but isolate the user communication between enterprises.
- If an enterprise has only one user host, configure the ports connected to the user hosts of such enterprises as the host ports of an isolated VLAN so as to implement isolation of user communication between the enterprises.

8.2.2 Layer-3 Application of PVLAN on a Single Device

As shown in the following figure, in the hosting service operation network, enterprise user hosts are connected to the network through the Layer-3 device Switch A. The main requirements are as follows:

- Users of an enterprise can communicate with each other but the user communication between enterprises is isolated.
- All enterprise users can access the server.
- All enterprise users share the same gateway address and can communicate with the external network.

Figure 7- 2



Remarks	Switch A is a gateway switch. When user hosts are connected to a single device, Port Gi 0/7 for connecting to the server is configured as a promiscuous port so that enterprise users can communicate with the server. Layer-3 mapping needs to be performed on the primary VLAN and secondary VLANs so that the users can communicate with the external network.

Deployment

- Configure the port that is directly connected to the server as a promiscuous port. Then, all enterprise users can communicate with the server through the promiscuous port.
- Configure the gateway address of PVLAN on the Layer-3 device (Switch A in this example) (in this example, set the SVI address of VLAN 2 to 192.168.1.1/24) and configure the mapping between the primary VLAN and secondary VLANs on the Layer-3 interface. Then, all enterprise users can communicate with the external network through the gateway address.

8.3 Features

Basic Concepts

↘ PVLAN

PVLAN supports three types of VLANs: primary VLANs, isolated VLANs, and community VLANs.

A PVLAN domain has only one primary VLAN. Secondary VLANs implement Layer-2 isolation in the same PVLAN domain. There are two types of secondary VLANs.

↘ Isolated VLAN

Ports in the same isolated VLAN cannot mutually make Layer-2 communication. A PVLAN domain has only one isolated VLAN.

Community VLAN

Ports in the same community VLAN can make Layer-2 communication with each other but cannot make Layer-2 communication with ports in other community VLANs. A PVLAN domain can have multiple community VLANs.

Layer-2 Association of PVLAN

PVLAN pairs exist only after Layer-2 association is performed among the three types of VLANs of PVLAN. Then, a primary VLAN has a specified secondary VLAN and a secondary VLAN has a specified primary VLAN. A primary VLAN and secondary VLANs are in the one-to-many relationship.

Layer-3 Association of PVLAN

In PVLAN, Layer-3 interfaces, that is, switched virtual interfaces (SVIs) can be created only in a primary VLAN. Users in a secondary VLAN can make Layer-3 communication only after Layer-3 association is performed between the secondary VLAN and the primary VLAN. Otherwise, the users can make only Layer-2 communication.

Isolated Port

A port in an isolated VLAN can communicate only with a promiscuous port. An isolated port can forward the received packets to a Trunk port but a Trunk port cannot forward the packets with the VID of an isolated VLAN to an isolated port.

Community Port

Community ports are ports in a community VLAN. Community ports in the same community VLAN can communicate with each other and can communicate with promiscuous ports. They cannot communicate with community ports in other community VLANs or isolated ports in an isolated VLAN.

Promiscuous Port

Promiscuous ports are ports in a primary VLAN. They can communicate with any ports, including isolated ports and community ports in secondary VLANs of the same PVLAN domain.

Promiscuous Trunk Port

A promiscuous Trunk port is a member port that belongs to multiple common VLANs and multiple PVLANs at the same time. It can communicate with any ports in the same VLAN.

- In a common VLAN, packet forwarding complies with 802.1Q.
- In PVLAN, for tagged packets to be forwarded by a promiscuous Trunk port, if the VID of the packets is a secondary VLAN ID, the VID is converted into the corresponding primary VLAN ID before packet forwarding.

Isolated Trunk Port

An isolated Trunk port is a member port that belongs to multiple common VLANs and multiple PVLANs at the same time.

- In an isolated VLAN, an isolated Trunk port can communicate only with a promiscuous port.
- In a community VLAN, an isolated Trunk port can communicate with community ports in the same community VLAN and promiscuous ports.
- In a common VLAN, packet forwarding complies with 802.1Q.
- An isolated Trunk port can forward the received packets of an isolated VLAN ID to a Trunk port but a Trunk port cannot forward the packets with the VID of an isolated VLAN to an isolated port.
- For tagged packets to be forwarded by an isolated Trunk port, if the VID of the packets is a primary VLAN ID, the VID is converted

into a secondary VLAN ID before packet forwarding.

In PVLAN, SVIs can be created only in a primary VLAN and SVIs cannot be created in secondary VLANs.

Ports in PVLAN can be used as mirroring source ports but cannot be used as mirroring destination ports.

Overview

Feature	Description
PVLAN Layer-2 Isolation and IP Address Saving	Ports of different PVLAN types can be configured to implement interworking and isolation of VLAN intermediate user hosts.
	After Layer-2 mapping is performed between a primary VLAN and secondary VLANs, only Layer-2 communication is supported. If Layer-3 communication is required, users in a secondary VLAN need to use SVIs of the primary VLAN to make Layer-3 communication.

8.3.1 PVLAN Layer-2 Isolation and IP Address Saving

Add users to subdomains of PVLAN to isolate communication between enterprises and between enterprise users.

Working Principle

Configure PVLAN, configure Layer-2 association and Layer-3 association between a primary VLAN and SubVLANs of PVLAN, and configure ports connected to user hosts, external network devices, and servers as different types of PVLAN ports. In this way, subdomain division and communication of users in subdomains with the external network and servers can be implemented.

Packet Forwarding Relationship Between Ports of Different Types

Output Port / Input Port	Promiscuous Port	Isolated Port	Community Port	Isolated Trunk Port (in the Same VLAN)	Promiscuous Trunk Port (in the Same VLAN)	Trunk Port (in the Same VLAN)
Promiscuous Port	Supported	Supported	Supported	Supported	Supported	Supported
Isolated Port	Supported	Unsupported	Unsupported	Unsupported	Supported	Supported
Community Port	Supported	Unsupported	Supported	Supported	Supported	Supported
Isolated Trunk Port (in the Same VLAN)	Supported	Unsupported	Supported	Unsupported (unsupported in an isolated VLAN but supported in a non-isolated VLAN)	Supported	Supported
Promiscuous Trunk Port (in the Same VLAN)	Supported	Supported	Supported	Supported	Supported	Supported
Trunk Port (in the Same VLAN)	Supported	Unsupported	Supported	Unsupported (unsupported in an isolated VLAN but supported in a	Supported	Supported

Output Port / Input Port	Promiscuous Port	Isolated Port	Community Port	Isolated Trunk Port (in the Same VLAN)	Promiscuous Trunk Port (in the Same VLAN)	Trunk Port (in the Same VLAN)
				non-isolated VLAN)		

VLAN Tag Changes After Packet Forwarding Between Ports of Different Types

Output Port / Input Port	Promiscuous Port	Isolated Port	Community Port	Isolated Trunk Port (in the Same VLAN)	Promiscuous Trunk Port (in the Same VLAN)	Trunk Port (in the Same VLAN)
Promiscuous Port	Unchanged	Unchanged	Unchanged	A secondary VLAN ID is added.	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	A primary VLAN ID tag is added.
Isolated Port	Unchanged	NA	NA	NA	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	An isolated VLAN ID tag is added.
Community Port	Unchanged	NA	Unchanged	A community VLAN ID tag is added.	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	A community VLAN ID tag is added.
Isolated Trunk Port (in the Same VLAN)	The VLAN tag is removed.	NA	The VLAN tag is removed.	The VLAN tag keeps unchanged in a non-isolated VLAN.	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	Unchanged
Promiscuous Trunk Port (in the Same VLAN)	The VLAN tag is removed.	Unchanged	Unchanged	A secondary VLAN ID is added.	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	Unchanged
Trunk Port (in the Same VLAN)	The VLAN tag is removed.	NA	The VLAN tag is removed.	The VLAN tag is converted into a secondary VLAN ID in a primary VLAN and the VLAN tag keeps unchanged in other non-isolated VLANs.	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	Unchanged

Output Port / Input Port	Promiscuous Port	Isolated Port	Community Port	Isolated Trunk Port (in the Same VLAN)	Promiscuous Trunk Port (in the Same VLAN)	Trunk Port (in the Same VLAN)
Switch CPU	Untag	Untag	Untag	A secondary VLAN ID tag is added.	A primary VLAN ID tag is added and the VLAN tag keeps unchanged in the non-PVLAN.	A primary VLAN ID tag is added.

8.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of PVLAN	(Mandatory) It is used to configure a primary VLAN and secondary VLANs.	
	private-vlan {**community** | **isolated** | **primary**}	Configures the PVLAN type.
	(Mandatory) It is used to configure Layer-2 association between a primary VLAN and secondary VLANs of PVLAN to form PVLAN pairs.	
	private-vlan association {*svlist* | **add** *svlist* | **remove** *svlist*}	Configures Layer-2 association between a primary VLAN and secondary VLANs to form PVLAN pairs.
	(Optional) It is used to allocate users to an isolated VLAN or community VLAN.	
	switchport mode private-vlan host	Configures a PVLAN host port.
	switchport private-vlan host-association *p_vid s_vid*	Associates Layer-2 ports with PVLAN and allocates ports to subdomains.
	(Optional) It is used to configure a port as a promiscuous port.	
	Switchport mode private-vlan promiscuous	Configures a PVLAN promiscuous port.
	switchport private-vlan mapping *p_vid* { *svlist* | **add** *svlist* | **remove** *svlist* }	Configures the primary VLAN to which a PVLAN promiscuous port belongs and a list of secondary VLANs. PVLAN packets can be transmitted or received through this port only after the configuration is performed.
	(Optional) It is used to configure Layer-3 communication for users in a secondary VLAN.	
	private-vlan mapping { *svlist* | **add** *svlist* | **remove** *svlist* }	Configures the SVI of the primary VLAN and configures Layer-3 association between the primary VLAN and secondary VLANs after PVLAN is created and Layer-2 association is performed. Users in a SubVLAN can make Layer-3 communication through the SVI of the primary VLAN.

8.4.1 Configuring Basic Functions of PVLAN

Configuration Effect

- Enable PVLAN subdomains to form to implement isolation between enterprises and between enterprise users.
- Implement Layer-3 mapping between multiple secondary VLANs and the primary VLAN so that and multiple VLANs uses the same IP gateway, thereby helping save IP addresses.

Notes

- After a primary VLAN and a secondary VLAN are configured, a PVLAN subdomain exist only after Layer-2 association is performed between them.
- A port connected to a use host must be configured as a specific PVLAN port so that the user host joins a subdomain to implement the real user isolation.
- The port connected to the external network and the port connected to a server must be configured as promiscuous ports so that upstream and downstream packets are forwarded normally.
- Users in a secondary VLAN can make Layer-3 communication through the SVI of the primary VLAN only after Layer-3 mapping is performed between the secondary VLAN and the primary VLAN.

Configuration Steps

Configuring PVLAN

- Mandatory.
- A primary VLAN and a secondary VLAN must be configured. The two types of VLANs cannot exist independently.
- Run the **private-vlan** { **community** | **isolated** | **primary** } command to configure a VLAN as the primary VLAN of PVLAN and other VLANs as secondary VLANs.

Command	**private-vlan** { **community** | **isolated** | **primary** }
Parameter Description	community: Specifies that the VLAN type is community VLAN. isolated: Specifies that the VLAN type is isolated VLAN. primary: Specifies that the VLAN type is the primary VLAN of a PVLAN pair.
Defaults	VLANs are common VLANs and do not have the attributes of PVLAN.
Command Mode	VLAN mode
Usage Guide	This command is used to specify the primary VLAN and secondary VLANs of PVLAN.

Configuring Layer-2 Association of PVLAN

- Mandatory.
- PVLAN subdomains form, and isolated ports, community ports, and Layer-3 association can be configured only after Layer-2 association is performed between the primary VLAN and secondary VLANs of PVLAN.
- By default, after various PVLANs are configured, the primary VLANs and secondary VLANs are independent of each other. A primary VLAN has a secondary VLAN and a secondary VLAN has a primary VLAN only after Layer-2 association is performed.
- Run the **private-vlan association** { *svlist* | **add** *svlist* | **remove** *svlist* } command to configure or cancel the Layer-2 association between the primary VLAN and secondary VLANs of PVLAN. A PVLAN subdomain forms only after Layer-2 association is configured,. The PVLAN subdomain does not exist after Layer-2 association is cancelled. If Layer-2 association is not performed, when isolated ports and promiscuous ports are used to configure associated PVLAN pairs, the configuration will fail or the association between

ports and VLANs will be cancelled.

Command	**private-vlan association** { *svlist* | **add** *svlist* | **remove** *svlist* }
Parameter Description	*svlist*: Specifies the list of secondary VLANs to be associated or disassociated. **add** *svlist*: Adds the secondary VLANs to be associated. **remove** *svlist*: Cancels the association between *svlist* and the primary VLAN.
Defaults	By default, the primary VLAN and secondary VLANs are not associated.
Command Mode	Primary VLAN mode of PVLAN
Usage Guide	This command is used to configure Layer-2 association between a primary VLAN and secondary VLANs to form PVLAN pairs. Each primary VLAN can be associated with only one isolated VLAN but can be associated with multiple community VLANs.

Configuring Layer-3 Association of PVLAN

- If users in a secondary VLAN domain need to make Layer-3 communication, configure a Layer-3 interface SVI for the primary VLAN and then configure Layer-3 association between the primary VLAN and secondary VLANs on the SVI.
- By default, SVIs can be configured only in a primary VLAN. Secondary VLANs do not support Layer-3 communication.
- If users in a secondary VLAN of PVLAN need to make Layer-3 communication, the SVI of the primary VLAN needs to be used to transmit and receive packets.
- Run the **private-vlan mapping** { *svlist* | **add** *svlist* | **remove** *svlist* } command to configure or cancel the Layer-3 association between the primary VLAN and secondary VLANs of PVLAN. Users in a secondary VLAN can make Layer-3 communication with the external network only after Layer-3 association is configured. After Layer-3 association is cancelled, users in a secondary VLAN cannot make Layer-3 communication.

Command	**private-vlan mapping** { *svlist* | **add** *svlist* | **remove** *svlist* }
Parameter Description	svlist: Indicates the list of secondary VLANs, for which Layer-3 mapping needs to be configured. add svlist: Adds the secondary VLANs to be associated with a Layer-3 interface. remove svlist: Cancels the secondary VLANs associated with a Layer-3 interface.
Defaults	By default, the primary VLAN and secondary VLANs are not associated.
Command Mode	Interface configuration mode of the primary VLAN
Usage Guide	A Layer-3 SVI must be configured for the primary VLAN first. Layer-3 interfaces can be configured only in a primary VLAN. Layer-2 association must be performed between associated secondary VLANs and the primary VLAN.

Configuring Isolated Ports and Community Ports

- After the primary VLAN and secondary VLANs of PVLAN as well as Layer-2 association are configured, allocate the device ports connected to user hosts so as to specify the subdomains to which the user hosts belong.
- If an enterprise has only one user host, set the port connected to the user host as an isolated port.
- If an enterprise has multiple user hosts, set the ports connected to the user hosts as community ports.

Command	**switchport mode private-vlan host** **switchport private-vlan host-association** *p_vid s_vid*
Parameter	p_vid: Indicates the primary VLAN ID in a PVLAN pair.

Description	s_vid: Indicates the secondary VLAN ID in a PVLAN pair. The port is an associated port if the VLAN is an isolated VLAN and the port is a community port if the VLAN is a community VLAN.
Defaults	By default, the interface works in Access mode; no private VLAN pairs are associated.
Command Mode	Both commands run in interface configuration mode.
Usage Guide	Both the preceding commands need to be configured. Before a port is configured as an isolated port or promiscuous port, and the port mode must be configured as the host port mode. Whether a port is configured as an isolated port or community port depends on the s_vid parameter. p_vid and s_vid must be respectively the IDs of the primary VLAN and secondary VLAN in a PVLAN pair, on which Layer-2 association is performed. One host port can be associated with only one PVLAN pair.

Configuring a Promiscuous Port

- According to the table listing port packet transmission and receiving rules in section "Features", the single port type of PVLAN cannot ensure symmetric forwarding of upstream and downstream packets. Ports for connecting to the external network or server need to be configured as promiscuous ports to ensure that users can successfully access the external network or server.

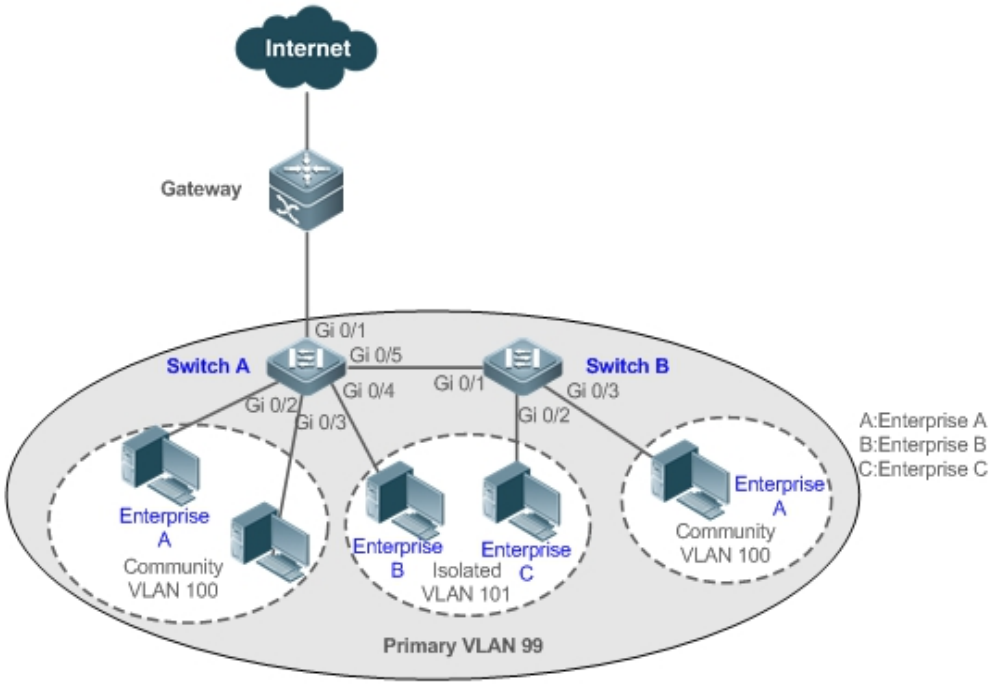
Command	**switchport mode private-vlan promiscuous** **switchport private-vlan mapping** *p_vid*{ *svlist* | **add** *svlist* | **remove** *svlist* }
Parameter Description	p_vid: Indicates the primary VLAN ID in a PVLAN pair. svlist: Indicates the secondary VLAN associated with a promiscuous port. Layer-2 association must be performed between it and p_vid. add svlist: Adds a secondary VLAN to be associated with a port. remove svlist: Cancels the secondary VLAN associated with a port.
Defaults	By default, an interface works in Access mode; a promiscuous port is not associated with a secondary VLAN.
Command Mode	Interface configuration mode
Usage Guide	The port mode must be configured as the promiscuous mode. If a port is configured as a promiscuous port, it must be associated with PVLN pairs. Otherwise, the port cannot bear or forward services. One promiscuous port can be associated with multiple PVLAN pairs within one primary VLAN but cannot be associated with multiple primary VLANs.

Verification

Make user hosts connected to PVLAN ports transmit and receive packets as per PVLAN port forwarding rules to implement isolation. Configure Layer-3 association to make users in the primary VLAN and secondary VLANs of the same PVLAN to share the same gateway IP address and make Layer-3 communication.

Configuration Example

Cross-Device Layer-2 Application of PVLAN

Figure 7- 3	
Configuration Steps	● Configure all enterprises to be in the same PVLAN (primary VLAN 99 in this example). All enterprise users share the same Layer-3 interface through this VLAN to communicate with the external network. ● If an enterprise has multiple user hosts, allocate each enterprise to a different community VLAN (in this example, allocate Enterprise A to Community VLAN 100) to implement user communication inside an enterprise and isolate user communication between enterprises. ● If an enterprise has only one user host, allocate such enterprises to the same isolated VLAN (in this example, allocate Enterprise B and Enterprise C to Isolated VLAN 101) to isolate user communication between enterprises.
A	SwitchA#configure terminal Enter configuration commands, one per line. End with CNTL/Z. SwitchA(config)#vlan 99 SwitchA(config-vlan)#private-vlan primary SwitchA(config-vlan)#exit SwitchA(config)#vlan 100 SwitchA(config-vlan)#private-vlan community SwitchA(config-vlan)#exit SwitchA(config)#vlan 101 SwitchA(config-vlan)#private-vlan isolated SwitchA(config-vlan)#exit SwitchA(config)#vlan 99 SwitchA(config-vlan)#private-vlan association 100-101 SwitchA(config-vlan)#exit

	SwitchA(config)#interface range gigabitEthernet 0/2-3 SwitchA(config-if-range)#switchport mode private-vlan host SwitchA(config-if-range)#switchport private-vlan host-association 99 100 SwitchA(config-if-range)#exit SwitchA(config)#interface gigabitEthernet 0/4 SwitchA(config-if-GigabitEthernet 0/4)#switchport mode private-vlan host SwitchA(config-if-GigabitEthernet 0/4)#switchport private-vlan host-association 99 101 SwitchA(config)#interface gigabitEthernet 0/5 SwitchA(config-if-GigabitEthernet 0/5)#switchport mode trunk SwitchA(config-if-GigabitEthernet 0/5)#exit
B	SwitchB#configure terminal Enter configuration commands, one per line. End with CNTL/Z. SwitchB(config)#vlan 99 SwitchB(config-vlan)#private-vlan primary SwitchB(config-vlan)#exit SwitchB(config)#vlan 100 SwitchB(config-vlan)#private-vlan community SwitchB(config-vlan)#exit SwitchB(config)#vlan 101 SwitchB(config-vlan)#private-vlan isolated SwitchB(config-vlan)#exit SwitchB(config)#vlan 99 SwitchB(config-vlan)#private-vlan association 100-101 SwitchB(config-vlan)#exit SwitchB(config)#interface gigabitEthernet 0/2 SwitchB(config-if-GigabitEthernet 0/2)#switchport mode private-vlan host SwitchB(config-if-GigabitEthernet 0/2)# switchport private-vlan host-association 99 101 SwitchB(config-if-GigabitEthernet 0/2)#exit SwitchB(config)#interface gigabitEthernet 0/3 SwitchB(config-if-GigabitEthernet 0/3)#switchport mode private-vlan host SwitchB(config-if-GigabitEthernet 0/3)# switchport private-vlan host-association 99 100 SwitchB(config-if-GigabitEthernet 0/3)#exit SwitchB(config)#interface gigabitEthernet 0/1 SwitchB(config-if-GigabitEthernet 0/1)#switchport mode trunk

	SwitchB(config-if-GigabitEthernet 0/1)#exit
Verification	Check whether VLANs and ports are correctly configured, and check whether packet forwarding is correct according to packet forwarding rules in section "Features".

A	SwitchA#show running-config
	!
	vlan 99
	  private-vlan primary
	  private-vlan association add 100-101
	!
	vlan 100
	  private-vlan community
	!
	vlan 101
	  private-vlan isolated
	!
	interface GigabitEthernet 0/1
	  switchport mode private-vlan promiscuous
	  switchport private-vlan mapping 99 add 100-101
	!
	interface GigabitEthernet 0/2
	  switchport mode private-vlan host
	  switchport private-vlan host-association 99 100
	!
	interface GigabitEthernet 0/3
	  switchport mode private-vlan host
	  switchport private-vlan host-association 99 100
	!
	interface GigabitEthernet 0/4
	  switchport mode private-vlan host
	  switchport private-vlan host-association 99 101
	!
	interface GigabitEthernet 0/5
	  switchport mode trunk

	``` ! SwitchA# show vlan private-vlan VLAN Type Status Routed Ports Associated VLANs ---------- ---------- 99 primary active Disabled Gi0/1, Gi0/5 100-101 100 community active Disabled Gi0/2, Gi0/3, Gi0/5 99 101 isolated active Disabled Gi0/4, Gi0/5 99 ... ```

B	```
	SwitchB#show running-config
	!
	vlan 99
	  private-vlan primary
	  private-vlan association add 100-101
	!
	vlan 100
	  private-vlan community
	!
	vlan 101
	  private-vlan isolated
	!
	interface GigabitEthernet 0/1
	  switchport mode trunk
	!
	interface GigabitEthernet 0/2
	  switchport mode private-vlan host
	  switchport private-vlan host-association 99 101
	!
	interface GigabitEthernet 0/3
	  switchport mode private-vlan host
	  switchport private-vlan host-association 99 100
	``

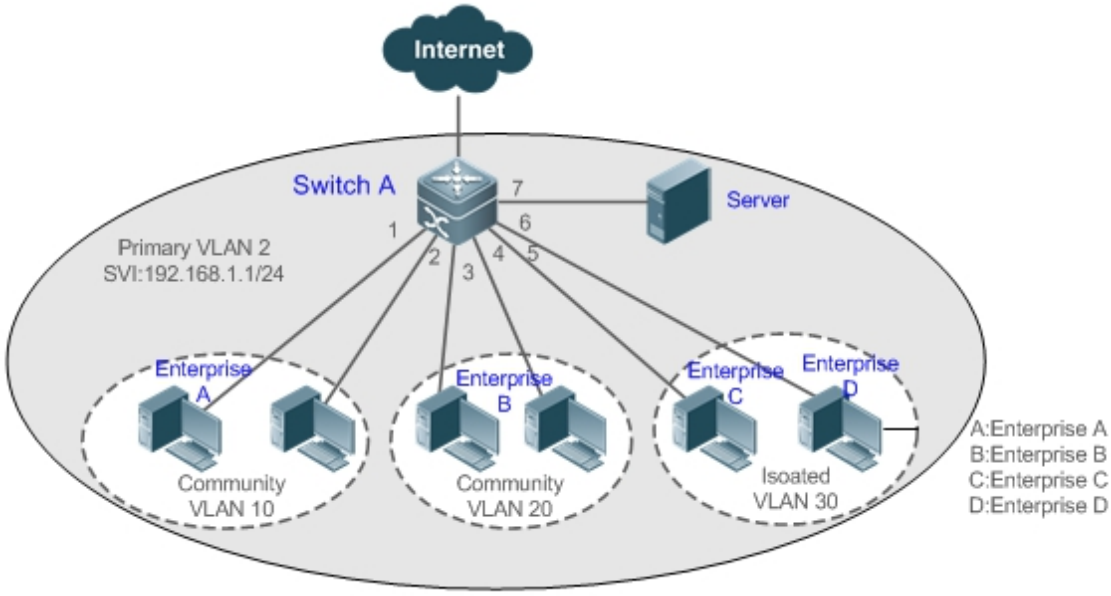
Common Errors

- Layer-2 association is not performed between the primary VLAN and secondary VLANs of PVLAN, and a port VLAN list fails to be added when isolated ports, promiscuous ports, and community ports are configured.

- One host port fails to be associated with multiple PVLAN pairs.

Configuration Example

Layer-3 Application of PVLAN on a Single Device

Figure 7- 4	
Configuration Steps	- Configure the PVLAN function on the device (Switch A in this example). For details about the configuration, see configuration tips in "Cross-Device Layer-2 Application of PVLAN." - Set the port that is directly connected to the server (Port Gi 0/7 in this example) as a promiscuous port. Then, all enterprise users can communicate with the server through the promiscuous port. - Configure the gateway address of PVLAN on the Layer-3 device (Switch A in this example) (in this example, set the SVI address of VLAN 2 to 192.168.1.1/24) and configure the Layer-3 interface mapping between the primary VLAN (VLAN 2 in this example) and secondary VLANs (VLAN 10, VLAN 20, and VLAN 30 in this example). Then, all enterprise users can communicate with the external network through the gateway address. Run PVLAN cross devices and configure the ports for connecting to the devices as Trunk ports.
A	SwitchA#configure terminal Enter configuration commands, one per line. End with CNTL/Z. SwitchA(config)#vlan 2 SwitchA(config-vlan)#private-vlan primary SwitchA(config-vlan)#exit SwitchA(config)#vlan 10 SwitchA(config-vlan)#private-vlan community SwitchA(config-vlan)#exit SwitchA(config)#vlan 20 SwitchA(config-vlan)#private-vlan community SwitchA(config-vlan)#exit

	SwitchA(config)#vlan 30 SwitchA(config-vlan)#private-vlan isolated SwitchA(config-vlan)#exit SwitchA(config)#vlan 2 SwitchA(config-vlan)#private-vlan association 10,20,30 SwitchA(config-vlan)#exit SwitchA(config)#interface range gigabitEthernet 0/1-2 SwitchA(config-if-range)#switchport mode private-vlan host SwitchA(config-if-range)#switchport private-vlan host-association 2 10 SwitchA(config-if-range)#exit SwitchA(config)#interface range gigabitEthernet 0/3-4 SwitchA(config-if-range)#switchport mode private-vlan host SwitchA(config-if-range)#switchport private-vlan host-association 2 20 SwitchA(config-if-range)#exit SwitchA(config)#interface range gigabitEthernet 0/5-6 SwitchA(config-if-range)#switchport mode private-vlan host SwitchA(config-if-range)#switchport private-vlan host-association 2 30 SwitchA(config-if-range)#exit SwitchA(config)#interface gigabitEthernet 0/7 SwitchA(config-if-GigabitEthernet 0/7)#switchport mode private-vlan promiscuous SwitchA(config-if-GigabitEthernet 0/7)#switchport private-vlan maping 2 10,20,30 SwitchA(config-if-GigabitEthernet 0/7)#exit SwitchA(config)#interface vlan 2 SwitchA(config-if-VLAN 2)#ip address 192.168.1.1 255.255.255.0 SwitchA(config-if-VLAN 2)#private-vlan mapping 10,20,30 SwitchA(config-if-VLAN 2)#exit
Verification	Ping the gateway address 192.168.1.1 from user hosts in different subdomains. The ping operation is successful.

A	SwitchA#show running-config
	!
	vlan 2
	  private-vlan primary
	  private-vlan association add 10,20,30

```
!
vlan 10
  private-vlan community
!
vlan 20
  private-vlan community
!
vlan 30
  private-vlan isolated
!
interface GigabitEthernet 0/1
  switchport mode private-vlan host
  switchport private-vlan host-association 2 10
!
interface GigabitEthernet 0/2
  switchport mode private-vlan host
  switchport private-vlan host-association 2 10
!
interface GigabitEthernet 0/3
  switchport mode private-vlan host
  switchport private-vlan host-association 2 20
!
interface GigabitEthernet 0/4
  switchport mode private-vlan host
  switchport private-vlan host-association 2 20
!
interface GigabitEthernet 0/5
  switchport mode private-vlan host
  switchport private-vlan host-association 2 30
!
interface GigabitEthernet 0/6
  switchport mode private-vlan host
  switchport private-vlan host-association 2 30
!
```

```
interface GigabitEthernet 0/7
 switchport mode private-vlan promiscuous
 switchport private-vlan mapping 2 add 10,20,30
!
interface VLAN 2
 no ip proxy-arp
 ip address 192.168.1.1 255.255.255.0
 private-vlan mapping add 10,20,30
!
SwitchA#show vlan private-vlan
VLAN   Type     Status     Routed     Ports    Associated VLANs
------------------------------ ------------------
2       primary      active     Enabled    Gi0/7                10,20,30
10      community   active     Enabled    Gi0/1, Gi0/2       2
20      community   active     Enabled    Gi0/3, Gi0/4       2
30      isolated     active     Enabled    Gi0/5, Gi0/6       2
```

Common Errors

- No Layer-2 association is performed on the primary VLAN and secondary VLANs of PVLAN and the Layer-3 association fails to be configured.
- The device is connected to the external network before Layer-3 association is configured. As a result, the device cannot communicate with the external network.
- The interfaces for connecting to the server and the external network are not configured as promiscuous interfaces, which results in asymmetric forwarding of upstream and downstream packets.

8.5 Monitoring

Displaying

Description	Command
Displays PVLAN configuration.	**show vlan private-vlan**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs PVLAN.	**debug bridge pvlan**

9 Configuring Voice VLAN

9.1 Overview

IP phones are widely used thanks to rapid development of technologies. The voice virtual local area network (VLAN) is a VLAN dedicated to voice data streams of users.

By creating a voice VLAN and add ports connected to voice devices to the voice VLAN, you can transmit voice data in a centralized manner in the voice VLAN, and configure Quality of Service (QoS) for voice streams to improve the transmission priority of voice streams and ensure the voice service quality.

9.2 Applications

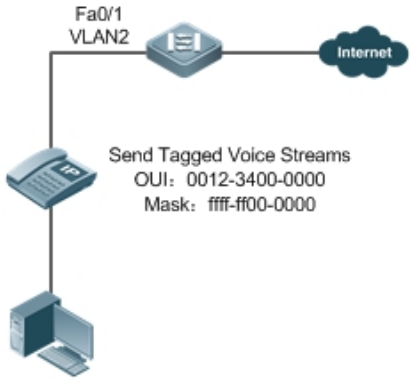
Application	Description
Configuring the Automatic Mode of the Voice VLAN	IP phones and PCs form a daisy chain and are connected to the network. Deployed IP phones can automatically obtain the IP addresses and voice VLAN information and send tagged voice streams.
Configuring the Manual Mode of the Voice VLAN	IP phones are directly connected to the network.
Isolating Voice Streams from Data Streams	PCs are connected to IP phones, and IP phones are connected to a switch. IP phones automatically obtain the IP addresses and send untagged voice streams.

9.2.1 Configuring the Automatic Mode of the Voice VLAN

Scenario

IP phones and PCs form a daisy chain and are connected to the network. Both voice and data streams are transmitted on this link. Voice streams are transmitted in the voice VLAN, whereas data streams are transmitted in the data VLAN. This ensures that voice and data streams do not interfere with each other. This networking is used when common office staff need to use PCs for data communication, and IP phones for voice communication.

Figure 8- 1 Networking When the Voice VLAN Works in Automatic Mode



The Fa 0/1 port is connected to an IP phone that automatically obtains the IP address. After obtaining the IP address in the voice VLAN, the IP phone can be used normally. The Fa 0/1 port is required to forward both voice and data streams and isolate voice streams from data streams. The port can be configured as a trunk port. The native VLAN forwards data streams, whereas the voice VLAN forwards voice streams.

A device supporting the voice VLAN can check whether a stream is a voice stream of a specified voice device based on the source MAC address field in each data packet received by the port. If the source MAC address of a packet in the stream matches the Organizationally Unique Identifier (OUI) configured on the device, the stream is treated as the voice stream and transmitted in the voice VLAN.

The OUI is the first 24 bits of the MAC address. It is a globally unique identifier allocated by the Institute of Electrical and Electronics Engineers (IEEE) to an equipment supplier. You can determine the supplier of a product based on the OUI.

Deployment

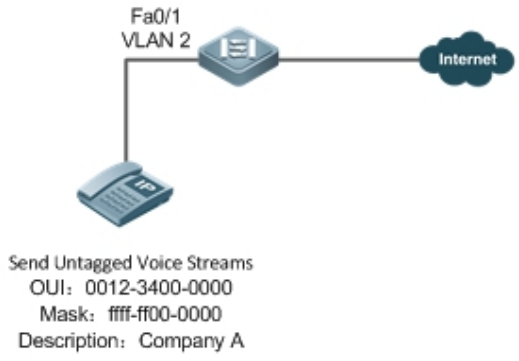
- Enable the port connected to IP phones to work in automatic mode and send tagged voice streams to devices.

9.2.2 Configuring the Manual Mode of the Voice VLAN

Scenario

An IP phone is directly connected to the voice VLAN, and only voice streams exist on the link. This type of networking is generally used when IP phones are deployed in conference rooms or when no PC is required to implement data services.

Figure 8- 2 Networking When the Voice VLAN Works in Manual Mode



The deployed IP phone automatically obtains the IP address, and sends untagged voice streams. As the Fa0/1 port is connected to the IP phone that sends only untagged voice streams, and untagged voice streams do not support the automatic mode, the port can only be set to work in manual mode. The Fa0/1 port is configured as a hybrid port. According to the matching relationship requirement (see "Features"), the native VLAN of the Fa0/1 port must be a voice VLAN, and the voice VLAN must be added to the allowed untagged VLAN list of the port.

Deployment

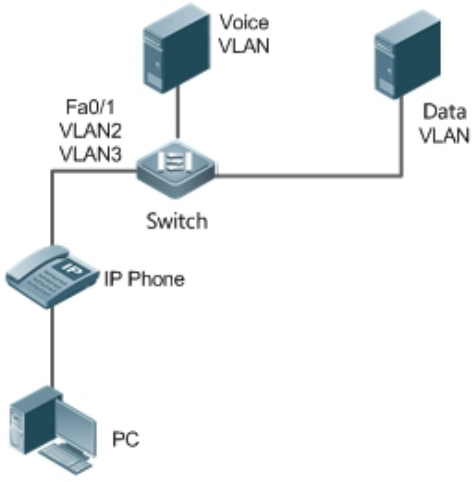
- Enable the port connected to IP phones to work in manual mode and send untagged voice streams to devices.

9.2.3 Isolating Voice Streams from Data Streams

Scenario

To ensure the quality of calls, voice data must be transmitted in the dedicated voice VLAN, and this voice VLAN cannot transmit non-voice data.

Figure 8- 3 Networking That Isolates Voice Streams from Data Streams



The Fa 0/1 port is required to forward both voice and data streams and isolate voice streams from data streams. As both IP phones and PCs send untagged streams, the port must be configured as a hybrid port. The native VLAN forwards data streams, whereas the voice VLAN forwards voice streams. The IP phone connected to the Fa0/1 port sends untagged voice streams. Therefore, the voice VLAN mode must be set to the manual mode. The PC sends untagged data streams. To isolate voice streams from data streams, the MAC VLAN function must be enabled on the Fa0/1 port. The native VLAN of the Fa0/1 port is a data VLAN. In addition, to ensure that received data and voice streams are untagged, both the data VLAN and voice VLAN must be added to the allowed untagged VLAN list of the port. The security mode of the port must be disabled to ensure that data streams can be forwarded.

Deployment

- PCs are connected to IP phones, and IP phones are connected to a switch. IP phones automatically obtain the IP addresses and send untagged voice streams. The security mode is disabled on devices.

9.3 Features

Basic Concepts

↘ Automatic and Manual Modes of the Voice VLAN

Ports in the voice VLAN can work either in automatic or manual mode. The way that ports are added to the voice VLAN varies according to the working mode.

- Automatic mode:

When a packet sent by an IP phone arrives at a device supporting the voice VLAN, the device identifies the source MAC address of the packet and compares this address with the OUI. If the source MAC address matches the OUI, the device automatically adds the input port of the voice packet to the voice VLAN, issues a policy to change the priority of the voice packet to the priority of the voice stream in the voice VLAN configured on the device, and uses the aging mechanism to maintain ports in the voice VLAN. If the system does not receive any voice packet from an input port before the aging timer expires, the system deletes this port from the voice VLAN.

- Manual mode:

In manual port, the administrator manually adds a port to or deletes a port from the voice VLAN. The device identifies the source MAC address of the voice packet sent by the IP phone and compares this address with the OUI configured on the device. If the source MAC address matches the OUI, the device issues a policy to change the priority of the voice packet to the priority of the voice stream in the

voice VLAN configured on the device.

The automatic mode is applicable to the scenario where the PC and IP phone are serially connected to the port and transmit both voice and data streams.

The manual mode is applicable to the scenario where the IP phone is directly connected to a switch and the port transmits only voice packets. In this networking mode, the port is dedicated to transmission of voice streams, which prevents data streams from affecting transmission of voice streams.

Cooperation Between Ports of the Voice VLAN and IP Phones

Based on the way that IP phones obtain IP addresses and voice VLAN information, IP phones are classified into the following types:

- The IP phone automatically obtains the IP address and the voice VLAN ID. This type of IP phones can send both tagged and untagged voice streams.
- The IP address and the voice VLAN ID are manually configured for the IP phone.

Working principle of an IP phone

Like any other network device, an IP phone needs an IP address before it can implement communication normally on the network. An IP phone obtains an IP address in either of the following ways:

- The IP address is automatically obtained through the Dynamic Host Configuration Protocol (DHCP).
- The IP address is manually configured.

When automatically obtaining an IP address, the IP phone can also request the voice VLAN information from the DHCP server. If the DHCP server returns the voice VLAN information, the IP phone can directly send the voice stream containing the voice VLAN tag. If the DHCP server does not return any voice VLAN information, the IP phone sends the voice stream without the voice VLAN tag. If the IP phone support manual configuration of the IP address and voice VLAN ID, you can manually configure the IP address and voice VLAN information on the IP hone. The IP phone sends the tagged or untagged voice streams based on your configuration.

- Voice VLAN Port and IP Phone That Sends Tagged Voice Streams

When sending the tagged voice stream, an IP hone must have obtain the voice VLAN information either automatically or through manual configuration. In this case, different types of ports must be configured accordingly so that voice packets can be transmitted normally in the voice VLAN without affecting forwarding of data streams by the switch. Unlike the IP phone that automatically obtains the voice VLAN information, the IP phone that supports manual configuration of the voice VLAN sends and receives only voice streams that contain the voice VLAN tag.

- Voice VLAN Port and IP Phone That Sends Untagged Voice Streams

An IP phone sends or receives untagged voice streams in either of the following cases:

6. The IP phone automatically obtains an IP address, but not the voice VLAN information.
7. The IP address is manually configured, but the voice VLAN information is not configured.

When the IP phone sends untagged voice streams, you must configure the default VLAN of the input port and add the default VLAN to the allowed VLAN list of the port. In addition, you must configure the default VLAN of the port as a voice VLAN so that the voice stream can be transmitted in the voice VLAN. In this case, the working mode of the voice VLAN of the port can only be set to the manual mode.

The way and process that an IP phone obtains IP address and voice VLAN information vary according to models of IP phones supplied by vendors. The working principle may differ from the preceding description. For details, see the user manual of the IP phone.

The following table describes the relationship between the working mode of the voice VLAN, IP phone type, and port type.

Working Mode of the Voice VLAN	Voice Stream Type	Port Type	Supported Or Not
Automatic mode	Tagged voice stream	Access Port	Not supported.
		Private VLAN host port	Not supported.
		Private VLAN hybrid port	Not supported.
		Trunk port	Supported. The native VLAN connected to the port must exist and cannot be a voice VLAN. In addition, the port allows packets of the native VLAN to pass through.
		Hybrid port	Supported. The native VLAN connected to the port must exist and cannot be a voice VLAN. In addition, the port allows packets of the native VLAN to pass through.
		Uplink port	Supported. The native VLAN connected to the port must exist and cannot be a voice VLAN. In addition, the port allows packets of the native VLAN to pass through.
	Untagged voice stream	Access port	Not supported.
		Private VLAN host port	Not supported.
		Private VLAN hybrid port	Not supported.
		Trunk port	Not supported.
		Hybrid port	Not supported.
		Uplink port	Not supported.
Manual mode	Tagged voice stream	Access port	Not supported.
		Private VLAN host port	Not supported.
		Private VLAN hybrid port	Not supported.
		Trunk port	Supported. The native VLAN connected to the port must exist and cannot be a voice VLAN. In addition, the port allows packets of the native VLAN and the voice VLAN to pass through.
		Hybrid port	Supported. The native VLAN connected to the port must exist and cannot be a voice VLAN. In addition, the port allows packets of the native VLAN to pass through, and the voice VLAN must be in the allowed tagged VLAN list of the port.
		Uplink port	Supported. The native VLAN connected to the port must exist and cannot be a voice VLAN. In addition, the port allows packets of the native VLAN and the voice VLAN to pass through.
	Untagged voice stream	Access port	Supported. The voice VLAN must one of the

Working Mode of the Voice VLAN	Voice Stream Type	Port Type	Supported Or Not
			VLANs to which the connected port is added.
		Private VLAN host port	Supported. The voice VLAN must be configured as the isolated VLAN or community VLAN of the port.
		Private VLAN hybrid port	Supported. The voice VLAN must be configured as the primary VLAN.
		Trunk port	Supported. The native VLAN connected to the port must be a voice VLAN, and the port allows packets of this VLAN to pass through.
		Hybrid port	Supported. The native VLAN connected to the port must be a voice VLAN. (If the MAC VLAN function is enabled on the port to isolate data streams from voice streams, the native VLAN is not necessary a voice VLAN.) In addition, the VLAN must be in the allowed untagged VLAN list of the port.
		Uplink port	Not supported.

If an IP phone sends tagged voice streams, and the 802.1x authentication and guest VLAN functions are enabled on the port connected to the IP phone, you must allocate different VLAN IDs to the voice VLAN, default VLAN of the port, and the guest VLAN of 802.1x to ensure that these functions take effect.

The protocol VLAN takes effect on untagged packets sent by a trunk port or hybrid port. In automatic mode of the voice VLAN, a trunk port or hybrid port can process only tagged voice streams. Therefore, do not configure a VLAN both as a protocol VLAN and a voice VLAN.

If the automatic mode is used, do not set the OUI as a static address; otherwise, the automatic mode is negatively affected.

Security Mode of the Voice VLAN

To better isolate voice streams from data streams during transmission, the voice VLAN provides the security mode. When the security mode is enabled, the voice VLAN allows transmission of only voice streams. In this case, the device checks the source MAC address of each packet. When the source MAC address of a packet is a voice VLAN OUI that can be identified, the packet can be transmitted in the voice VLAN; otherwise, the packet is dropped. When the security mode is disabled, the device does not check the source MAC address of each packet, and all packets can be transmitted in the voice VLAN.

In security mode, the device checks the source MAC address of only an untagged packet or a packet containing the voice VLAN tag. For other packets that do not contain the voice VLAN tag, the device forwards or drops these packets according to the VLAN rules.

You are advised not to transmit voice and data streams concurrently in a voice VLAN. If concurrent transmission of voice and data streams is necessary, confirm that the security mode of the voice VLAN has been disabled.

Overview

Feature	Description

Voice VLAN	Transmit data streams and voice streams respectively in the data VLAN and the voice VLAN to prevent mutual interference between voice calls and data services.

9.3.1 Voice VLAN

Working Principle

A device supporting the voice VLAN transmits data streams and voice streams respectively in the data VLAN and the voice VLAN to prevent mutual interference between voice calls and data services. In addition, the device issues a priority policy to improve the priority of voice streams and ensure the quality of calls. The working principle of the voice VLAN is as follows:

Step 1: The user creates a voice VLAN dedicated to transmission of voice packets on the device, and enable the voice VLAN function on the port that is connected to the IP phone.

Step 2: Add the port connected to the IP phone to the voice VLAN. This step is crucial. The way of adding a port to the voice VLAN varies according to the working mode (automatic or manual) of the voice VLAN:

- In automatic mode, when receiving an untagged packet from the port, the device compares the source MAC address of this packet with the valid OUI. If the source MAC address matches the OUI, the packet is a voice packet. Then, the device automatically adds the port to the voice VLAN, and meanwhile learns the MAC address from this port.
- In manual mode, the user manually adds the port connected to the IP phone to the voice VLAN.

Step 3: Regardless of the working mode (automatic or manual), when a port is added to the voice VLAN, the device issues a policy to improve the priority of every packet with the source MAC address matching the OUI in the voice VLAN. For every voice packet with the source MAC address matching the OUI, the CoS is set to **6**, and DSCP is set to **46**.

After the preceding steps are complete, the port connected to the IP phone is added to the dedicated voice VLAN. Voice packets are transmitted in a centralized manner in the voice VLAN, and is forwarded to other devices with a high priority.

If the IP phone supports the Link Layer Discovery Protocol (LLDP), you do not need to configure the OUI. The device can capture the LLDP packet sent by the IP phone, and identifies the device capability field of the LLDP packet. If the device capability field is "telephone", the device extracts the source MAC address from the LLDP packet as the MAC address of the voice device. In this way, the voice device can be automatically identified.

Related Configuration

Enabling the Voice VLAN Function

By default, the voice VLAN function is disabled.

VLAN 1 cannot be configured as a voice VLAN.

Run the **vlan** *vlan-id* command to create a VLAN.

Run the **voice vlan** *vlan-id* command to enable the voice VLAN function and configure a VLAN as the voice VLAN.

Enabling the Voice VLAN Function on a Port

By default, the voice VLAN function is disabled on a port.

Run the **voice vlan enable** command to enable the voice VLAN function of a port.

Configuring the Voice VLAN Working Mode of a Port

By default, the voice VLAN working mode of a port is set to the automatic mode.

Run the **voice vlan mode auto** command to set the voice VLAN working mode of a port to the automatic mode.

Run the **no voice vlan mode auto** command to set the voice VLAN working mode of a port to the manual mode.

The voice VLAN working modes of ports are independent of each other. You can configure different voice VLAN working modes for different ports.

Configuring the Aging Time of the Voice VLAN

By default, the aging time is 1,440 minutes. The aging time takes effect on ports only in automatic mode. If the port does not receive any voice packet within the aging time, the port is automatically deleted from the voice VLAN. A longer aging time indicates that a port can reside in the voice VLAN for a longer time before it receives any voice packet.

Run the **voice vlan aging** command to configure the aging time of a port.

Configuring the OUI of the Voice VLAN

By default, the OUI is not configured.

Run the **voice vlan mac-address** command to configure the OUI of the voice VLAN that can be identified by devices.

Configuring the Security Mode of the Voice VLAN

By default, the security mode of the voice VLAN is enabled.

Run the **voice vlan security enable** command to enable the security mode of the voice VLAN.

When the security mode is enabled, only voice streams can be transmitted in the voice VLAN.

Configuring the Priority of Voice Streams in the Voice VLAN

By default, the CoS of voice streams is **6** and DSCP is **46**. A higher priority of voice streams indicates a higher priority for transmitting voice packets, which improves the quality of calls.

Run the **voice vlan security cos** command to set the CoS of voice streams in the voice VLAN.

Run the **voice vlan security dscp** command to set the DSCP of voice streams in the voice VLAN.

9.4 Configuration

Configuration Item	Description and Command	
Enabling the Voice VLAN Function	(Mandatory). It is used to globally enable the voice VLAN function.	
	voice vlan	Enables the voice VLAN function and configures a VLAN as the voice VLAN.
Enabling the Voice VLAN Function on a Port	(Mandatory). It is used to enable the voice VLAN function on a port.	
	voice vlan enable	Enables the voice VLAN function on a port.
Configuring the Aging Time of the Voice VLAN	(Optional) It is used to configure the aging time of the voice VLAN.	
	voice vlan aging	Configures the aging time of the voice VLAN. The value ranges from 5 to 10,000 minutes. The default value is 1,440 minutes.
Configuring the OUI of the	(Optional) It is used to configure the OUI of the voice VLAN.	

Configuration Item	Description and Command	
Voice VLAN	**voice vlan mac-address**	Configures the OUI of the voice VLAN that can be identified by devices.
Configuring the Security Mode of the Voice VLAN	(Optional) It is used to configure the security mode of the voice VLAN.	
	voice vlan security enable	Configures the security mode of the voice VLAN.
Configuring the Priority of Voice Streams in the Voice VLAN	(Optional) It is used to configure the priority of voice streams in the voice VLAN.	
	voice vlan cos *cos-value* **voice vlan dscp** *dscp-value*	Configures the priority of voice streams in the voice VLAN.
Configuring the Voice VLAN Working Mode of a Port	(Optional) It is used to configure the voice VLAN working mode of a port.	
	voice vlan mode auto	Sets the voice VLAN working mode of a port to the automatic mode.

9.4.1 Enabling the Voice VLAN Function

Configuration Effect

- Configure a VLAN as the voice VLAN to transmit voice streams.

Notes

- Create a VLAN before configuring the voice VLAN.
- VLAN 1 is the default VLAN and does not need to be created, but VLAN 1 cannot be configured as the voice VLAN.
- A VLAN cannot be configured both as the voice VLAN and super VLAN.
- If 802.1x authentication with VLAN assignment is enabled on the port, do not configure the issued VLAN ID as the voice VLAN ID; otherwise, the function of 802.1x authentication with VLAN assignment is negatively affected.
- Do not configure the same VLAN as the remote VLAN and voice VLAN of the remote switched port analyzer (RSPAN); otherwise the RSPAN and voice VLAN functions may be negatively affected.

Configuration Steps

Configuring a Voice VLAN

- Mandatory.
- Create a VLAN, and configure this VLAN as the voice VLAN for transmission of voice streams.
- Perform this configuration on a switch.

Command	**voice vlan** *vlan-id*
Parameter Description	*vlan-id*: Indicates the ID of a VLAN. The value ranges from 2 to 4,094.
Defaults	By default, the voice VLAN function is disabled.
Command Mode	Global configuration mode
Usage Guide	Run the **no voice vlan** command in global configuration mode to disable the voice VLAN function.

Assume that both the 802.1x and voice VLAN functions are enabled on a port. If the device MAC address of an IP phone matches the OUI of the voice VLAN configured on the FS device, the IP phone can use the voice VLAN for communication without being

authenticated. For example, if a PC and an IP phone are connected to the same port, and the 802.1x function is enabled, the PC must pass the 802.1x authentication before using the network for communication, but the IP phone does not need to be authenticated.

Verification

- Run the **show voice vlan** command to check whether the configuration takes effect.

Command	**show voice vlan**
Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS#show voice vlan Voice VLAN status : **ENABLE** Voice VLAN ID : **10** Voice VLAN security mode: Security Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: PORT MODE ---------- ----------

Configuration Example

Configuring a Voice VLAN

Configuration Steps	• Create VLAN 2. • Globally enable the voice VLAN function, and configure VLAN 2 as a voice VLAN.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# vlan 2 FS(config-vlan)# exit FS(config)# voice vlan 2
Verification	Run the **show voice vlan** command to check whether the configuration is correct.
	FS#show voice vlan

	Voice VLAN status : **ENABLE** Voice VLAN ID : **2** Voice VLAN security mode: Security Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: PORT MODE ---------- ----------

9.4.2 Enabling the Voice VLAN Function on a Port

Configuration Effect

Enable the voice VLAN function of a port connected to an IP phone. This step is mandatory to enable a port to transmit voice streams.

Notes

- The voice VLAN function can be enabled only on a Layer-2 (L2) port, such as the access port, trunk port, hybrid port, uplink port, and private VLAN port. It cannot be enabled on an AP port or a routed port.
- After the voice VLAN function is enabled on a port, to ensure normal operation of the function, do not switch the L2 mode (such as the access port, trunk port, and hybrid port) of the port. If L2 mode switching of a port is necessary, disable the voice VLAN port on this port first.

Configuration Steps

Enabling the Voice VLAN Function on a Port

- Mandatory.
- You must enable the voice VLAN function on a port if you want to use this port for IP phone communication.
- Perform this configuration on a switch.

Command	**voice vlan enable**
Parameter Description	N/A
Defaults	By default, the voice VLAN function is disabled on a port.
Command Mode	Interface configuration mode
Usage Guide	Run the **no voice vlan enable** command to disable the voice VLAN function on a port.

When the voice VLAN function is globally disabled, you can enable the voice VLAN function on a port, but this configuration does not take effect.

Verification

- Run the **show voice vlan** command to check whether the configuration takes effect.

Command	**show voice vlan**

Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: Security Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: **PORT MODE** **--- ---** **Gi0/1 MANUAL**

Configuration Example

Enabling the Voice VLAN Function on a Port

Configuration Steps	● Enter the port configuration mode, and enable the voice VLAN function on a physical port.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface gigabitEthernet 0/1 FS(config-if)# voice vlan enable
Verification	Run the **show voice vlan** command to check whether the configuration is correct.
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: Security Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46

	Current voice VLAN enabled port mode: **PORT MODE** **---------- ----------** **Gi0/1 MANUAL**

9.4.3 Configuring the Aging Time of the Voice VLAN

Configuration Effect

- Configure the aging time of the voice VLAN on a device. If the device does not receive any voice packet from the input port within the aging time, the port is automatically deleted from the voice VLAN. The aging time takes effect only in automatic mode.

Configuration Steps

Configuring the Aging Time of the Voice VLAN

- Optional.
- Perform this configuration if you need to change the time that a port resides in the voice VLAN before the port receives any voice stream.
- Perform this configuration on a switch.

Command	**voice vlan aging** *minutes*
Parameter Description	***minute :*** Indicates the aging time of the voice VLAN.
Defaults	By default, the aging time is 1,440 minutes.
Command Mode	Global configuration mode
Usage Guide	Run the **no voice vlan aging** command in global configuration mode to restore the default aging time.

Verification

- Run the **show voice vlan** command to check whether the configuration takes effect.

Command	**show voice vlan**
Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: Security Voice VLAN aging time : **1440 minutes** Voice VLAN cos : 6

	Voice VLAN dscp : 46 Current voice VLAN enabled port mode: PORT MODE ---------- ----------

Configuration Example

Configuring the Aging Time of the Voice VLAN

Configuration Steps	● Set the aging time of the voice VLAN to 10 minutes.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# voice vlan aging 10
Verification	Run the **show voice vlan** command to check whether the configuration is correct.
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: Security Voice VLAN aging time : **10 minutes** Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: PORT MODE ---------- ----------

9.4.4 Configuring the OUI of the Voice VLAN

Configuration Effect

- FS products support configuration of the OUI of a voice VLAN that can be identified. For details about the OUI, see "Overview". A device supporting the voice VLAN function can compare the source MAC address contained in a received packet with the OUI of the voice VLAN configured on the device to check whether the stream is a voice stream sent from a specified voice device.

Notes

- The OUI of the voice VLAN cannot be a multicast address, and the configured mask must be continuous.

Configuration Steps

Configuring the OUI of the Voice VLAN

- Optional.
- After an IP phone is connected to the device that supports the voice VLAN, you need to configure the OUI of the IP phone so that the IP phone can implement communication on the network.
- Perform this configuration on a switch.

Command	**voice vlan mac-address** *mac-addr* **mask** *oui-mask* [**description** *text*]
Parameter Description	*mac-addr*: Indicates the source MAC address in a voice packet. *oui-mask*: Indicates the valid length of the OUI, which is expressed by a mask. ***text:*** Indicates the description about the OUI.
Defaults	By default, no OUI is configured.
Command Mode	Global configuration mode
Usage Guide	Run the **no voice vlan mac-address oui** command in global configuration mode to delete an OUI configured on a device.

Verification

- Run the **show voice vlan oui** command to check whether the configuration takes effect.

Command	**show voice vlan oui**
Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS(config)# show voice vlan oui **Oui Address Mask Description** **0012.3400.0000 ffff.ff00.0000 Company A**

Configuration Example

Configuring the OUI of the Voice VLAN

Configuration Steps	• Set the OUI of the voice VLAN to 0012.3400.0000, and the supplier is Company A.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# voice vlan mac-address 0012.3400.0000 mask ffff.ff00.0000 description Company A
Verification	Run the **show voice vlan oui** command to check whether the configuration is correct.
	FS(config)# show voice vlan oui **Oui Address Mask Description** **0012.3400.0000 ffff.ff00.0000 Company A**

9.4.5 Configuring the Security Mode of the Voice VLAN

Configuration Effect

- To better isolate voice streams from data streams during transmission, FS products support the security mode of the voice VLAN. When the security mode is enabled, only voice streams can be transmitted in the voice VLAN, which better ensures the quality of voice stream transmission.

Configuration Steps

Configuring the Security Mode of the Voice VLAN

- Optional.
- The security mode of the voice VLAN is configured to isolate voice streams from data streams. Perform this configuration if only voice streams can be transmitted in the voice VLAN.
- Perform this configuration on a switch.

Command	**voice vlan security enable**
Parameter Description	N/A
Defaults	By default, the security mode of the voice VLAN is enabled.
Command Mode	Global configuration mode
Usage Guide	Run the **no voice vlan security enable** command in global configuration mode to disable the security mode of the voice VLAN.

Verification

- Run the **show voice vlan** command to check whether the configuration takes effect.

Command	**show voice vlan**
Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: **Security** Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: PORT MODE

	---------- ----------

Configuration Example

Configuring the Security Mode of the Voice VLAN

Configuration Steps	● Enter the global configuration mode, and enable the security mode of the voice VLAN.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# voice vlan security enable
Verification	Run the **show voice vlan** command to check whether the configuration is correct.
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: **Security** Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: PORT MODE ---------- ----------

9.4.6 Configuring the Priority of Voice Streams in the Voice VLAN

Configuration Effect

- Modify the CoS and DSCP values of voice streams in the voice VLAN to improve the priority of voice streams and ensure the quality of calls. For details about the CoS and DSCP, see "Configuring the QoS".

Configuration Steps

Configuring the Priority of Voice Streams in the Voice VLAN

- Optional.
- Perform this configuration if you need to improve the priority of voice streams.
- Perform this configuration on a switch.

Command	**voice vlan cos** *cos-value* **voice vlan dscp** *dscp-value*
Parameter	*cos-value:* Indicates the CoS value of voice streams in the voice VLAN. The value ranges from 0 to 7.

Description	*dscp-value* **:** Indicates the DSCP value of voice streams in the voice VLAN. The value ranges from 0 to 63.
Defaults	By default, the CoS is **6** and DSCP is **46**.
Command Mode	Global configuration mode
Usage Guide	Run the **no voice vlan cos** command to restore the default CoS value, or the **no voice vlan dscp** command to restore the default DSCP value.

Verification

- Run the **show voice vlan** command to check whether the configuration takes effect.

Command	**show voice vlan**
Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: Security Voice VLAN aging time : 1440 minutes Voice VLAN cos : **6** Voice VLAN dscp : **46** Current voice VLAN enabled port mode: PORT MODE ---------- ----------

Configuration Example

Configuring the Priority of Voice Streams in the Voice VLAN

Configuration Steps	• Set the CoS of voice streams in the voice VLAN to **5** and the DSCP to **40**.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# voice vlan cos 5 FS(config)# voice vlan dscp 40
Verification	Run the **show voice vlan** command to check whether the configuration is correct.
	FS#show voice vlan

```
Voice VLAN status          : ENABLE
Voice VLAN ID              : 10
Voice VLAN security mode: Security
Voice VLAN aging time      : 1440 minutes
Voice VLAN cos             : 5
Voice VLAN dscp            : 40
Current voice VLAN enabled port mode:
PORT                       MODE
-------------------- ----------
```

9.4.7 Configuring the Voice VLAN Working Mode of a Port

Configuration Effect

- The voice VLAN may work in either automatic or manual mode. The working mode of the voice VLAN is configured in port configuration mode. For details about the automatic and manual modes, see "Automatic and Manual Modes of the Voice VLAN".

Notes

- If the voice VLAN function is enabled on a port and the voice VLAN works in manual mode, you must manually add the port to the voice VLAN to ensure that the voice VLAN function can take effect.
- When the voice VLAN function is enabled on a port and the voice VLAN works in automatic mode, do not configure the native VLAN of the port as the voice VLAN; otherwise, the voice VLAN function may be negatively affected.
- By default, the trunk or hybrid port of a FS product can transmit packets of all VLANs. You need to delete the voice VLAN from the allowed VLAN list of a port and then enable the voice VLAN function. In this way, ports that are not connected to any voice device will not be added to the voice VLAN, and ports that are not in use for a long time always reside in the voice VLAN.

Configuration Steps

Setting the Voice VLAN Working Mode of a Port to the Automatic Mode

- Optional.
- Perform this configuration if you want a port to be automatically added to the voice VLAN when the port receive voice streams and automatically deleted from the voice VLAN when the aging time expires.
- Perform this configuration on a switch.

Command	**voice vlan mode auto**
Parameter Description	N/A
Defaults	By default, the voice VLAN of a port works in automatic mode.
Command Mode	Interface configuration mode
Usage Guide	Run the **no voice vlan mode auto** command to set the voice VLAN working mode of a port to the manual mode.

After the voice VLAN function is enabled on a port, the working mode of the voice VLAN cannot be changed. To change the

working mode of the voice VLAN, you must first disable the voice VLAN function on the port.

In automatic mode, you cannot use the manual configuration command (**switchport trunk allow vlan add**) to add a port to or delete a port from the voice VLAN.

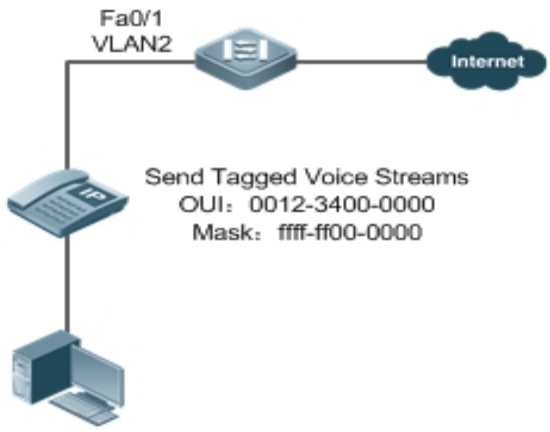
Verification

- Run the **show voice vlan** command to check whether the configuration takes effect.

Command	**show voice vlan**
Parameter Description	N/A
Command Mode	All configuration modes
Usage Guide	N/A
	FS#show voice vlan Voice VLAN status : ENABLE Voice VLAN ID : 10 Voice VLAN security mode: Security Voice VLAN aging time : 1440 minutes Voice VLAN cos : 6 Voice VLAN dscp : 46 Current voice VLAN enabled port mode: **PORT MODE** **--- ---** **Gi0/1 AUTO**

Configuration Example

Adding a Port to the Voice VLAN That Works in Automatic Mode

Scenario **Figure 8- 4**	

Configuration Steps	Configuration Tips ● The Fa 0/1 port is connected to an IP phone that automatically obtains the IP address. After obtaining the IP address in the voice VLAN, the IP phone can be used normally. The Fa 0/1 port is required to forward both voice and data streams and isolate voice streams from data streams. The port can be configured as a trunk port. The native VLAN forwards data streams, whereas the voice VLAN forwards voice streams. ● The PC sends untagged packets. Therefore, the packets will be transmitted in the native VLAN of the port. Configure VLAN 5 as the native VLAN to transmit data streams sent by the PC. ● The network is required to isolate voice streams from data streams. When the port is configured as a trunk port, and the automatic mode is configured as the working mode of the voice VLAN, the native VLAN of the Fa0/1 port must exist and cannot be a voice VLAN according to the matching relationship. In addition, the port must allow packets of the native VLAN to pass through. The ID of the native VLAN is 5, and the native VLAN is not a voice VLAN (VLAN 2). Therefore, the networking requirement for isolating voice streams from data streams can be met. As the trunk port contains all VLANs by default, to better use the automatic mode and prevent ports that are not connected with any voice device from being added to the voice VLAN, the voice VLAN (VLAN 2) must be deleted from the allowed VLAN list of the Fa0/1 port.
	Step 1: Create VLAN 2, and configure this VLAN as the voice VLAN. FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# vlan 2 FS(config-vlan)# exit FS(config)# voice vlan 2 Step 2: Configure data on the device so that the device allows voice packets with the OUI set to 0012.3400.0000 and mask set to ffff.ff00.0000 to be forwarded through the voice VLAN. FS(config)# voice vlan mac-address 0012.3400.0000 mask ffff.ff00.0000 Step 3: Configure Fa0/1 as the trunk port, and VLAN 5 as the native VLAN of the port. FS(config)# interface fastEthernet 0/1 FS(config-if)# switchport mode trunk FS(config-if)# switchport trunk native vlan 5 Step 4: Delete the voice VLAN from the allowed VLAN list of the Fa0/1 port, and enable the voice VLAN function of the Fa0/1 port. FS(config-if)# switchport trunk allowed vlan remove 2 FS(config-if)# voice vlan enable
Verification	● Run the **show voice vlan** command to check the current status of the voice VLAN on the device.
	FS(config)# show voice vlan Voice Vlan status: **ENABLE**

```
Voice Vlan ID              : 2
Voice Vlan security mode: Security
Voice Vlan aging time: 1440 minutes
Voice Vlan cos             : 6
Voice Vlan dscp            : 46
Current voice vlan enabled port mode:
PORT                       MODE
-------------------- ----------
Fa0/1                      AUTO
# Check Voice VLAN OUI Address
FS(config)# show voice vlan oui
Oui Address       Mask                Description
0012.3400.0000    ffff.ff00.0000
```

9.5 Monitoring

Displaying

Description	Command
Displays the voice VLAN configuration.	**show voice vlan**
Displays the OUI configuration of the voice VLAN.	**show voice vlan oui**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the voice VLAN.	**debug bridge vvlan**

10 Configuring MSTP

10.1 Overview

Spanning Tree Protocol (STP) is a Layer-2 management protocol. It cannot only selectively block redundant links to eliminate Layer-2 loops but also can back up links.

Similar to many protocols, STP is continuously updated from Rapid Spanning Tree Protocol (RSTP) to Multiple Spanning Tree Protocol (MSTP) as the network develops.

For the Layer-2 Ethernet, only one active link can exist between two local area networks (LANs). Otherwise, a broadcast storm will occur. To enhance the reliability of a LAN, it is necessary to establish a redundant link and keep some paths in backup state. If the network is faulty and a link fails, you must switch the redundant link to the active state. STP can automatically activate the redundant link without any manual operations. STP enables devices on a LAN to:

- Discover and start the best tree topology on the LAN.
- Troubleshoot a fault and automatically update the network topology so that the possible best tree topology is always selected.

The LAN topology is automatically calculated based on a set of bridge parameters configured by the administrator. The best topology tree can be obtained by properly configuring these parameters.

RSTP is completely compatible with 802.1D STP. Similar to traditional STP, RSTP provides loop-free and redundancy services. It is characterized by rapid speed. If all bridges in a LAN support RSTP and are properly configured by the administrator, it takes less than 1 second (about 50 seconds if traditional STP is used) to re-generate a topology tree after the network topology changes.

STP and RSTP have the following defects:

- STP migration is slow. Even on point-to-point links or edge ports, it still takes two times of the forward delay for ports to switch to the forwarding state.
- RSTP can rapidly converge but has the same defect with STP: Since all VLANs in a LAN share the same spanning tree, packets of all VLANs are forwarded along this spanning tree. Therefore, redundant links cannot be blocked according to specific VLANs and data traffic cannot be balanced among VLANs.

MSTP, defined by the IEEE in 802.1s, resolves defects of STP and RSTP. It cannot only rapidly converge but also can enable traffic of different VLANs to be forwarded along respective paths, thereby providing a better load balancing mechanism for redundant links.

In general, STP/RSTP works based on ports while MSTP works based on instances. An instance is a set of multiple VLANs. Binding multiple VLANs to one instance can reduce the communication overhead and resource utilization.

FS devices support STP, RSTP, and MSTP, and comply with IEEE 802.1D, IEEE 802.1w, and IEEE 802.1s.

Protocols and Standards

- IEEE 802.1D: Media Access Control (MAC) Bridges
- IEEE 802.1w: Part 3: Media Access Control (MAC) Bridges—Amendment 2: Rapid Reconfiguration
- IEEE 802.1s: Virtual Bridged Local Area Networks—Amendment 3: Multiple Spanning Trees

10.2 Applications

Application	Description
MSTP+VRRP Dual-Core Topology	With a hierarchical network architecture model, the MSTP+VRRP mode is used to implement redundancy and load balancing to improve system availability of the network.
BPDU Tunnel	In QinQ network environment, Bridge Protocol Data Unit (BPDU) Tunnel is used to implement tunnel-based transparent transmission of STP packets.

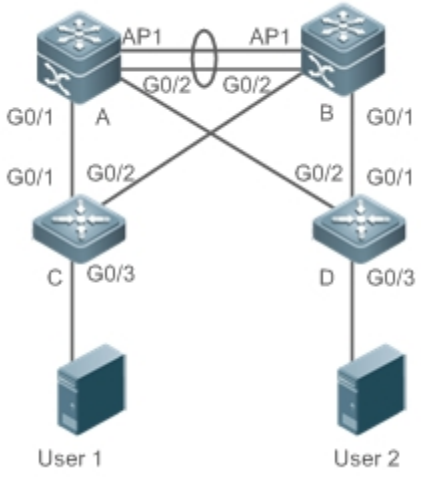
10.2.1 MSTP+VRRP Dual-Core Topology

Scenario

The typical application of MSTP is the MSTP+VRRP dual-core solution. This solution is an excellent solution to improve system availability of the network. Using a hierarchical network architecture model, it is generally divided into three layers (core layer, convergence layer, and access layer) or two layers (core layer and access layer). They form the core network system to provide data exchange service.

The main advantage of this architecture is its hierarchical structure. In the hierarchical network architecture, all capacity indicators, characteristics, and functions of network devices at each layer are optimized based on their network locations and roles, enhancing their stability and availability.

Figure 9- 1 MSTP+VRRP Dual-Core Topology



Remarks	The topology is divided into two layers: core layer (Devices A and B) and access layer (Devices C and D).

Deployment

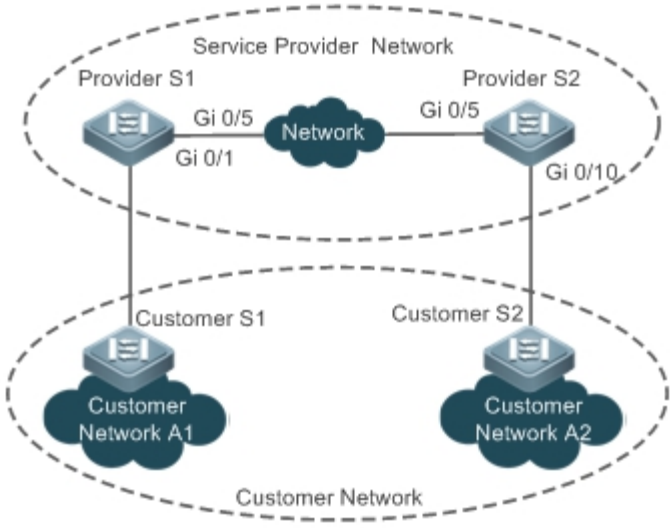
- Core layer: Multiple MSTP instances are configured to realize load balancing. For example, two instances are created: Instance 1 and Instance 2. Instance 1 maps VLAN 10 while Instance 2 maps VLAN 20. Device A is the root bridge of Instances 0 and 1 (Instance 0 is CIST, which exists by default). Device B is the root bridge of Instance 2.
- Core layer: Devices A and B are the active VRRP devices respectively on VLAN 10 and VLAN 20.
- Access layer: Configure the port directly connected to the terminal (PC or server) as a PortFast port, and enable BPDU guard to prevent unauthorized users from accessing illegal devices.

10.2.2 BPDU Tunnel

Scenario

The QinQ network is generally divided into two parts: customer network and service provider (SP) network. You can enable BPDU Tunnel to calculate STP packets of the customer network independently of the SP network, thereby preventing STP packets between the customer network from affecting the SP network.

Figure 9-2 BPDU Tunnel Topology



Remarks	As shown in the above figure, the upper part is the SP network and the lower part is the customer network. The SP network consists of two provider edges (PEs): Provider S1 and Provider S2. Customer Network A1 and Customer Network A2 are a user's two sites in different regions. Customer S1 and Customer S2, access devices from the customer network to the SP network, access the SP network respectively through Provider S1 and Provider S2. Using BPDU Tunnel, Customer Network A1 and Customer Network A2 in different regions can perform unified spanning tree calculation across the SP network, not affecting the spanning tree calculation of the SP network.

Deployment

- Enable basic QinQ on the PEs (Provider S1/Provider S2 in this example) so that data packets of the customer network are transmitted within the specified VLAN on the SP network.
- Enable STP transparent transmission on the PEs (Provider S1/Provider S2 in this example) so that the SP network can transmit STP packets of the customer network through BPDU Tunnel.

10.3 Features

Basic Concepts

BPDU

To generate a stable tree topology network, the following conditions must be met:

- Each bridge has a unique ID consisting of the bridge priority and MAC address.
- The overhead of the path from the bridge to the root bridge is called root path cost.

- A port ID consists of the port priority and port number.

Bridges exchange BPDU packets to obtain information required for establishing the best tree topology. These packets use the multicast address 01-80-C2-00-00-00 (hexadecimal) as the destination address.

A BPDU consists of the following elements:

- Root bridge ID assumed by the local bridge
- Root path cost of the local bridge
- Bridge ID (ID of the local bridge)
- Message age (age of a packet)
- Port ID (ID of the port sending this packet)
- **Forward-Delay Time**, **Hello Time**, **Max-Age Time** are time parameters specified in the MSTP.
- Other flags, such as flags indicating network topology changes and local port status.

If a bridge receives a BPDU with a higher priority (smaller bridge ID and lower root path cost) at a port, it saves the BPDU information at this port and transmits the information to all other ports. If the bridge receives a BPDU with a lower priority, it discards the information.

Such a mechanism allows information with higher priorities to be transmitted across the entire network. BPDU exchange results are as follows:

- A bridge is selected as the root bridge.
- Except the root bridge, each bridge has a root port, that is, a port providing the shortest path to the root bridge.
- Each bridge calculates the shortest path to the root bridge.
- Each LAN has a designated bridge located in the shortest path between the LAN and the root bridge. A port designated to connect the bridge and the LAN is called designated port.
- The root port and designated port enter the forwarding status.

Bridge ID

According to IEEE 802.1W, each bridge has a unique ID. The spanning tree algorithm selects the root bridge based on the bridge ID. The bridge ID consists of eight bytes, of which the last six bytes are the MAC address of the bridge. In its first two bytes (as listed in the following table), the first four bits indicate the priority; the last eight bits indicate the system ID for use in extended protocol. In RSTP, the system ID is 0. Therefore, the bridge priority should be a integral multiple of 4,096.

	Bit	Value
Priority value	16	32,768
	15	16,384
	14	8,192
	13	4,096
System ID	12	2,048
	11	1,024
	10	512
	9	256
	8	128
	7	64
	6	32

	5	16
	4	8
	3	4
	2	2
	1	1

Spanning-Tree Timers

The following three timers affect the performance of the entire spanning tree:

- Hello timer: Interval for periodically sending a BPDU packet.
- Forward-Delay timer: Interval for changing the port status, that is, interval for a port to change from the listening state to the learning state or from the learning state to the forwarding state when RSTP runs in STP-compatible mode.
- Max-Age timer: The longest time-to-live (TTL) of a BPDU packet. When this timer elapses, the packet is discarded.

Port Roles and Port States

Each port plays a role on a network to reflect different functions in the network topology.

- Root port: Port providing the shortest path to the root bridge.
- Designated port: Port used by each LAN to connect the root bridge.
- Alternate port: Alternative port of the root port. Once the root port loses effect, the alternate port immediately changes to the root port.
- Backup port: Backup port of the designated port. When a bridge has two ports connected to a LAN, the port with the higher priority is the designated port while the port with the lower priority is the backup port.
- Disabled port: Inactive port. All ports with the operation state being down play this role.

The following figures show the roles of different ports:

R = Root port D = Designated port A = Alternate port B = Backup port

Unless otherwise specified, port priorities decrease from left to right.

Figure 9-3

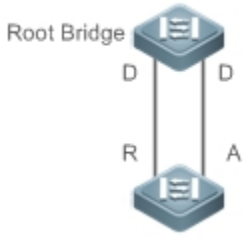


Figure 9-4

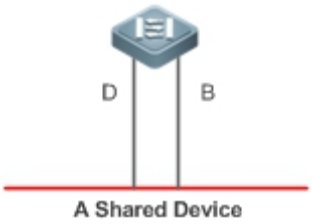
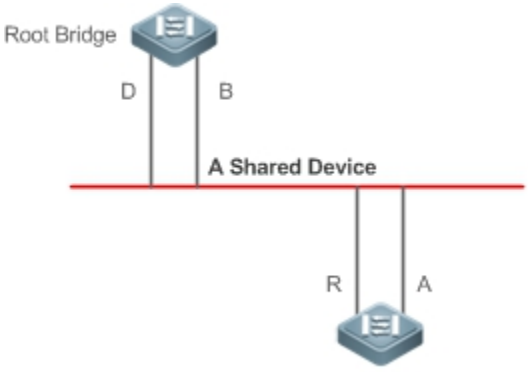


Figure 9-5



Each port has three states indicating whether to forward data packets so as to control the entire spanning tree topology.

- Discarding: Neither forwards received packets nor learns the source MAC address.
- Learning: Does not forward received packets but learns the source MAC address, which is a transitive state.
- Forwarding: Forwards received packets and learns the source MAC address.

For a stable network topology, only the root port and designated port can enter the forwarding state while other ports are always in discarding state.

Hop Count

Internal spanning trees (ISTs) and multiple spanning tree instances (MSTIs) calculate whether the BPDU packet time expires based on an IP TTL-alike mechanism Hop Count, instead of Message Age and Max Age.

It is recommended to run the **spanning-tree max-hops** command in global configuration mode to configure the hop count. In a region, every time a BPDU packet passes through a device from the root bridge, the hop count decreases by 1. When the hop count becomes 0, the BPDU packet time expires and the device discards the packet.

To be compatible with STP and RSTP outside the region, MSTP also retains the Message Age and Max Age mechanisms.

Overview

Feature	Description
STP	STP, defined by the IEEE in 802.1D, is used to eliminate physical loops at the data link layer in a LAN.
RSTP	RSTP, defined by the IEEE in 802.1w, is optimized based on STP to rapidly converge the network topology.
MSTP	MSTP, defined by the IEEE in 802.1s, resolves defects of STP, RSTP, and Per-VLAN Spanning Tree (PVST). It cannot only rapidly converge but also can forward traffic of different VLANs along respective paths, thereby providing a better load balancing mechanism for redundant links.
MSTP Optical Features	MSTP includes the following features: PortFast, BPDU guard, BPDU filter, TC protection, TC guard, TC filter, BPDU check based on the source MAC address, BPDU filter based on the illegal length, Auto Edge, root guard, and loop guard.

10.3.1 STP

STP is used to prevent broadcast storms incurred by loops and provide link redundancy.

Working Principle

For the Layer-2 Ethernet, only one active link can exist between two LANs. Otherwise, a broadcast storm will occur. To enhance the reliability of a LAN, it is necessary to establish a redundant link and keep some paths in backup state. If the network is faulty and a link fails, you must switch the redundant link to the active state. STP can automatically activate the redundant link without any manual operations. STP enables devices on a LAN to:

- Discover and start the best tree topology on the LAN.
- Troubleshoot a fault and automatically update the network topology so that the possible best tree topology is always selected.

The LAN topology is automatically calculated based on a set of bridge parameters configured by the administrator. The best topology tree can be obtained by properly configuring these parameters.

10.3.2 RSTP

RSTP is completely compatible with 802.1D STP. Similar to traditional STP, RSTP provides loop-free and redundancy services. It is characterized by rapid speed. If all bridges in a LAN support RSTP and are properly configured by the administrator, it takes less than 1 second (about 50 seconds if traditional STP is used) to re-generate a topology tree after the network topology changes.

Working Principle

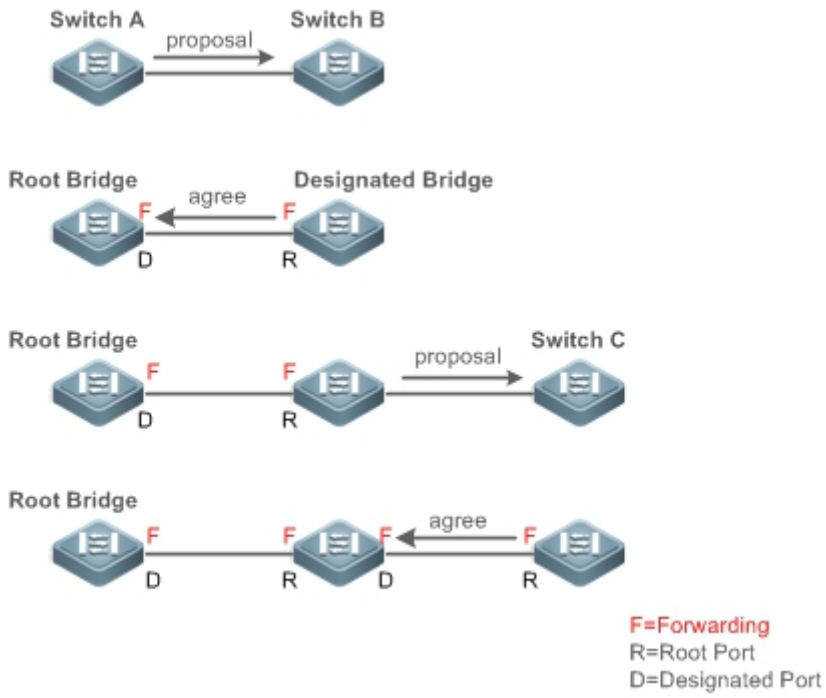
Fast RSTP Convergence

RSTP has a special feature, that is, to make ports quickly enter the forwarding state.

STP enables a port to enter the forwarding state 30 seconds (two times of the Forward-Delay Time; the Forward-Delay Time can be configured, with a default value of 15 seconds) after selecting a port role. Every time the topology changes, the root port and designated port reselected by each bridge enter the forwarding state 30 seconds later. Therefore, it takes about 50 seconds for the entire network topology to become a tree.

RSTP differs greatly from STP in the forwarding process. As shown in Figure 9-6, Switch A sends an RSTP Proposal packet to Switch B. If Switch B finds the priority of Switch A higher, it selects Switch A as the root bridge and the port receiving the packet as the root port, enters the forwarding state, and then sends an Agree packet from the root port to Switch A. If the designated port of Switch A is agreed, the port enters the forwarding state. Switch B's designated port resends a Proposal packet to extend the spanning tree by sequence. Theoretically, RSTP can recover the network tree topology to rapidly converge once the network topology changes.

Figure 9-6



The above handshake process is implemented only when the connection between ports is in point-to-point mode. To give the devices their full play, it is recommended not to enable point-to-point connection between devices.

Figure 9-7 and Figure 9-8 show the examples of non point-to-point connection.

Example of non point-to-point connection:

Figure 9-7

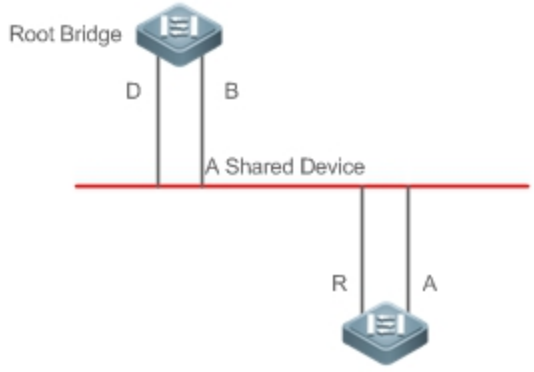


Figure 9-8

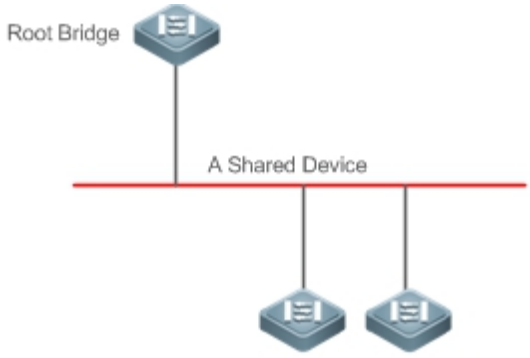
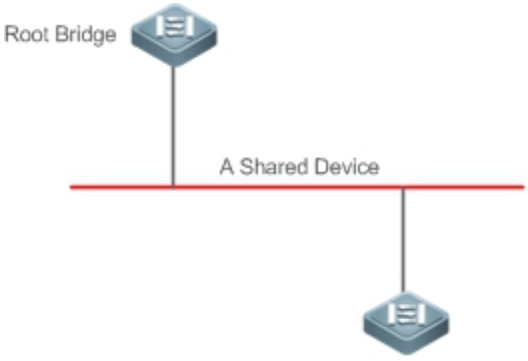


Figure 9-9 shows an example of point-to-point connection.

Figure 9-9



Compatibility Between RSTP and STP

RSTP is completely compatible with STP. RSTP automatically checks whether the connected bridge supports STP or RSTP based on the received BPDU version number. If the port connects to an STP bridge, the port enters the forwarding state 30 seconds later, which cannot give RSTP its full play.

Another problem may occur when RSTP and STP are used together. As shown in the following figures, Switch A (RSTP) connects to Switch B (STP). If Switch A finds itself connected to an STP bridge, it sends an STP BPDU packet. However, if Switch B is replaced with Switch C (RSTP) but Switch A still sends STP BPDU packets, Switch C will assume itself connected to the STP bridge. As a result, two RSTP devices work under STP, greatly reducing the efficiency.

RSTP provides the protocol migration feature to forcibly send RSTP BPDU packets (the peer bridge must support RSTP). In this case, Switch A is enforced to send an RSTP BPDU and Switch C then finds itself connected to the RSTP bridge. As a result, two RSTP devices work under RSTP, as shown in Figure 9-11.

Figure 9-10

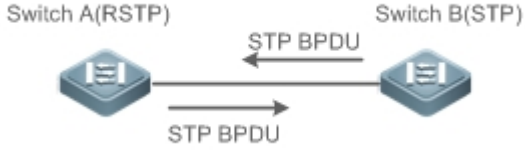


Figure 9-11

10.3.3 MSTP

MSTP resolves defects of STP and RSTP. It cannot only rapidly converge but also can forward traffic of different VLANs along respective paths, thereby providing a better load balancing mechanism for redundant links.

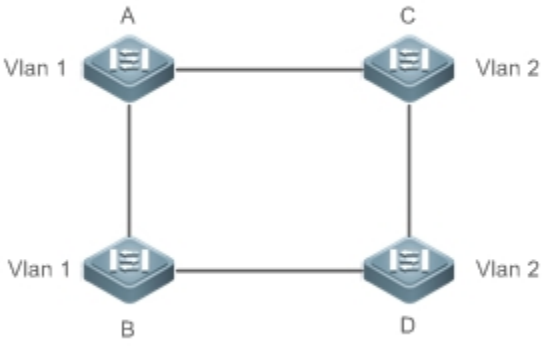
Working Principle

FS devices support MSTP. MSTP is a new spanning tree protocol developed from traditional STP and RSTP and includes the fast RSTP forwarding mechanism.

Since traditional spanning tree protocols are irrelevant to VLANs, problems may occur in specific network topologies:

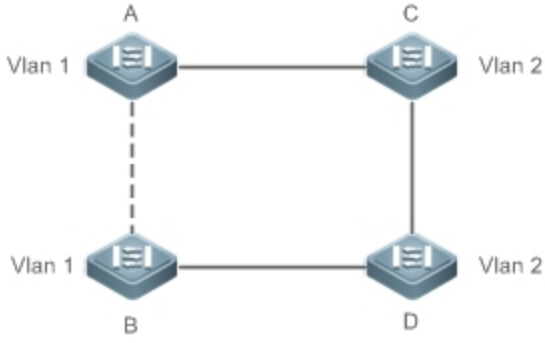
As shown in Figure 9-12, Devices A and B are in VLAN 1 while Devices C and D are in VLAN 2, forming a loop.

Figure 9-12



If the link from Device A to Device B through Devices C and D costs less than the link from Device A direct to Device B, the link between Device A and Device B enters the discarding state (as shown in Figure 9-13). Since Devices C and D do not include VLAN 1 and cannot forward data packets of VLAN 1, VLAN 1 of Device A fails to communicate with VLAN 1 of Device B.

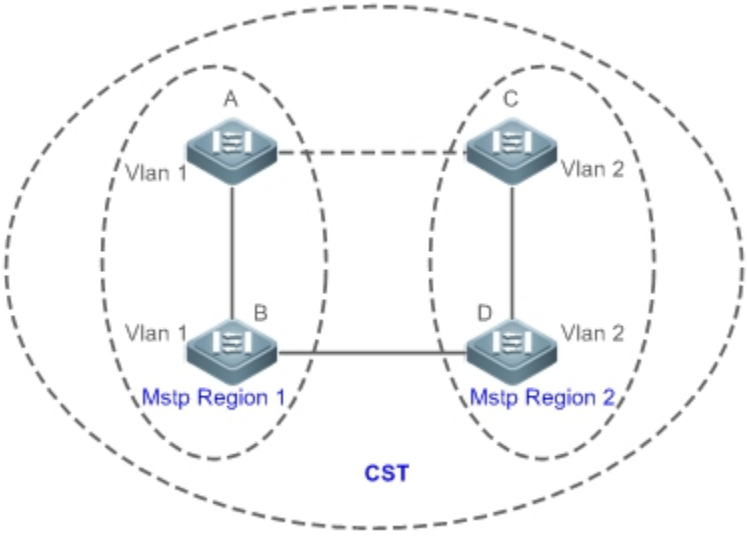
Figure 9-13



MSTP is developed to resolve this problem. It divides one or multiple VLANs of a device into an instance. Devices configured with the same instance form an MST region to run an independent spanning tree (called IST). This MST region, like a big device, implements the spanning tree algorithm with other MST regions to generate a complete spanning tree called common spanning tree (CST).

Based on this algorithm, the above network can form the topology shown in Figure 9-14 under the MSTP algorithm: Devices A and B are in MSTP region 1 in which no loop occurs, and therefore no link enters the discarding state. This also applies to MSTP Region 2. Region 1 and Region 2, like two big devices having loops, select a link to enter the discarding state based on related configuration.

Figure 9-14



This prevents loops to ensure proper communication between devices in the same VLAN.

MSTP Region Division

To give MSTP its due play, properly divide MSTP regions and configure the same MST configuration information for devices in the same MSTP region.

MST configuration information include:

- MST configuration name: Consists of at most 32 bytes to identify an MSTP region.
- MST Revision Number: Consists of 16 bits to identify an MSTP region.
- MST instance-VLAN mapping table: A maximum number of 64 instances (with their IDs ranging from 1 to 64) are created for each device and Instance 0 exists mandatorily. Therefore, the system supports a maximum number of 65 instances. Users can assign 1 to 4,994 VLANs belonging to different instances (ranging from 0 to 64) as required. Unassigned VLANs belong to Instance 0 by default. In this case, each MSTI is a VLAN group and implements the spanning tree algorithm of the MSTI specified in the BPDU packet, not affected by CIST and other MSTIs.

Run the **spanning-tree mst configuration** command in global configuration mode to enter the MST configuration mode to configure the above information.

MSTP BPDUs carry the above information. If the BPDU received by a device carries the same MST configuration information with the information on the device, it regards that the connected device belongs to the same MST region with itself. Otherwise, it regards the connected device originated from another MST region.

It is recommended to configure the instance-VLAN mapping table after disabling MSTP. After the configuration, re-enable MSTP to ensure stability and convergence of the network topology.

IST (Spanning Tree in an MSTP Region)

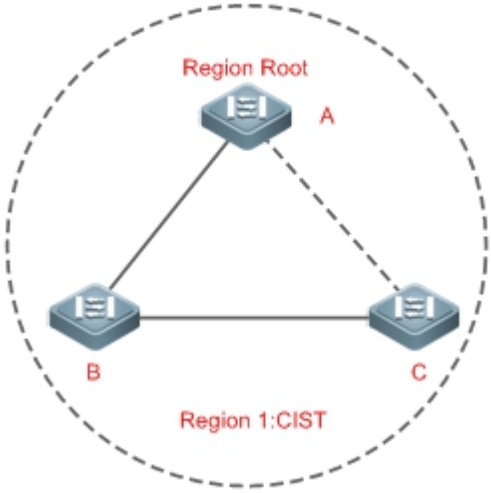
After MSTP regions are divided, each region selects an independent root bridge for each instance based on the corresponding parameters such as bridge priority and port priority, assigns roles to each port on each device, and specifies whether the port is in forwarding or discarding state in the instance based on the port role.

Through MSTP BPDU exchange, an IST is generated and each instance has their own spanning trees (MSTIs), in which the spanning tree corresponding to Instance 0 and CST are uniformly called Common Instance Spanning Tree (CIST). That is, each instance provides a single and loop-free network topology for their own VLAN groups.

As shown in Figure 9-15, Devices A, B, and C form a loop in Region 1.

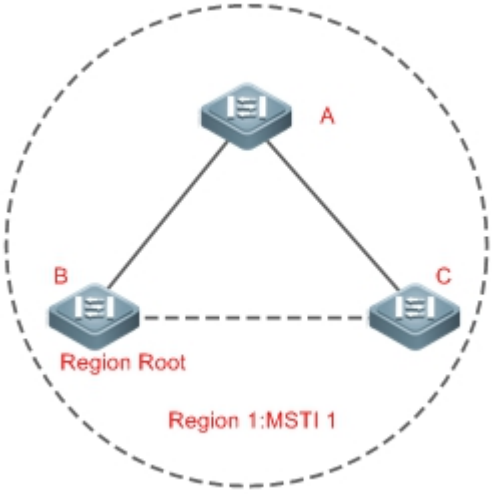
As shown in Figure 9-15, Device A has the highest priority in the CIST (Instance 0) and thereby is selected as the region root. Then MSTP enables the link between A and C to enter the discarding state based on other parameters. Therefore, for the VLAN group of Instance 0, only links from A to B and from B to C are available, interrupting the loop of this VLAN group.

Figure 9-15



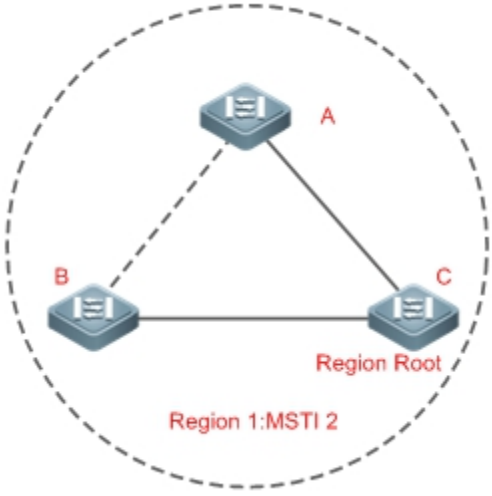
As shown in Figure 9-16, Device B has the highest priority in the MSTI 1 (Instance 1) and thereby is selected as the region root. Then MSTP enables the link between B and C to enter the discarding state based on other parameters. Therefore, for the VLAN group of Instance 1, only links from A to B and from A to C are available, interrupting the loop of this VLAN group.

Figure 9-16



As shown in Figure 9-17, Device C has the highest priority in the MSTI 2 (Instance 2) and thereby is selected as the region root. Then MSTP enables the link between B and C to enter the discarding state based on other parameters. Therefore, for the VLAN group of Instance 2, only links from B to C and from A to C are available, interrupting the loop of this VLAN group.

Figure 9-17

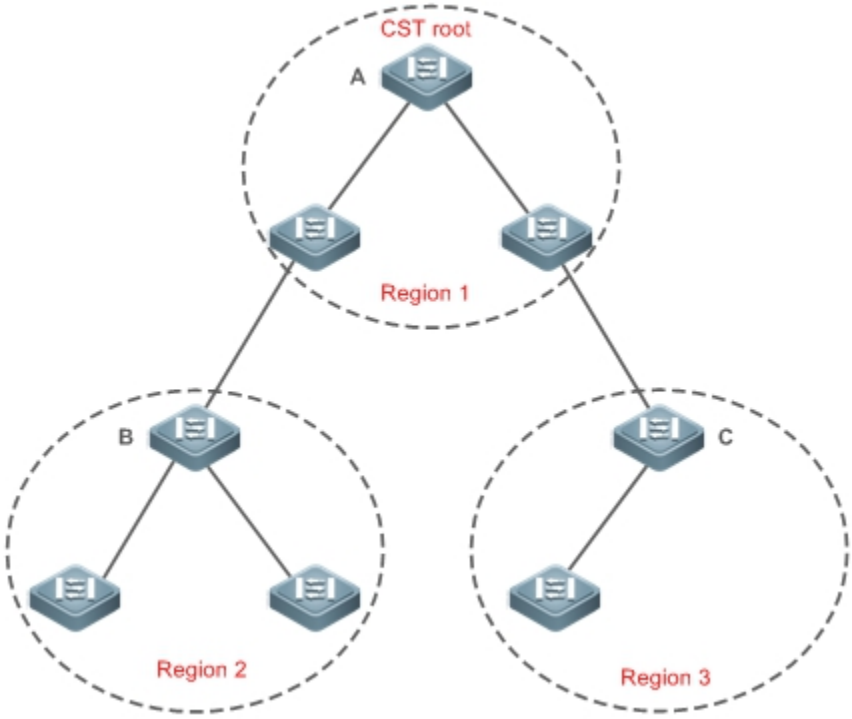


Note that MSTP does not care which VLAN a port belongs to. Therefore, users should configure the path cost and priority of a related port based on the actual VLAN configuration to prevent MSTP from interrupting wrong loops.

↘ CST (Spanning Tree Between MSTP Regions)

Each MSTP region is like a big device for the CST. Different MSTP regions form a bit network topology tree called CST. As shown in Figure 9-18, Device A, of which the bridge ID is the smallest, is selected as the root in the entire CST and the CIST regional root in this region. In Region 2, since the root path cost from Device B to the CST root is lowest, Device B is selected as the CIST regional root in this region. For the same reason, Device C is selected as the CIST regional root.

Figure 9-18



The CIST regional root may not be the device of which the bridge ID is the smallest in the region but indicates the device of which the root path cost from this region to the CST root is the smallest.

For the MSTI, the root port of the CIST regional root has a new role "master port". The master port acts as the outbound port of all instances and is in forwarding state for all instances. To make the topology more stable, we suggest that the master port of each region to the CST root be on the same device of the region if possible.

Compatibility Among MSTP, RSTP, and STP

Similar to RSTP, MSTP sends STP BPDUs to be compatible with STP. For details, see "Compatibility Between RSTP and STP".

Since RSTP processes MSTP BPDUs of the CIST, MSTP does not need to send RSTP BPDUs to be compatible with it.

Each STP or RSTP device is a single region and does not form the same region with any devices.

10.3.4 MSTP Optional Features

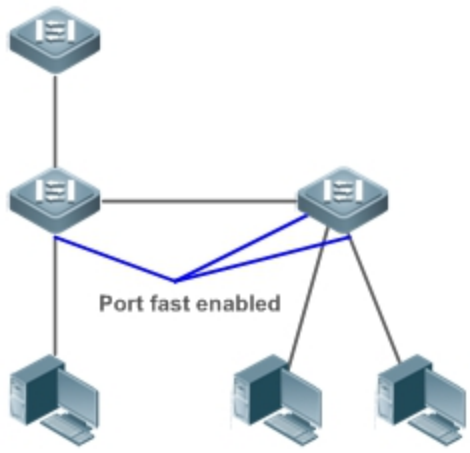
MSTP optional features mainly include PortFast port, BPDU guard, BPDU filter, TC guard, and guard. The optional features are mainly used to deploy MSTP configurations based on the network topology and application characteristics in the MSTP network. This enhances the stability, robustness, and anti-attack capability of MSTP, meeting application requirements of MSTP in different customer scenarios.

Working Principle

PortFast

If a port of a device connects directly to the network terminal, this port is configured as a PortFast port to directly enter the forwarding state. If the PortFast port is not configured, the port needs to wait for 30 seconds to enter the forwarding state. Figure 9-19 shows which ports of a device can be configured as PortFast ports.

Figure 9-19



If a PortFast port still receives BPDUs, its Port Fast Operational State is Disabled and the port enters the forwarding state according to the normal STP algorithm.

BPDU Guard

BPDU guard can be enabled globally or enabled on an interface.

It is recommended to run the **spanning-tree portfast bpduguard default** command in global configuration mode to enable global BPDU guard. If PortFast is enabled on a port or this port is automatically identified as an edge port, this port enters the error-disabled state to indicate the configuration error immediately after receiving a BPDU. At the same time, the port is disabled, indicating that a network device may be added by an unauthorized user to change the network topology.

It is also recommended to run the **spanning-tree bpduguard enable** command in interface configuration mode to enable BPDU guard on a port (whether PortFast is enabled or not on the port). In this case, the port enters the error-disabled state immediately after receiving a BPDU.

BPDU Filter

BPDU filter can be enabled globally or enabled on an interface.

It is recommended to run the **spanning-tree portfast bpdufilter default** command in global configuration mode to enable global BPDU filter. In this case, the PortFast port neither receives nor sends BPDUs and therefore the host connecting directly to the PortFast port receives no BPDUs. If the port changes its Port Fast Operational State to Disabled after receiving a BPDU, BPDU filter automatically loses effect.

It is also recommended to run the **spanning-tree bpdufilter enable** command in interface configuration mode to enable BPDU filter on a port (whether PortFast is enabled or not on the port). In this case, the port neither receives nor sends BPDUs but directly enters the forwarding state.

TC Protection

TC BPDUs are BPDU packets carrying the TC. If a switch receives such packets, it indicates the network topology changes and the switch will delete the MAC address table. For Layer-3 switches in this case, the forwarding module is re-enabled and the port status in the ARP entry changes. When a switch is attacked by forged TC BPDUs, it will frequently perform the above operations, causing heavy load and affecting network stability. To prevent this problem, you can enable TC protection.

TC protection can only be globally enabled or disabled. This function is disabled by default.

When TC protection is enabled, the switch deletes TC BPDUs within a specified period (generally 4 seconds) after receiving them and

monitors whether any TC BPDU packet is received during the period. If a device receives TC BPDU packets during this period, it deletes them when the period expires. This can prevent the device from frequently deleting MAC address entries and ARP entries.

TC Guard

TC protection ensures less dynamic MAC addresses and ARP entries removed when a large number of TC packets are generated on the network. However, a device receiving TC attack packets still performs many removal operations and TC packets can be spread, affecting the entire network. Users can enable TC guard to prevent TC packets from spreading globally or on a port. If TC guard is enabled globally or on a port, a port receiving TC packets filters these TC packets or TC packets generated by itself so that TC packets will not be spread to other ports. This can effectively control possible TC attacks in the network to ensure network stability. Particularly on Layer-3 devices, this function can effectively prevent the access-layer device from flapping and interrupting the core route.

- If TC guard is used incorrectly, the communication between networks is interrupted.
- It is recommended to enable this function only when illegal TC attack packets are received in the network.
- If TC guard is enabled globally, no port spreads TC packets to others. This function can be enabled only on laptop access devices.
- If TC guard is enabled on a port, the topology changes incurred and TC packets received on the port will not be spread to other ports. This function can be enabled only on uplink ports, particularly on ports of the convergence core.

TC Filter

If TC guard is enabled on a port, the port does not forward TC packets received and generated by the port to other ports performing spanning tree calculation on the device. When the status of a port changes (for example, from blocking to forwarding), the port generates TC packets, indicating that the topology may have changed.

In this case, since TC guard prevents TC packets from spreading, the device may not clear the MAC addresses of the port when the network topology changes, causing a data forwarding error.

To resolve this problem, TC filter is introduced. TC filter does not process TC packets received by ports but processes TC packets in case of normal topology changes. If TC filter is enabled, the address removal problem will be avoided and the core route will not be interrupted when ports not enabled with PortFast frequently go up or down, and the core routing entries can be updated in a timely manner when the topology changes.

- TC filter is disabled by default.

BPDU Source MAC Address Check

BPDU source MAC address check prevents BPDU packets from maliciously attacking switches and causing MSTP abnormal. When the switch connected to a port on a point-to-point link is determined, you can enable BPDU source MAC address check to receive BPDU packets sent only by the peer switch and discard all other BPDU packets, thereby preventing malicious attacks. You can enable the BPDU source MAC address check in interface configuration mode for a specific port. One port can only filter one MAC address. If you run the **no bpdu src-mac-check** command to disable BPDU source MAC address check on a port, the port receives all BPDU packets.

BPDU Filter

If the Ethernet length of a BPDU exceeds 1,500, this BPDU will be discarded, preventing receipt of illegal BPDU packets.

Auto Edge

If the designated port of a device does not receive a BPDU from the downlink port within a specific period (3 seconds), the device regards a network device connected to the designated port, configures the port as an edge port, and switches the port directly into the forwarding state. The edge port will be automatically identified as a non-edge port after receiving a BPDU.

You can run the **spanning-tree autoedge disabled** command to disable Auto Edge.

This function is enabled by default.

- If Auto Edge conflicts with the manually configured PortFast, the manual configuration prevails.
- Since this function is used for rapid negotiation and forwarding between the designated port and the downlink port, STP does not support this function. If the designated port is in forwarding state, the Auto Edge configuration does not take effect on this port. It takes only when rapid negotiation is re-performed, for example, when the network cable is removed and plugged.
- If BPDU filter has been enabled on a port, the port directly enters the forwarding state and is not automatically identified as an edge port.
- This function applies only to the designated port.

Root Guard

In the network design, the root bridge and backup root bridge are usually divided into the same region. Due to incorrect configuration of maintenance personnel or malicious attacks in the network, the root bridge may receive configuration information with a higher priority and thereby switches to the backup root bridge, causing incorrect changes in the network topology. Root guard is to resolve this problem.

If root guard is enabled on a port, its roles on all instances are enforced as the designated port. Once the port receives configuration information with a higher priority, it enters the root-inconsistent (blocking) state. If the port does not receive configuration information with a higher priority within a period, it returns to its original state.

If a port enters the blocking state due to root guard, you can manually restore the port to the normal state by disabling root guard on this port or disabling spanning tree guard (running **spanning-tree guard none** in interface configuration mode).

- If root guard is used incorrectly, the network link will be interrupted.
- If root guard is enabled on a non-designated port, this port will be enforced as a designated port and enter the BKN state. This indicates that the port enters the blocking state due to root inconsistency.
- If a port enters the BKN state due to receipt of configuration information with a higher priority in MST0, this port will be enforced in the BKN state in all other instances.
- Root guard and loop guard cannot take effect on a port at the same time.

Loop Guard

Due to the unidirectional link failure, the root port or backup port becomes the designated port and enters the forwarding state if it does not receive BPDUs, causing a network loop. Loop guard is to prevent this problem.

If a port enabled with loop guard does not receive BPDUs, the port switches its role but stays in discarding state till it receives BPDUs and recalculates the spanning tree.

- You can enable loop guard globally or on a port.
- Root guard and loop guard cannot take effect on a port at the same time.
- Before MSTP is restarted on a port, the port enters the blocking state in loop guard. If the port still receives no BPDU after MSTP is restarted, the port will become a designated port and enter the forwarding state. Therefore, it is recommended to identify the cause why a port enters the blocking state in loop protection and rectify the fault as soon as possible before restarting MSTP. Otherwise, the spanning tree topology will still become abnormal after MSTP is restarted.

BPDU Transparent Transmission

In IEEE 802.1Q, the destination MAC address 01-80-C2-00-00-00 of the BPDU is used as a reserved address. That is, devices compliant with IEEE 802.1Q do not forward the BPDU packets received. However, devices may need to transparently transmit BPDU packets in actual network deployment. For example, if STP is disabled on a device, the device needs to transparently transmit BPDU packets so that the spanning tree between devices is properly calculated.

BPDU transparent transmission is disabled by default.

BPDU transparent transmission takes effect only when STP is disabled. If STP is enabled on a device, the device does not transparently transmit BPDU packets.

BPDU Tunnel

The QinQ network is generally divided into two parts: customer network and SP network. Before a user packet enters the SP network, it is encapsulated with the VLAN tag of an SP network and also retains the original VLAN tag as data. As a result, the packet carries two VLAN tags to pass through the SP network. In the SP network, packets are transmitted only based on the outer-layer VLAN tag. When packets leave the SP network, the outer-layer VLAN tag is removed.

The STP packet transparent transmission feature, namely BPDU Tunnel, can be used to realize the transmission of STP packets between the customer network without any impact on the SP network. If an STP packet sent from the customer network enters a PE, the PE changes the destination MAC address of the packet to a private address before the packet is forwarded by the SP network. When the packet reaches the PE at the peer end, the PE changes the destination MAC address to a public address and returns the packet to the customer network at the peer end, realizing transparent transmission across the SP network. In this case, STP on the customer network is calculated independently of that on the SP network.

10.4 Configuration

Configuration	Description and Command	
Enabling STP	(Mandatory) It is used to enable STP.	
	spanning-tree	Enables STP and configures basic attributes.
	spanning-tree mode	Configures the STP mode.
Configuring STP Compatibility	(Optional) It is used to be compatible with competitor devices.	
	spanning-tree compatible enable	Enables the compatibility mode of a port.
	clear spanning-tree detected-protocols	Performs mandatory version check for BPDUs.
Configuring an MSTP Region	(Optional) It is used to configure an MSTP region.	
	spanning-tree mst configuration	Enters the MST configuration mode.
Enabling Fast RSTP Convergence	(Optional) It is used to configure whether the link type of a port is point-to-point connection.	
	spanning-tree link-type	Configures the link type.
Configuring Priorities	(Optional) It is used to configure the switch priority or port priority.	
	spanning-tree priority	Configures the switch priority.
	spanning-tree port-priority	Configures the port priority.
Configuring the Port Path Cost	(Optional) It is used to configure the path cost of a port or the default path cost calculation method.	

Configuration	Description and Command	
	spanning-tree cost	Configures the port path cost.
	spanning-tree pathcost method	Configures the default path cost calculation method.
Configuring the Maximum Hop Count of a BPDU Packet	(Optional) It is used to configure the maximum hop count of a BPDU packet.	
	spanning-tree max-hops	Configures the maximum hop count of a BPDU packet.
Enabling PortFast-related Features	(Optional) It is used to enable PortFast-related features.	
	spanning-tree portfast	Enables PortFast.
	spanning-tree portfast bpduguard default	Enables BPDU guard on all ports.
	spanning-tree bpduguard enabled	Enables BPDU guard on a port.
	spanning-tree portfast bpdufilter default	Enables BPDU filter on all ports.
	spanning-tree bpdufilter enabled	Enables BPDU filter on a port.
Enabling TC-related Features	(Optional) It is used to enable TC-related features.	
	spanning-tree tc-protection	Enables TC protection.
	spanning-tree tc-protection tc-guard	Enables TC guard on all ports.
	spanning-tree tc-guard	Enables TC guard on a port.
	spanning-tree ignore tc	Enables TC filter on a port.
Enabling BPDU Source MAC Address Check	(Optional) It is used to enable BPDU source MAC address check.	
	bpdu src-mac-check	Enables BPDU source MAC address check on a port.
Configuring Auto Edge	(Optional) It is used to configure Auto Edge.	
	spanning-tree autoedge	Enables Auto Edge on a port. This function is enabled by default.
Enabling Guard-related Features	(Optional) It is used to enable port guard features.	
	spanning-tree guard root	Enables root guard on a port.
	spanning-tree loopguard default	Enables loop guard on all ports.
	spanning-tree guard loop	Enables loop guard on a port.
	spanning-tree guard none	Disables the guard feature on a port.
Enabling BPDU Transparent Transmission	(Optional) It is used to enable BPDU transparent transmission	
	bridge-frame forwarding protocol bpdu	Enables BPDU transparent transmission.
Enabling BPDU Tunnel	(Optional) It is used to enable BPDU Tunnel.	
	l2protocol-tunnel stp	Enables BPDU Tunnel globally.
	l2protocol-tunnel stp enable	Enables BPDU Tunnel on a port.
	l2protocol-tunnel stp tunnel-dmac	Configures the transparent transmission address of BPDU Tunnel.

10.4.1 Enabling STP

Configuration Effect

- Enable STP globally and configure the basic attributes.
- Configure the STP mode.

Notes

- STP is disabled by default. Once STP is enabled, the device starts to run STP. The device runs MSTP by default.
- The default STP mode is MSTP mode.
- STP and Transparent Interconnection of Lots of Links (TRILL) of the data center cannot be enabled at the same time.

Configuration Steps

Enabling STP

- Mandatory.
- Unless otherwise specified, enable STP on each device.
- Run the **spanning-tree** [**forward-time** *seconds* | **hello-time** *seconds* | **max-age** *seconds*] command to enable STP and configure basic attributes.
- The forward-time ranges from 4 to 30. The hello-time ranges from 1 to 10. The max-age ranges from 6 to 40.

Running the **clear** commands may lose vital information and thus interrupt services. The value ranges of forward-time, hello-time, and max-age are related. If one of them is modified, the other two ranges are affected. The three values must meet the following condition: 2 x (Hello Time + 1 second) <= Max-Age Time <= 2 x (Forward-Delay Time –1 second). Otherwise, the configuration will fail.

Command	**spanning-tree** [**forward-time** *seconds* | **hello-time** *seconds* | **max-age** *seconds* | **tx-hold-count** *numbers*]
Parameter Description	**forward-time** *seconds*: Indicates the interval when the port status changes. The value ranges from 4 to 30 seconds. The default value is 15 seconds. **hello-time** *seconds*: Indicates the interval when a device sends a BPDU packet. The value ranges from 1 to 10 seconds. The default value is 2 seconds. **max-age** *second*: Indicates the longest TTL of a BPDU packet. The value ranges from 6 to 40 seconds. The default value is 20 seconds. **tx-hold-count** *numbers*: Indicates the maximum number of BPDUs sent per second. The value ranges from 1 to 10. The default value is 3.
Defaults	STP is disabled by default.
Command Mode	Global configuration mode
Usage Guide	The value ranges of forward-time, hello-time, and max-age are related. If one of them is modified, the other two ranges are affected. The three values must meet the following condition: 2 x (Hello Time + 1 second) <= Max-Age Time <= 2 x (Forward-Delay Time – 1 second) Otherwise, the topology may become unstable and the configuration will fail.

Configuring the STP Mode

- Optional.

- According to related 802.1 protocol standards, STP, RSTP, and MSTP are mutually compatible, without any configuration by the administrator. However, some vendors' devices do not work according to 802.1 protocol standards, possibly causing incompatibility. Therefore, FS provides a command for the administrator to switch the STP mode to a lower version if other vendors' devices are incompatible with FS devices.
- Run the **spanning-tree mode** [**stp | rstp | mstp**] command to modify the STP mode.

Command	**spanning-tree mode** [**stp** | **rstp** | **mstp**]
Parameter Description	**stp**: Spanning Tree Protocol (IEEE 802.1d) **rstp**: Rapid Spanning Tree Protocol (IEEE 802.1w) **mstp**: Multiple Spanning Tree Protocol (IEEE 802.1s)
Defaults	The default value is **mstp**.
Command Mode	Global configuration mode
Usage Guide	However, some vendors' devices do not work according to 802.1 protocol standards, possibly causing incompatibility. If other vendors' devices are incompatible with FS devices, run this command to switch the STP mode to a lower version.

Verification

- Display the configuration.

Configuration Example

Enabling STP and Configuring Timer Parameters

Scenario **Figure 9-20**	
Configuration Steps	- Enable STP and set the STP mode to STP on the devices. - Configure the timer parameters of root bridge DEV A as follows: Hello Time=4s, Max Age=25s, Forward Delay=18s.
DEV A	Step 1: Enable STP and set the STP mode to STP. FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#spanning-tree FS(config)#spanning-tree mode stp Step 2: Configure the timer parameters of root bridge DEV A.

	FS(config)#spanning-tree hello-time 4 FS(config)#spanning-tree max-age 25 FS(config)#spanning-tree forward-time 18
DEV B	Enable STP and set the STP mode to STP. FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#spanning-tree FS(config)#spanning-tree mode stp
Verification	● Run the **show spanning-tree summary** command to display the spanning tree topology and protocol configuration parameters.

DEV A	FS#show spanning-tree summary
	
	Spanning tree enabled protocol stp
	  Root ID Priority 0
	  Address 00d0.f822.3344
	  this bridge is root
	  Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec
	
	  Bridge ID Priority 0
	  Address 00d0.f822.3344
	  Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec
	
	Interface Role Sts Cost Prio OperEdge Type
	---------- ---- --- ---------- -------- -------- ----------
	Gi0/2 Desg FWD 20000 128 False P2p
	Gi0/1 Desg FWD 20000 128 False P2p

DEV B	FS#show spanning-tree summary
	
	Spanning tree enabled protocol stp
	  Root ID Priority 0
	  Address 00d0.f822.3344
	  this bridge is root
	  Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec

```
  Bridge ID    Priority       32768
                 Address        001a.a917.78cc
                 Hello Time     2 sec    Forward Delay 15 sec    Max Age 20 sec

Interface            Role Sts Cost          Prio        OperEdge Type
---------------- ---- --- ---------- -------- -------- ----------------
Gi0/2                Altn BLK 20000         128         False       P2p Bound(STP)
Gi0/1                Root FWD 20000         128         False       P2p Bound(STP)
```

Common Errors

N/A

10.4.2 Configuring STP Compatibility

Configuration Effect

- Enable the compatibility mode of a port to realize interconnection between FS devices and other SPs' devices.
- Enable protocol migration to perform forcible version check to affect the compatibility between RSTP and STP.

Notes

- If the compatibility mode is enabled on a port, this port will add different MSTI information into the to-be-sent BPDU based on the current port to realize interconnection between FS devices and other SPs' devices.

Configuration Steps

Enabling the Compatibility Mode on a Port

- Optional.

Command	**spanning-tree compatible enable**
Parameter Description	N/A
Defaults	The compatibility mode is disabled on a port by default.
Command Mode	Interface configuration mode
Usage Guide	If the compatibility mode is enabled on a port, this port will add different MSTI information into the to-be-sent BPDU based on the current port to realize interconnection between FS devices and other SPs' devices.

Enabling Protocol Migration

- Optional.
- If the peer device supports RSTP, you can enforce version check on the local device to force the two devices to run RSTP.
- Run the **clear spanning-tree detected-protocols** [**interface** *interface-id*] command to enforce version check on a port. For details,

see "Compatibility Between RSTP and STP".

Command	**clear spanning-tree detected-protocols** [**interface** *interface-id*]
Parameter Description	**interface** *interface-id*: Indicates a port.
Defaults	N/A
Command Mode	Privileged EXEC mode
Usage Guide	This command is used to enforce a port to send RSTP BPDU packets and perform forcible check on them.

Verification

- Display the configuration.

Configuration Example

Enabling STP Compatibility

Scenario Figure 9-21	
Configuration Steps	• Configure Instances 1 and 2 on Devices A and B, and map Instance 1 with VLAN 10 and Instance 2 with VLAN 20. • Configure Gi0/1 and Gi0/2 to respectively belong to VLAN 10 and VLAN 20, and enable STP compatibility.
DEV A	Step 1: Configure Instances 1 and 2, and map Instances 1 and 2 respectively with VLANs 10 and 20. FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#spanning-tree mst configuration FS(config-mst)#instance 1 vlan 10 FS(config-mst)#instance 2 vlan 20 Step 2: Configure the VLAN the port belongs to, and enable STP compatibility on the port. FS(config)#int gi 0/1 FS(config-if-GigabitEthernet 0/1)#switchport access vlan 10 FS(config-if-GigabitEthernet 0/1)#spanning-tree compatible enable FS(config-if-GigabitEthernet 0/1)#int gi 0/2 FS(config-if-GigabitEthernet 0/2)#switchport access vlan 20

	FS(config-if-GigabitEthernet 0/2)#spanning-tree compatible enable
DEV B	Perform the same steps as DEV A.
Verification	● Run the **show spanning-tree summary** command to check whether the spanning tree topology is correctly calculated.
DEV A	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : 1-9, 11-19, 21-4094 Root ID Priority 32768 Address 001a.a917.78cc this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Bridge ID Priority 32768 Address 001a.a917.78cc Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Desg FWD 20000 128 False P2p Gi0/1 Desg FWD 20000 128 False P2p MST 1 vlans map : 10 Region Root Priority 32768 Address 001a.a917.78cc this bridge is region root Bridge ID Priority 32768 Address 001a.a917.78cc Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/1 Desg FWD 20000 128 False P2p

	MST 2 vlans map : 20 Region Root Priority 32768 Address 001a.a917.78cc this bridge is region root Bridge ID Priority 32768 Address 001a.a917.78cc Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Desg FWD 20000 128 False P2p
DEV B	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : 1-9, 11-19, 21-4094 Root ID Priority 32768 Address 001a.a917.78cc this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Bridge ID Priority 32768 Address 00d0.f822.3344 Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Altn BLK 20000 128 False P2p Gi0/1 Root FWD 20000 128 False P2p MST 1 vlans map : 10 Region Root Priority 32768 Address 001a.a917.78cc this bridge is region root

```
  Bridge ID    Priority       32768
                 Address        00d0.f822.3344

Interface            Role Sts Cost          Prio        OperEdge Type
---------------- ---- --- ---------- -------- -------- ----------------
Gi0/1                Root FWD 20000         128         False       P2p

MST 2 vlans map : 20
  Region Root Priority      32768
                 Address        001a.a917.78cc
                 this bridge is region root

  Bridge ID    Priority       32768
                 Address        00d0.f822.3344

Interface            Role Sts Cost          Prio        OperEdge Type
---------------- ---- --- ---------- -------- -------- ----------------
Gi0/2                Root FWD 20000         128         False       P2p
```

Common Errors

N/A

10.4.3 Configuring an MSTP Region

Configuration Effect

- Configure an MSTP region to adjust which devices belong to the same MSTP region and thereby affect the network topology.

Notes

- To make multiple devices belong to the same MSTP region, configure the same name, revision number, and instance-VLAN mapping table for them.
- You can configure VLANs for Instances 0 to 64, and then the remaining VLANs are automatically allocated to Instance 0. One VLAN belongs to only one instance.
- It is recommended to configure the instance-VLAN mapping table after disabling STP. After the configuration, re-enable MSTP to ensure stability and convergence of the network topology.

Configuration Steps

↘ **Configuring an MSTP Region**

- Optional.
- Configure an MSTP region when multiple devices need to belong to the same MSTP region.
- Run the **spanning-tree mst configuration** command to enter the MST configuration mode.
- Run the **instance** *instance-id* **vlan** *vlan-range* command to configure the MSTI-VLAN mapping.
- Run the **name** *name* command to configure the MST name.
- Run the **revision** *version* command to configure the MST version number.

Command	**spanning-tree mst configuration**
Parameter Description	N/A
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	Run this command to enter the MST configuration mode.

Command	**instance** *instance-id* **vlan** *vlan-range*
Parameter Description	*instance-id*: Indicates the MSTI ID, ranging from 0 to 64. *vlan-range*: Indicates the VLAN ID, ranging from 1 to 4,094.
Defaults	The default instance-VLAN mapping is that all VLANs are in Instance 0.
Command Mode	MST configuration mode
Usage Guide	To add a VLAN group to an MSTI, run this command. For example, instance 1 vlan 2-200: Adds VLANs 2 to 200 to Instance 1. instance 1 vlan 2,20,200: Adds VLANs 2, 20, and 200 to Instance 1. You can use the **no** form of this command to remove VLANs from an instance. Removed VLANs are automatically forwarded to Instance 0.

Command	**name** *name*
Parameter Description	*name*: Indicates the MST name. It consists of a maximum of 32 bytes.
Defaults	The default name is an empty character string.
Command Mode	MST configuration mode
Usage Guide	N/A

Command	**revision** *version*
Parameter Description	*version*: Indicates the MST revision number, ranging from 0 to 65,535.
Defaults	The default revision number is 0.
Command	MST configuration mode

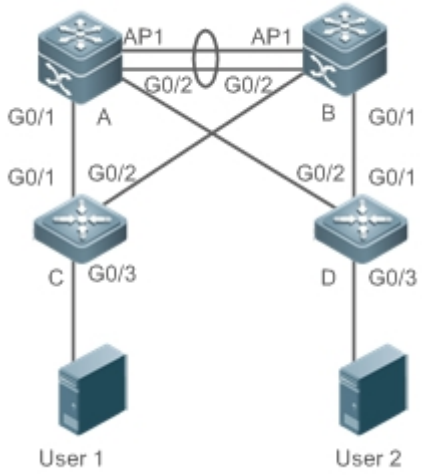
Mode	
Usage Guide	N/A

Verification

- Display the configuration.
- Run the **show spanning-tree mst configuration** command to display the MSTP region configuration.

Configuration Example

Enabling MSTP to Achieve VLAN Load Balancing in the MSTP+VRRP Topology

Scenario **Figure 9-22**	
Configuration Steps	- Enable MSTP and create Instances 1 and 2 on Switches A, B, C, and D. - Configure Switch A as the root bridge of Instances 0 and 1 and Switch B as the root bridge of Instance 2. - Configure Switch A as the VRRP master device of VLANs 1 and 10 and Switch B as the VRRP master device of VLAN 20.
A	Step 1: Configure VLANs 10 and 20, and configure ports as Trunk ports. A(config)#vlan 10 A(config-vlan)#vlan 20 A(config-vlan)#exit A(config)#int range gi 0/1-2 A(config-if-range)#switchport mode trunk A(config-if-range)#int ag 1 A(config-if-AggregatePort 1)# switchport mode trunk Step 2: Enable MSTP and create Instances 1 and 2. A(config)#spanning-tree A(config)# spanning-tree mst configuration

	A(config-mst)#instance 1 vlan 10 A(config-mst)#instance 2 vlan 20 A(config-mst)#exit Step 3: Configure Switch A as the root bridge of Instances 0 and 1. A(config)#spanning-tree mst 0 priority 4096 A(config)#spanning-tree mst 1 priority 4096 A(config)#spanning-tree mst 2 priority 8192 Step 4: Configure VRRP priorities to enable Switch A to act as the VRRP master device of VLAN 10, and configure the virtual gateway IP address of VRRP. A(config)#interface vlan 10 A(config-if-VLAN 10)ip address 192.168.10.2 255.255.255.0 A(config-if-VLAN 10) vrrp 1 priority 120 A(config-if-VLAN 10) vrrp 1 ip 192.168.10.1 Step 5 Set the VRRP priority to the default value 100 to enable Switch A to act as the VRRP backup device of VLAN 20. A(config)#interface vlan 20 A(config-if-VLAN 20)ip address 192.168.20.2 255.255.255.0 A(config-if-VLAN 20) vrrp 1 ip 192.168.20.1
B	Step 1: Configure VLANs 10 and 20, and configure ports as Trunk ports. B(config)#vlan 10 B(config-vlan)#vlan 20 B(config-vlan)#exit B(config)#int range gi 0/1-2 B(config-if-range)#switchport mode trunk B(config-if-range)#int ag 1 B(config-if-AggregatePort 1)# switchport mode trunk Step 2: Enable MSTP and create Instances 1 and 2. B(config)#spanning-tree B(config)# spanning-tree mst configuration B(config-mst)#instance 1 vlan 10 B(config-mst)#instance 2 vlan 20

	B(config-mst)#exit Step 3: Configure Switch A as the root bridge of Instance 2. B(config)#spanning-tree mst 0 priority 8192 B(config)#spanning-tree mst 1 priority 8192 B(config)#spanning-tree mst 2 priority 4096 Step 4: Configure the virtual gateway IP address of VRRP. B(config)#interface vlan 10 B(config-if-VLAN 10)ip address 192.168.10.3 255.255.255.0 B(config-if-VLAN 10) vrrp 1 ip 192.168.10.1 Step 5 Set the VRRP priority to 120 to enable Switch B to act as the VRRP backup device of VLAN 20. B(config)#interface vlan 20 B(config-if-VLAN 20)vrrp 1 priority 120 B(config-if-VLAN 20)ip address 192.168.20.3 255.255.255.0 B(config-if-VLAN 20) vrrp 1 ip 192.168.20.1
C	Step 1: Configure VLANs 10 and 20, and configure ports as Trunk ports. C(config)#vlan 10 C(config-vlan)#vlan 20 C(config-vlan)#exit C(config)#int range gi 0/1-2 C(config-if-range)#switchport mode trunk Step 2: Enable MSTP and create Instances 1 and 2. C(config)#spanning-tree C(config)# spanning-tree mst configuration C(config-mst)#instance 1 vlan 10 C(config-mst)#instance 2 vlan 20 C(config-mst)#exit Step 3: Configure the port connecting Device C directly to users as a PortFast port and enable BPDU guard. C(config)#int gi 0/3 C(config-if-GigabitEthernet 0/3)#spanning-tree portfast

	C(config-if-GigabitEthernet 0/3)#spanning-tree bpduguard enable
D	Perform the same steps as Device C.
Verification	● Run the **show spanning-tree summary** command to check whether the spanning tree topology is correctly calculated. ● Run the **show vrrp brief** command to check whether the VRRP master/backup devices are successfully created.
A	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : 1-9, 11-19, 21-4094 Root ID Priority 4096 Address 00d0.f822.3344 this bridge is root Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec Bridge ID Priority 4096 Address 00d0.f822.3344 Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Ag1 Desg FWD 19000 128 False P2p Gi0/1 Desg FWD 200000 128 False P2p Gi0/2 Desg FWD 200000 128 False P2p MST 1 vlans map : 10 Region Root Priority 4096 Address 00d0.f822.3344 this bridge is region root Bridge ID Priority 4096 Address 00d0.f822.3344 Interface Role Sts Cost Prio OperEdge Type

	---------- ---- --- ---------- -------- -------- ---------- Ag1 Desg FWD 19000 128 False P2p Gi0/1 Desg FWD 200000 128 False P2p Gi0/2 Desg FWD 200000 128 False P2p MST 2 vlans map : 20 Region Root Priority 4096 Address 001a.a917.78cc this bridge is region root Bridge ID Priority 8192 Address 00d0.f822.3344 Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Ag1 Root FWD 19000 128 False P2p Gi0/1 Desg FWD 200000 128 False P2p Gi0/2 Desg FWD 200000 128 False P2p
B	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : 1-9, 11-19, 21-4094 Root ID Priority 4096 Address 00d0.f822.3344 this bridge is root Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec Bridge ID Priority 8192 Address 001a.a917.78cc Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Ag1 Root FWD 19000 128 False P2p

```
Gi0/1                Desg FWD 200000      128       False     P2p
Gi0/2                Desg FWD 200000      128       False     P2p

MST 1 vlans map : 10
  Region Root Priority    4096
                Address       00d0.f822.3344
                this bridge is region root

  Bridge ID   Priority      8192
                Address       001a.a917.78cc

Interface          Role Sts Cost          Prio      OperEdge Type
---------------- ---- --- ---------- -------- -------- ----------------
Ag1                  Root FWD 19000       128       False     P2p
Gi0/1                Desg FWD 200000      128       False     P2p
Gi0/2                Desg FWD 200000      128       False     P2p

MST 2 vlans map : 20
  Region Root Priority    4096
                Address       001a.a917.78cc
                this bridge is region root

  Bridge ID   Priority      4096
                Address       001a.a917.78cc

Interface          Role Sts Cost          Prio      OperEdge Type
---------------- ---- --- ---------- -------- -------- ----------------
Ag1                  Desg FWD 19000       128       False     P2p
Gi0/1                Desg FWD 200000      128       False     P2p
Gi0/2                Desg FWD 200000      128       False     P2p
```

C

```
FS#show spanning-tree summary

Spanning tree enabled protocol mstp
MST 0 vlans map : 1-9, 11-19, 21-4094
```

```
  Root ID    Priority    4096
             Address     00d0.f822.3344
             this bridge is root
             Hello Time  4 sec  Forward Delay 18 sec  Max Age 25 sec

  Bridge ID  Priority    32768
             Address     001a.a979.00ea
             Hello Time  2 sec  Forward Delay 15 sec  Max Age 20 sec

Interface        Role Sts Cost      Prio    Type    OperEdge
---------------- ---- --- ---------- -------- ----- ----------------
Fa0/2            Altn BLK 200000    128     P2p     False
Fa0/1            Root FWD 200000    128     P2p     False

MST 1 vlans map : 10
  Region Root Priority  4096
             Address     00d0.f822.3344
             this bridge is region root

  Bridge ID  Priority    32768
             Address     001a.a979.00ea

Interface        Role Sts Cost      Prio    Type    OperEdge
---------------- ---- --- ---------- -------- ----- ----------------
Fa0/2            Altn BLK 200000    128     P2p     False
Fa0/1            Root FWD 200000    128     P2p     False

MST 2 vlans map : 20
  Region Root Priority  4096
             Address     001a.a917.78cc
             this bridge is region root

  Bridge ID  Priority    32768
             Address     001a.a979.00ea
```

	Interface Role Sts Cost Prio Type OperEdge ---------- ---- --- ---------- -------- ----- ---------- Fa0/2 Root FWD 200000 128 P2p False Fa0/1 Altn BLK 200000 128 P2p False
D	Omitted.

Common Errors

- MST region configurations are inconsistent in the MSTP topology.
- VLANs are not created before you configure the mapping between the instance and VLAN.
- A device runs STP or RSTP in the MSTP+VRRP topology, but calculates the spanning tree according to the algorithms of different MST regions.

10.4.4 Enabling Fast RSTP Convergence

Configuration Effect

- Configure the link type to make RSTP rapidly converge.

Notes

- If the link type of a port is point-to-point connection, RSTP can rapidly converge. For details, see "Fast RSTP Convergence". If the link type is not configured, the device automatically sets the link type based on the duplex mode of the port. If a port is in full duplex mode, the device sets the link type to point-to-point. If a port is in half duplex mode, the device sets the link type to shared. You can also forcibly configure the link type to determine whether the port connection is point-to-point connection.

Configuration Steps

Configuring the Link Type

- Optional.

Command	**spanning-tree link-type** [**point-to-point** | **shared**]
Parameter Description	**point-to-point**: Forcibly configures the link type of a port to be point-to-point. **shared**: Forcibly configures the link type of a port to be shared.
Defaults	If a port is in full duplex mode, the link type of the port is point-to-point. If a port is in half duplex mode, the link type of the port is shared.
Command Mode	Interface configuration mode
Usage Guide	If the link type of a port is point-to-point connection, RSTP can rapidly converge. If the link type is not configured, the device automatically sets the link type based on the duplex mode of the port.

Verification

- Display the configuration.
- Run the **show spanning-tree** [**mst** *instance-id*] **interface** *interface-id* command to display the spanning tree configuration of the

port.

Configuration Example

Enabling Fast RSTP Convergence

Configuration Steps	Set the link type of a port to point-to-point.
	FS(config)#int gi 0/1 FS(config-if-GigabitEthernet 0/1)#spanning-tree link-type point-to-point
Verification	● Run the **show spanning-tree summary** command to display the link type of the port.
	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : ALL Root ID Priority 32768 Address 001a.a917.78cc this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Bridge ID Priority 32768 Address 00d0.f822.3344 Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/1 Root FWD 20000 128 False P2p

Common Errors

N/A

10.4.5 Configuring Priorities

Configuration Effect

- Configure the switch priority to determine a device as the root of the entire network and to determine the topology of the entire network.
- Configure the port priority to determine which port enters the forwarding state.

Notes

- It is recommended to set the priority of the core device higher (to a smaller value) to ensure stability of the entire network. You can assign different switch priorities to different instances so that each instance runs an independent STP based on the assigned priorities. Devices in different regions use the priority only of the CIST (Instance 0). As described in bridge ID, the switch priority has 16 optional values: 0, 4,096, 8,192, 12,288, 16,384, 20,480, 24,576, 28,672, 32,768, 36,864, 40,960, 45,056, 49,152, 53,248, 57,344, 61,440. They are integral multiples of 4,096. The default value is 32,768.
- If two ports are connected to a shared device, the device selects a port with a higher priority (smaller value) to enter the forwarding state and a port with a lower priority (larger value) to enter the discarding state. If the two ports have the same priority, the device selects the port with a smaller port ID to enter the forwarding state. You can assign different port priorities to different instances on a port so that each instance runs an independent STP based on the assigned priorities.
- Similar to the switch priority, the port priority also has 16 optional values: 0, 16, 32, 48, 64, 80, 96, 112, 128, 144, 160, 176, 192, 208, 224, 240. They are integral multiples of 16. The default value is 128.

Configuration Steps

Configuring the Switch Priority

- Optional.
- To change the root or topology of a network, configure the switch priority.

Command	**spanning-tree** [**mst** *instance-id*] **priority** *priority*
Parameter Description	**mst** *instance-id*: Indicates the instance ID, ranging from 0 to 64. **priority** *priority*: Indicates the switch priority. There are 16 optional values: 0, 4,096, 8,192, 12,288, 16,384, 20,480, 24,576, 28,672, 32,768, 36,864, 40,960, 45,056, 49,152, 53,248, 57,344, 61,440. They are integral multiples of 4,096.
Defaults	The default value of *instance-id* is 0 while that of *priority* is 32,768.
Command Mode	Global configuration mode
Usage Guide	Configure the switch priority to determine a device as the root of the entire network and to determine the topology of the entire network.

Configuring the Port Priority

- Optional.
- To change the preferred port entering the forwarding state, configure the port priority.

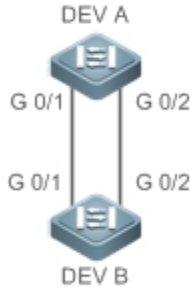
Command	spanning-tree [**mst** *instance-id*] **port-priority** *priority*
Parameter Description	**mst** *instance-id*: Indicates the instance ID, ranging from 0 to 64. **port-priority** *priority*: Indicates the port priority. There are 16 optional values: 0, 16, 32, 48, 64, 80, 96, 112, 128, 144, 160, 176, 192, 208, 224, 240. They are integral multiples of 4,096.
Defaults	The default value of *instance-id* is 0. The default value of *priority* is 128.
Command Mode	Interface configuration mode
Usage Guide	If a loop occurs in a region, the port with a higher priority is preferred to enter the forwarding state. If two ports have the same priority, the port with a smaller port ID is selected to enter the forwarding state. Run this command to determine which port in the loop of a region enters the forwarding state.

Verification

- Display the configuration.
- Run the **show spanning-tree** [**mst** *instance-id*] **interface** *interface-id* command to display the spanning tree configuration of the port.

Configuration Example

Configuring the Port Priority

Scenario **Figure 9-23**	
Configuration Steps	● Configure the bridge priority so that DEV A becomes the root bridge of the spanning tree. ● Configure the priority of Gi0/2 on DEV A is 16 so that Gi0/2 on DEV B can be selected as the root port.
DEV A	Step 1: Enable STP and configure the bridge priority. FS(config)#spanning-tree FS(config)#spanning-tree mst 0 priority 0 Step 2: Configure the priority of Gi 0/2. FS(config)# int gi 0/2 FS(config-if-GigabitEthernet 0/2)#spanning-tree mst 0 port-priority 16
DEV B	FS(config)#spanning-tree
Verification	● Run the **show spanning-tree summary** command to display the topology calculation result of the spanning tree.
DEV A	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : ALL Root ID Priority 0 Address 00d0.f822.3344 this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

	Bridge ID Priority 0 Address 00d0.f822.3344 Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Desg FWD 20000 16 False P2p Gi0/1 Desg FWD 20000 128 False P2p
DEV B	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : ALL Root ID Priority 0 Address 00d0.f822.3344 this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Bridge ID Priority 32768 Address 001a.a917.78cc Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Root FWD 20000 128 False P2p Gi0/1 Altn BLK 20000 128 False P2p

Common Errors

N/A

10.4.6 Configuring the Port Path Cost

Configuration Effect

- Configure the path cost of a port to determine the forwarding state of the port and the topology of the entire network.
- If the path cost of a port uses its default value, configure the path cost calculation method to affect the calculation result.

Notes

- A device selects a port as the root port if the path cost from this port to the root bridge is the lowest. Therefore, the port path cost determines the root port of the local device. The default port path cost is automatically calculated based on the port rate (Media Speed). A port with a higher rate will have a low path cost. Since this method can calculate the most scientific path cost, do not change the path cost unless required. You can assign different path costs to different instances on a port so that each instance runs an independent STP based on the assigned path costs.
- If the port path cost uses the default value, the device automatically calculates the port path cost based on the port rate. However, IEEE 802.1d-1998 and IEEE 802.1t define different path costs for the same link rate. The value is a short integer ranging from 1 to 65,535 in 802.1d-1998 while is a long integer ranging from 1 to 200,000,000 in IEEE 802.1t. The path cost of an aggregate port (AP) has two solutions: 1. FS solution: Port Path Cost x 95%; 2. Solution recommended in standards: 20,000,000,000/Actual link bandwidth of the AP, in which Actual link bandwidth of the AP = Bandwidth of a member port x Number of active member ports. The administrator must unify the path cost calculation method in the entire network. The default standard is the private long integer standard.
- The following table lists path costs automatically configured for different link rate in two solutions.

Port Rate	Port	IEEE 802.1d (short)	IEEE 802.1t (long)	IEEE 802.1t (long standard)
10M	Common port	100	2000000	2000000
	AP	95	1900000	2000000÷linkupcnt
100M	Common port	19	200000	200000
	AP	18	190000	200000÷linkupcnt
1000M	Common port	4	20000	20000
	AP	3	19000	20000÷linkupcnt
10000M	Common port	2	2000	2000
	AP	1	1900	20000÷linkupcnt

- FS's long integer standard is used by default. After the solution is changed to the path cost solution recommended by the standards, the path cost of an AP changes with the number of member ports in UP state. If the port path cost changes, the network topology also will change.
- If an AP is static, linkupcnt in the table is the number of active member ports. If an AP is an LACP AP, linkupcnt in the table is the number of member ports forwarding AP data. If no member port in the AP goes up, linkupcnt is 1. For details about AP and LACP, see the *Configuring AP*.

Configuration Steps

Configuring the Port Path Cost

- Optional.
- To determine which port or path data packets prefer to pass through, configure the port path cost.

Command	**spanning-tree** [**mst** *instance-id*] **cost** *cost*
Parameter Description	**mst** *instance-id*: Indicates the instance ID, ranging from 0 to 64. **cost** *cost*: Indicates the path cost, ranging from 1 to 200,000,000.
Defaults	The default value of *instance-id* is 0. The default value is automatically calculated based on the port rate. 1000 Mbps—20000

	100 Mbps—200000 10 Mbps—2000000
Command Mode	Interface configuration mode
Usage Guide	A larger value of *cost* indicates a higher path cost.

Configuring the Default Path Cost Calculation Method

- Optional.
- To change the path cost calculation method, configure the default path cost calculation method.

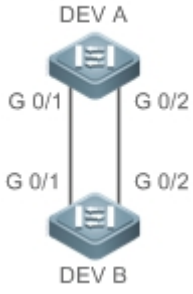
Command	**spanning-tree pathcost method** { *long* [*standard*] | *short* }
Parameter Description	*long*: Uses the path cost specified in 802.1t. *standard*: Uses the cost calculated according to the standard. *short*: Uses the path cost specified in 802.1d.
Defaults	The path cost specified in 802.1t is used by default.
Command Mode	Global configuration mode
Usage Guide	If the port path cost uses the default value, the device automatically calculates the port path cost based on the port rate.

Verification

- Display the configuration.
- Run the **show spanning-tree** [**mst** *instance-id*] **interface** *interface-id* command to display the spanning tree configuration of the port.

Configuration Example

Configuring the Port Path Cost

Scenario **Figure 9-24**	
Configuration Steps	● Configure the bridge priority so that DEV A becomes the root bridge of the spanning tree. ● Configure the path cost of Gi 0/2 on DEV B is 1 so that Gi 0/2 can be selected as the root port.
DEV A	FS(config)#spanning-tree FS(config)#spanning-tree mst 0 priority 0

DEV B	FS(config)#spanning-tree FS(config)# int gi 0/2 FS(config-if-GigabitEthernet 0/2)# spanning-tree cost 1
Verification	● Run the **show spanning-tree summary** command to display the topology calculation result of the spanning tree.
DEV A	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : ALL Root ID Priority 0 Address 00d0.f822.3344 this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Bridge ID Priority 0 Address 00d0.f822.3344 Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Desg FWD 20000 128 False P2p Gi0/1 Desg FWD 20000 128 False P2p
DEV B	FS#show spanning-tree summary Spanning tree enabled protocol mstp MST 0 vlans map : ALL Root ID Priority 0 Address 00d0.f822.3344 this bridge is root Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Bridge ID Priority 32768 Address 001a.a917.78cc

	Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec Interface Role Sts Cost Prio OperEdge Type ---------- ---- --- ---------- -------- -------- ---------- Gi0/2 Root FWD 1 128 False P2p Gi0/1 Altn BLK 20000 128 False P2p

Common Errors

- N/A

10.4.7 Configuring the Maximum Hop Count of a BPDU Packet

Configuration Effect

- Configure the maximum hop count of a BPDU packet to change the BPDU TTL and thereby affect the network topology.

Notes

- The default maximum hop count of a BPDU packet is 20. Generally, it is not recommended to change the default value.

Configuration Steps

Configuring the Maximum Hop Count

- (Optional) If the network topology is so large that a BPDU packet exceeds the default 20 hops, it is recommended to change the maximum hop count.

Command	**spanning-tree max-hops** *hop-count*
Parameter Description	*hop-count*: Indicates the number of devices a BPDU passes through before being discarded. It ranges from 1 to 40.
Defaults	The default value of *hop-count* is 20.
Command Mode	Global configuration mode
Usage Guide	In a region, the BPDU sent by the root bridge includes a hop count. Every time a BPDU passes through a device from the root bridge, the hop count decreases by 1. When the hop count becomes 0, the BPDU times out and the device discards the packet. This command specifies the number of devices a BPDU passes through in a region before being discarded. Changing the maximum hop count will affect all instances.

Verification

- Display the configuration.
- Run the **show spanning-tree max-hops** command to display the configured maximum hop count.

Configuration Example

Configuring the Maximum Hop Count of a BPDU Packet

Configuration Steps	● Set the maximum hop count of a BPDU packet to 25.
	FS(config)# spanning-tree max-hops 25
Verification	● Run the **show spanning-tree** command to display the configuration.
	FS# show spanning-tree StpVersion : MSTP SysStpStatus : ENABLED MaxAge : 20 HelloTime : 2 ForwardDelay : 15 BridgeMaxAge : 20 BridgeHelloTime : 2 BridgeForwardDelay : 15 MaxHops: 25 TxHoldCount : 3 PathCostMethod : Long BPDUGuard : Disabled BPDUFilter : Disabled LoopGuardDef : Disabled ###### mst 0 vlans map : ALL BridgeAddr : 00d0.f822.3344 Priority: 0 TimeSinceTopologyChange : 2d:0h:46m:4s TopologyChanges : 25 DesignatedRoot : 0.001a.a917.78cc RootCost : 0 RootPort : GigabitEthernet 0/1 CistRegionRoot : 0.001a.a917.78cc CistPathCost : 20000

10.4.8 Enabling PortFast-related Features

Configuration Effect

- After PortFast is enabled on a port, the port directly enters the forwarding state. However, since the Port Fast Operational State

becomes disabled due to receipt of BPDUs, the port can properly run the STP algorithm and enter the forwarding state.

- If BPDU guard is enabled on a port, the port enters the error-disabled state after receiving a BPDU.
- If BPDU filter is enabled on a port, the port neither sends nor receives BPDUs.

Notes

- The global BPDU guard takes effect only when PortFast is enabled on a port.
- If BPDU filter is enabled globally, a PortFast-enabled port neither sends nor receives BPDUs. In this case, the host connecting directly to the PortFast-enabled port does not receive any BPDUs. If the port changes its Port Fast Operational State to Disabled after receiving a BPDU, BPDU filter automatically fails.
- The global BPDU filter takes effect only when PortFast is enabled on a port.

Configuration Steps

Enabling PortFast

- Optional.
- If a port connects directly to the network terminal, configure this port as a PortFast port.
- In global configuration mode, run the **spanning-tree portfast default** command to enable PortFast on all ports and the **no spanning-tree portfast default** command to disable PortFast on all ports.
- In interface configuration mode, run the **spanning-tree portfast** command to enable PortFast on a port and the **spanning-tree portfast disabled** command to disable PortFast on a port.

Command	**spanning-tree portfast default**
Parameter Description	N/A
Defaults	PortFast is disabled on all ports by default.
Command Mode	Global configuration mode
Usage Guide	N/A

Command	**spanning-tree portfast**
Parameter Description	N/A
Defaults	PortFast is disabled on a port by default.
Command Mode	Interface configuration mode
Usage Guide	After PortFast is enabled on a port, the port directly enters the forwarding state. However, since the Port Fast Operational State becomes disabled due to receipt of BPDUs, the port can properly run the STP algorithm and enter the forwarding state.

Enabling BPDU Guard

- Optional.
- If device ports connect directly to network terminals, you can enable BPDU guard on these ports to prevent BPDU attacks from

causing abnormality in the spanning tree topology. A port enabled with BPDU guard enters the error-disabled state after receiving a BPDU.

- If device ports connect directly to network terminals, you can enable BPDU guard to prevent loops on the ports. The prerequisite is that the downlink device (such as the hub) can forward BPDU packets.
- In global configuration mode, run the **spanning-tree portfast bpduguard default** command to enable BPDU guard on all ports and the **no spanning-tree portfast bpduguard default** command to disable BPDU guard on all ports.
- In interface configuration mode, run the **spanning-tree bpduguard enabled** command to enable BPDU guard on a port and the **spanning-tree bpduguard disabled** command to disable BPDU guard on a port.

Command	**spanning-tree portfast bpduguard default**
Parameter Description	N/A
Defaults	BPDU guard is globally disabled by default.
Command Mode	Global configuration mode
Usage Guide	If BPDU guard is enabled on a port, the port enters the error-disabled state after receiving a BPDU. Run the **show spanning-tree** command to display the configuration.

Command	**spanning-tree bpduguard enabled**
Parameter Description	N/A
Defaults	BPDU guard is disabled on a port by default.
Command Mode	Interface configuration mode
Usage Guide	If BPDU guard is enabled on a port, the port enters the error-disabled state after receiving a BPDU.

Enabling BPDU Filter

- Optional.
- To prevent abnormal BPDU packets from affecting the spanning tree topology, you can enable BPDU filter on a port to filter abnormal BPDU packets.
- In global configuration mode, run the **spanning-tree portfast bpdufilter default** command to enable BPDU filter on all ports and the **no spanning-tree portfast bpdufilter default** command to disable BPDU filter on all ports.
- In interface configuration mode, run the **spanning-tree bpdufilter enabled** command to enable BPDU filter on a port and the **spanning-tree bpdufilter disabled** command to disable BPDU filter on a port.

Command	**spanning-tree portfast bpdufilter default**
Parameter Description	N/A
Defaults	BPDU filter is globally disabled by default.
Command Mode	Global configuration mode
Usage Guide	If BPDU filter is enabled, corresponding ports neither send nor receive BPDUs.

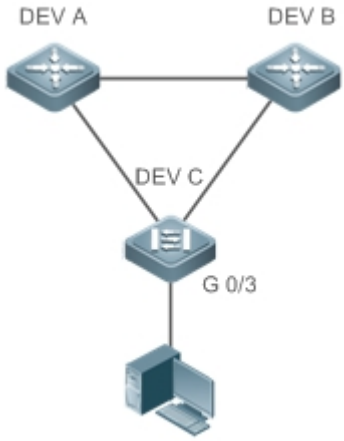
Command	**spanning-tree bpdufilter enabled**
Parameter Description	N/A
Defaults	BPDU filter is disabled on a port by default.
Command Mode	Interface configuration mode
Usage Guide	If BPDU filter is enabled on a port, the port neither sends nor receives BPDUs.

Verification

- Display the configuration.
- Run the **show spanning-tree** [**mst** *instance-id*] **interface** *interface-id* command to display the spanning tree configuration of the port.

Configuration Example

Enabling PortFast on a Port

Scenario **Figure 9-25**	
Configuration Steps	- Configure Gi 0/3 of DEV C as a PortFast port and enable BPDU guard.
DEV C	FS(config)# int gi 0/3 FS(config-if-GigabitEthernet 0/3)# spanning-tree portfast %Warning: portfast should only be enabled on ports connected to a single host. Connecting hubs, switches, bridges to this interface when portfast is enabled,can cause temporary loops. FS(config-if-GigabitEthernet 0/3)#spanning-tree bpduguard enable
Verification	- Run the **show spanning-tree interface** command to display the port configuration.
DEV C	FS#show spanning-tree int gi 0/3

```
PortAdminPortFast : Enabled
PortOperPortFast : Enabled
PortAdminAutoEdge : Enabled
PortOperAutoEdge : Enabled
PortAdminLinkType : auto
PortOperLinkType : point-to-point
PortBPDUGuard : Enabled
PortBPDUFilter : Disabled
PortGuardmode    : None

###### MST 0 vlans mapped :ALL
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 0.00d0.f822.3344
PortDesignatedCost : 0
PortDesignatedBridge :0.00d0.f822.3344
PortDesignatedPortPriority : 128
PortDesignatedPort : 4
PortForwardTransitions : 1
PortAdminPathCost : 20000
PortOperPathCost : 20000
Inconsistent states : normal
PortRole : designatedPort
```

10.4.9 Enabling TC-related Features

Configuration Effect

- If TC protection is enabled on a port, the port deletes TC BPDU packets within a specified time (generally 4 seconds) after receiving them, preventing MAC and ARP entry from being removed.
- If TC guard is enabled, a port receiving TC packets filters TC packets received or generated by itself so that TC packets are not spread to other ports. In this way, possible TC attacks are efficiently prevented to keep the network stable.
- TC filter does not process TC packets received by ports but processes TC packets in case of normal topology changes.

Notes

- It is recommended to enable TC guard only when illegal TC attack packets are received in the network.

Configuration Steps

Enabling TC Protection

- Optional.
- TC protection is disabled by default.
- In global configuration mode, run the **spanning-tree tc-protection** command to enable TC protection on all ports and the **no spanning-tree tc-protection** command to disable TC protection on all ports.
- TC protection can only be enabled or disabled globally.

Command	**spanning-tree tc-protection**
Parameter Description	N/A
Defaults	TC protection is disabled by default.
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling TC Guard

- Optional.
- TC guard is disabled by default.
- To filter TC packets received or generated due to topology changes, you can enable TC guard.
- In global configuration mode, run the **spanning-tree tc-protection tc-guard** command to enable TC guard on all ports and the **no spanning-tree tc-protection tc-guard** command to disable TC guard on all ports.
- In interface configuration mode, run the **spanning-tree tc-guard** command to enable TC guard on a port and the **no spanning-tree tc-guard** command to disable TC guard on a port.

Command	**spanning-tree tc-protection tc-guard**
Parameter Description	N/A
Defaults	TC guard is globally disabled by default.
Command Mode	Global configuration mode
Usage Guide	Enable TC guard to prevent TC packets from spreading.

Command	**spanning-tree tc-guard**
Parameter Description	N/A
Defaults	TC guard is disabled on a port by default.
Command Mode	Interface configuration mode
Usage Guide	Enable TC guard to prevent TC packets from spreading.

Enabling TC Filter

- Optional.
- TC filter is disabled by default.
- To filter TC packets received on a port, you can enable TC filter on the port.
- In interface configuration mode, run the **spanning-tree ignore tc** command to enable TC filter on a port and the **no spanning-tree ignore tc** command to disable it on a port.

Command	**spanning-tree ignore tc**
Parameter Description	N/A
Defaults	TC filter is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	If TC filter is enabled on a port, the port does not process received TC packets.

Verification

- Display the configuration.

Configuration Example

Enabling TC Guard on a Port

Configuration Steps	Enable TC guard on a port.
	FS(config)#int gi 0/1 FS(config-if-GigabitEthernet 0/1)#spanning-tree tc-guard
Verification	• Run the **show run interface** command to display the TC guard configuration of the port.
	FS#show run int gi 0/1 Building configuration... Current configuration : 134 bytes interface GigabitEthernet 0/1 switchport mode trunk spanning-tree tc-guard

Common Errors

- If TC guard or TC filter is incorrectly configured, an error may occur during packet forwarding of the network device. For example, when the topology changes, the device fails to clear MAC address in a timely manner, causing packet forwarding errors.

10.4.10 Enabling BPDU Source MAC Address Check

Configuration Effect

- Enable BPDU source MAC address check. After this, a device receives only BPDU packets with the source MAC address being the specified MAC address and discards other BPDU packets.

Notes

- When the switch connected to a port on a point-to-point link is determined, you can enable BPDU source MAC address check so that the switch receives the BPDU packets sent only by the peer switch.

Configuration Steps

Enabling BPDU Source MAC Address Check

- Optional.
- To prevent malicious BPDU attacks, you can enable BPDU source MAC address check.
- In interface configuration mode, run the **bpdu src-mac-check** *H.H.H* command to enable BPDU source MAC address check on a port and the **no bpdu src-mac-check** command to disable it on a port.

Command	**bpdu src-mac-check** *H.H.H*
Parameter Description	*H.H.H*: Indicates an MAC address. The device receives only BPDU packets with this address being the source MAC address.
Defaults	BPDU source MAC address check is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	BPDU source MAC address check prevents BPDU packets from maliciously attacking switches and causing MSTP abnormal. When the switch connected to a port on a point-to-point link is determined, you can enable BPDU source MAC address check to receive BPDU packets sent only by the peer switch and discard all other BPDU packets, thereby preventing malicious attacks. You can enable BPDU source MAC address check in interface configuration mode for a specific port. One port can only filter one MAC address.

Verification

- Display the configuration.

Configuration Example

Enabling BPDU Source MAC Address Check on a Port

Configuration Steps	Enable BPDU source MAC address check on a port.
	FS(config)#int gi 0/1 FS(config-if-GigabitEthernet 0/1)#bpdu src-mac-check 00d0.f800.1234
Verification	- Run the **show run interface** command to display the spanning tree configuration of the port.
	FS#show run int gi 0/1 Building configuration...

	Current configuration : 170 bytes interface GigabitEthernet 0/1 switchport mode trunk bpdu src-mac-check 00d0.f800.1234 spanning-tree link-type point-to-point

Common Errors

- If BPDU source MAC address check is enabled on a port, the port receives only BPDU packets with the configured MAC address being the source MAC address and discards all other BPDU packets.

10.4.11 Configuring Auto Edge

Configuration Effect

- Enable Auto Edge. If a designated port does not receive any BPDUs within a specified time (3 seconds), it is automatically identified as an edge port. However, if the port receives BPDUs, its Port Fast Operational State will become Disabled.

Notes

- Unless otherwise specified, do not disable Auto Edge.

Configuration Steps

Configuring Auto Edge

- Optional.
- Auto Edge is enabled by default.
- In interface configuration mode, run the **spanning-tree autoedge** command to enable Auto Edge on a port and the **spanning-tree autoedge disabled** command to disable it on a port.

Command	**spanning-tree autoedge**
Parameter Description	N/A
Defaults	Auto Edge is enabled by default.
Command Mode	Interface configuration mode
Usage Guide	If the designated port of a device does not receive a BPDU from the downlink port within a specific period (3 seconds), the device regards a network device connected to the designated port, configures the port as an edge port, and switches the port directly into the forwarding state. The edge port will be automatically identified as a non-edge port after receiving a BPDU. You can run the **spanning-tree autoedge disabled** command to disable Auto Edge.

Verification

- Display the configuration.

Configuration Example

Disabling Auto Edge on a Port

Configuration Steps	Disable Auto Edge on a port.
	FS(config)#int gi 0/1 FS(config-if-GigabitEthernet 0/1)#spanning-tree autoedge disabled
Verification	● Run the **show spanning-tree interface** command to display the spanning tree configuration of the port.
	FS#show spanning-tree interface gi 0/1 PortAdminPortFast : Disabled PortOperPortFast : Disabled PortAdminAutoEdge : Disabled PortOperAutoEdge : Disabled PortAdminLinkType : point-to-point PortOperLinkType : point-to-point PortBPDUGuard : Disabled PortBPDUFilter : Disabled PortGuardmode : None ###### MST 0 vlans mapped :ALL PortState : forwarding PortPriority : 128 PortDesignatedRoot : 0.00d0.f822.3344 PortDesignatedCost : 0 PortDesignatedBridge :0.00d0.f822.3344 PortDesignatedPortPriority : 128 PortDesignatedPort : 2 PortForwardTransitions : 6 PortAdminPathCost : 20000 PortOperPathCost : 20000 Inconsistent states : normal PortRole : designatedPort

Common Errors

N/A

10.4.12 Enabling Guard-related Features

Configuration Effect

- If root guard is enabled on a port, its roles on all instances are enforced as the designated port. Once the port receives configuration information with a higher priority, it enters the root-inconsistent (blocking) state. If the port does not receive configuration information with a higher priority within a period, it returns to its original state.
- Due to the unidirectional link failure, the root port or backup port becomes the designated port and enters the forwarding state if it does not receive BPDUs, causing a network loop. Loop guard is to prevent this problem.

Notes

- Root guard and loop guard cannot take effect on a port at the same time.

Configuration Steps

Enabling Root Guard

- Optional.
- The root bridge may receive configuration with a higher priority due to incorrect configuration by maintenance personnel or malicious attacks in the network. As a result, the current root bridge may lose its role, causing incorrect topology changes. To prevent this problem, you can enable root guard on a designated port of a device.
- In interface configuration mode, run the **spanning-tree guard root** command to enable root guard on a port and the **no spanning-tree guard root** command to disable it on a port.

Command	**spanning-tree guard root**
Parameter Description	N/A
Defaults	Root guard is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	If root guard is enabled, the current root bridge will not change due to incorrect configuration or illegal packet attacks.

Enabling Loop Guard

- Optional.
- You can enable loop guard on a port (root port, master port, or AP) to prevent it from failing to receive BPDUs sent by the designated bridge, increasing device stability. Otherwise, the network topology will change, possibly causing a loop.
- In global configuration mode, run the **spanning-tree loopguard default** command to enable loop guard on all ports and the **no spanning-tree loopguard default** command to disable it on all ports.
- In interface configuration mode, run the **spanning-tree guard loop** command to enable loop guard on a port and the **no spanning-tree guard loop** command to disable it on a port.

Command	**spanning-tree loopguard default**
Parameter	N/A

Description	
Defaults	Loop guard is disabled by default.
Command Mode	Global configuration mode
Usage Guide	Enabling loop guard on a root port or backup port will prevent possible loops caused by BPDU receipt failure.

Command	**spanning-tree guard loop**
Parameter Description	N/A
Defaults	Loop guard is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	Enabling loop guard on a root port or backup port will prevent possible loops caused by BPDU receipt failure.

Disabling Guard

- Optional.

Command	**spanning-tree guard none**
Parameter Description	N/A
Defaults	Guard is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

- Display the configuration.

Configuration Example

Enabling Loop Guard on a Port

Scenario Figure 9-26	
Configuration Steps	• Configure DEV A as the root bridge and DEV B as a non-root bridge on a spanning tree. • Enable loop guard on ports Gi 0/1 and Gi 0/2 of DEV B.

DEV A	FS(config)#spanning-tree FS(config)#spanning-tree mst 0 priority 0
DEV B	FS(config)#spanning-tree FS(config)# int range gi 0/1-2 FS(config-if-range)#spanning-tree guard loop
Verification	● Run the **show spanning-tree interface** command to display the spanning tree configuration of the port.
DEV A	Omitted.
DEV B	FS#show spanning-tree int gi 0/1 PortAdminPortFast : Disabled PortOperPortFast : Disabled PortAdminAutoEdge : Enabled PortOperAutoEdge : Disabled PortAdminLinkType : auto PortOperLinkType : point-to-point PortBPDUGuard : Disabled PortBPDUFilter : Disabled PortGuardmode : Guard loop ###### MST 0 vlans mapped :ALL PortState : forwarding PortPriority : 128 PortDesignatedRoot : 0.001a.a917.78cc PortDesignatedCost : 0 PortDesignatedBridge :0.001a.a917.78cc PortDesignatedPortPriority : 128 PortDesignatedPort : 17 PortForwardTransitions : 1 PortAdminPathCost : 20000 PortOperPathCost : 20000 Inconsistent states : normal PortRole : rootPort

```
FS#show spanning-tree int gi 0/2

PortAdminPortFast : Disabled
PortOperPortFast : Disabled
PortAdminAutoEdge : Enabled
PortOperAutoEdge : Disabled
PortAdminLinkType : auto
PortOperLinkType : point-to-point
PortBPDUGuard : Disabled
PortBPDUFilter : Disabled
PortGuardmode    : Guard loop

###### MST 0 vlans mapped :ALL
PortState : discarding
PortPriority : 128
PortDesignatedRoot : 0.001a.a917.78cc
PortDesignatedCost : 0
PortDesignatedBridge :0.001a.a917.78cc
PortDesignatedPortPriority : 128
PortDesignatedPort : 18
PortForwardTransitions : 1
PortAdminPathCost : 20000
PortOperPathCost : 20000
Inconsistent states : normal
PortRole : alternatePort
```

Common Errors

- If root guard is enabled on the root port, master port, or AP, the port may be incorrectly blocked.

10.4.13 Enabling BPDU Transparent Transmission

Configuration Effect

- If STP is disabled on a device, the device needs to transparently transmit BPDU packets so that the spanning tree between devices is properly calculated.

Notes

- BPDU transparent transmission takes effect only when STP is disabled. If STP is enabled on a device, the device does not transparently transmit BPDU packets.

Configuration Steps

Enabling BPDU Transparent Transmission

- Optional.
- If STP is disabled on a device that needs to transparently transmit BPDU packets, enable BPDU transparent transmission.
- In global configuration mode, run the **bridge-frame forwarding protocol bpdu** command to enable BPDU transparent transmission and the **no bridge-frame forwarding protocol bpdu** command to disable it.
- BPDU transparent transmission takes effect only when STP is disabled. If STP is enabled on a device, the device does not transparently transmit BPDU packets.

Command	**bridge-frame forwarding protocol bpdu**
Parameter Description	N/A
Defaults	BPDU transparent transmission is disabled by default.
Command Mode	Global configuration mode
Usage Guide	In IEEE 802.1Q, the destination MAC address 01-80-C2-00-00-00 of the BPDU is used as a reserved address. That is, devices compliant with IEEE 802.1Q do not forward the BPDU packets received. However, devices may need to transparently transmit BPDU packets in actual network deployment. For example, if STP is disabled on a device, the device needs to transparently transmit BPDU packets so that the spanning tree between devices is properly calculated. BPDU transparent transmission takes effect only when STP is disabled. If STP is enabled on a device, the device does not transparently transmit BPDU packets.

Verification

- Display the configuration.

Configuration Example

Enabling BPDU Transparent Transmission

Scenario Figure 9-27	
	STP is enabled on DEV A and DEV C while is disabled on DEV B.
Configuration Steps	- Enable BPDU transparent transmission on DEV B so that STP between DEV A and DEV C can be correctly calculated.
DEV B	FS(config)#bridge-frame forwarding protocol bpdu
Verification	- Run the **show run** command to check whether BPDU transparent transmission is enabled.
DEV B	FS#show run

	Building configuration... Current configuration : 694 bytes bridge-frame forwarding protocol bpdu

10.4.14 Enabling BPDU Tunnel

Configuration Effect

- Enable BPDU Tunnel so that STP packets from the customer network can be transparently transmitted across the SP network. STP packet transmission between the customer network does not affect the SP network, causing STP on the customer network to be calculated independently of that on the SP network.

Notes

- BPDU Tunnel takes effect only when it is enabled in both global configuration mode and interface configuration mode.

Configuration Steps

Enabling BPDU Tunnel

- (Optional) In a QinQ network, you can enable BPDU Tunnel if STP needs to be calculated separately between customer networks and SP networks.
- BPDU Tunnel is disabled by default.
- In global configuration mode, run the **l2protocol-tunnel stp** command to globally enable BPDU Tunnel and the **no l2protocol-tunnel stp** command to globally disable it.
- In interface configuration mode, run the **l2protocol-tunnel stp enable** command to enable BPDU Tunnel on a port and the **no l2protocol-tunnel stp enable** command to disable it on a port.
- Run the **l2protocol-tunnel stp tunnel-dmac** *mac-address* command in global configuration mode to configure the transparent transmission address of BPDU Tunnel.
- BPDU Tunnel takes effect only when it is enabled in both global configuration mode and interface configuration mode.

Command	**l2protocol-tunnel stp**
Parameter Description	N/A
Defaults	BPDU Tunnel is disabled by default.
Command Mode	Global configuration mode
Usage Guide	BPDU Tunnel takes effect only when it is enabled in both global configuration mode and interface configuration mode.

Command	**l2protocol-tunnel stp enable**
Parameter Description	N/A

Defaults	BPDU Tunnel is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	BPDU Tunnel takes effect only when it is enabled in both global configuration mode and interface configuration mode.

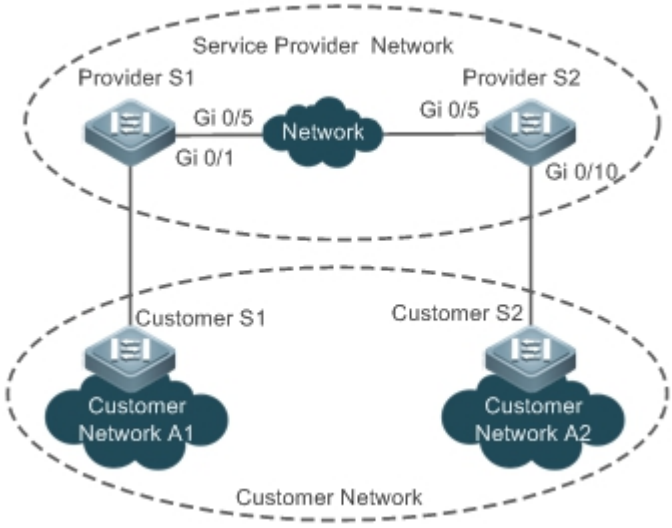
Command	**l2protocol-tunnel stp tunnel-dmac** *mac-address*
Parameter Description	*mac-address*: Indicates the STP address for transparent transmission.
Defaults	The default MAC address is 01d0.f800.0005.
Command Mode	Global configuration mode
Usage Guide	If an STP packet sent from a customer network enters a PE, the PE changes the destination MAC address of the packet to a private address before the packet is forwarded by the SP network. When the packet reaches the PE at the peer end, the PE changes the destination MAC address to a public address and returns the packet to the customer network at the peer end, realizing transparent transmission across the SP network. This private address is the transparent transmission address of BPDU Tunnel. Optional transparent transmission addresses of STP packets include 01d0.f800.0005, 011a.a900.0005, 010f.e200.0003, 0100.0ccd.cdd0, 0100.0ccd.cdd1, and 0100.0ccd.cdd2. If no transparent transmission address is configured, BPDU Tunnel uses the default address 01d0.f800.0005.

Verification

- Run the **show l2protocol-tunnel stp** command to display the BPDU Tunnel configuration.

Configuration Example

Enabling BPDU Tunnel

Scenario **Figure 9-28**	
Configuration	- Enable basic QinQ on the PEs (Provider S1/Provider S2 in this example) so that data packets of the customer

Steps	network are transmitted within VLAN 200 on the SP network. ● Enable STP transparent transmission on the PEs (Provider S1/Provider S2 in this example) so that the SP network can transmit STP packets of the customer network through BPDU Tunnel.
Provider S1	Step 1: Create VLAN 200 on the SP network. FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#vlan 200 FS(config-vlan)#exit Step 2: Enable basic QinQ on the port connected to the customer network and use VLAN 20 for tunneling. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel FS(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 200 FS(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel allowed vlan add untagged 200 Step 3: Enable STP transparent transmission on the port connected to the customer network. FS(config-if-GigabitEthernet 0/1)#l2protocol-tunnel stp enable FS(config-if-GigabitEthernet 0/1)#exit Step 4: Enable STP transparent transmission in global configuration mode. FS(config)#l2protocol-tunnel stp Step 5: Configure an Uplink port. FS(config)# interface gigabitEthernet 0/5 FS(config-if-GigabitEthernet 0/5)#switchport mode uplink
Provider S2	Configure Provider S2 by performing the same steps.
Verification	● Check whether the BPDU Tunnel configuration is correct. ● Verify the Tunnel port configuration by checking whether: 1. The port type is dot1q-tunnel; 2. The outer tag VLAN is consistent with the native VLAN and added to the VLAN list of the Tunnel port; 3. The port that accesses the SP network is configured as an Uplink port.
Provider S1	Step 1: Check whether the BPDU Tunnel configuration is correct. FS#show l2protocol-tunnel stp L2protocol-tunnel: stp Enable L2protocol-tunnel destination mac address: 01d0.f800.0005 GigabitEthernet 0/1 l2protocol-tunnel stp enable Step 2: Check whether the QinQ configuration is correct.

	FS#show running-config interface GigabitEthernet 0/1 switchport mode dot1q-tunnel switchport dot1q-tunnel allowed vlan add untagged 200 switchport dot1q-tunnel native vlan 200 l2protocol-tunnel stp enable spanning-tree bpdufilter enable ! interface GigabitEthernet 0/5 switchport mode uplink
Provider S2	Verify Provider S2 configuration by performing the same steps.

Common Errors

- In the SP network, BPDU packets can be correctly transparently transmitted only when the transparent transmission addresses of BPDU Tunnel are consistent.

10.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the statistics of packets sent and received on a port.	**clear spanning-tree counters** [**interface** *interface-id*]
Clears the STP topology change information.	**clear spanning-tree mst** *instance-id* **topochange record**

Displaying

Description	Command
Displays MSTP parameters and spanning tree topology information.	**show spanning-tree**
Displays the count of sent and received MSTP packets.	**show spanning-tree counters** [**interface** *interface-id*]
Displays MSTP instances and corresponding port forwarding status.	**show spanning-tree summary**
Displays the ports that are blocked by root guard or loop guard.	**show spanning-tree inconsistentports**
Displays the configuration of an MST region.	**show spanning-tree mst configuration**
Displays MSTP information of an instance.	**show spanning-tree mst** *instance-id*
Displays MSTP information of the instance corresponding to a port.	**show spanning-tree mst** *instance-id* **interface** *interface-id*
Displays topology changes of a port in an instance.	**show spanning-tree mst** *instance-id* **topochange record**

Description	Command
Displays MSTP information of all instances corresponding to a port.	**show spanning-tree interface** *interface-id*
Displays the forwarding time.	**show spanning-tree forward-time**
Displays the hello time.	**show spanning-tree hello time**
Displays the maximum hop count.	**show spanning-tree max-hops**
Displays the maximum number of BPDU packets sent per second.	**show spanning-tree tx-hold-count**
Displays the path cost calculation method.	**show spanning-tree pathcost method**
Displays BPDU Tunnel information.	**show l2protocol-tunnel stp**

Debugging

System resources are occupied when debugging information is output. Therefore, disable the debugging switch immediately after use.

Description	Command
Debugs all STPs.	**debug mstp all**
Debugs MSTP Graceful Restart (GR).	**debug mstp gr**
Debugs BPDU packet receiving.	**debug mstp rx**
Debugs BPDU packet sending.	**debug mstp tx**
Debugs MSTP events.	**debug mstp event**
Debugs loop guard.	**debug mstp loopguard**
Debugs root guard.	**debug mstp rootguard**
Debugs the bridge detection state machine.	**debug mstp bridgedetect**
Debugs the port information state machine.	**debug mstp portinfo**
Debugs the port protocol migration state machine.	**debug mstp protomigrat**
Debugs MSTP topology changes.	**debug mstp topochange**
Debugs the MSTP receiving state machine.	**debug mstp receive**
Debugs the port role transition state machine.	**debug mstp roletran**
Debugs the port state transition state machine.	**debug mstp statetran**
Debugs the MSTP sending state machine.	**debug mstp transmit**

11 Configuring GVRP

11.1 Overview

The GARP VLAN Registration Protocol (GVRP) is an application of the Generic Attribute Registration Protocol (GARP) used to dynamically configure and proliferate VLAN memberships.

GVRP simplifies VLAN configuration and management. It reduces the workload of manually configuring VLANs and adding ports to VLANs, and reduces the possibility of network disconnection due to inconsistent configuration. With GVRP, you can dynamically maintain VLANs and add/remove ports to/from VLANs to ensure VLAN connectivity in a topology.

Protocols and Standards

IEEE standard 802.1D

IEEE standard 802.1Q

11.2 Applications

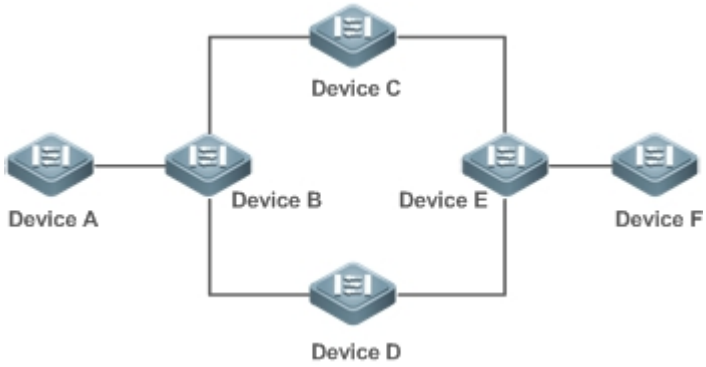
Application	Description
GVRP Configuration in a LAN	Connect two switches in a local area network (LAN) and realize VLAN synchronization.
GVRP PDUs Tunnel Application	Use the GVRP Protocol Data Units (PDUs) Tunnel feature to transparently transmit GVRP packets through a tunnel in a QinQ network environment.

11.2.1 GVRP Configuration in a LAN

Scenario

Enable GVRP and set the GVRP registration mode to Normal to register and deregister all dynamic and static VLANs between Device A and Device F.

Figure 10-1



Remarks	Device A, Device B, Device C, Device D, Device E, and Device F are switches. The ports connected between two devices are Trunk ports. On Device A and Device F, configure static VLANs used for communication. Enable GVRP on all switches.

Deployment

- On each device, enable the GVRP and dynamic VLAN creation features, and ensure that dynamic VLANs can be created on

intermediate devices.

- On Device A and Device F, configure static VLANs used for communication. Device B, Device C, Device D, and Device E will dynamically learn the VLANs through GVRP.

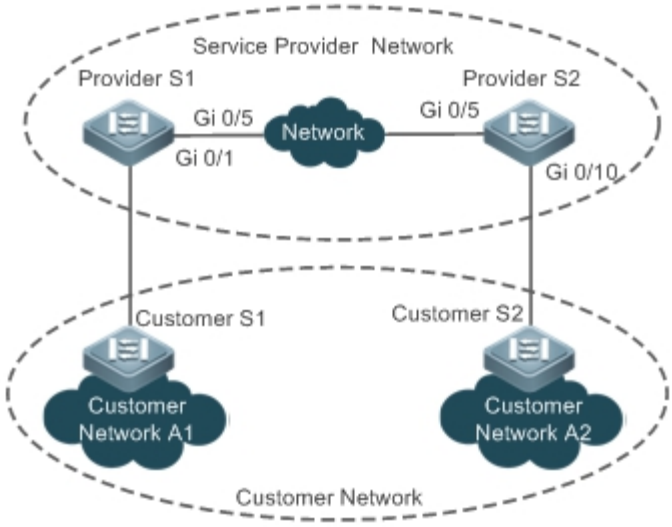
It is recommended that the Spanning Tree Protocol (STP) be enabled to avoid loops in the customer network topology.

11.2.2 GVRP PDUs Tunnel Application

Scenario

A QinQ network environment is generally divided into a customer network and a service provider (SP) network. The GVRP PDUs Tunnel feature allows GVRP packets to be transmitted between customer networks without impact on SP networks. The GVRP calculation in customer networks is separated from that in SP networks without interference.

Figure 10-2 GVRP PDUs Tunnel Application Topology



Remarks	Figure 10-2 shows an SP network and a customer network. The SP network contains the provider edge (PE) devices Provider S1 and Provider S2. Customer Network A1 and Customer Network A2 are the same customer's two sites in different locations. Customer S1 and Customer S2 are the access devices in the customer network, which are connected to the SP network through Provider S1 and Provider S2 respectively. The GVRP PDUs Tunnel feature allows Customer Network A1 and Customer Network A2 to perform unified GVRP calculation across the SP network, without impact on the SP network's GVRP calculation.

Deployment

- Enable basic QinQ on the PEs (Provider S1 and Provider S2) in the SP network to transmit data packets from the customer network through a specified VLAN in the SP network.
- Enable GVRP transparent transmission on the PEs (Provider S1 and Provider S2) in the SP network to allow the SP network to tunnel GVRP packets from the customer network via the GVRP PDUs Tunnel feature.

11.3 Features

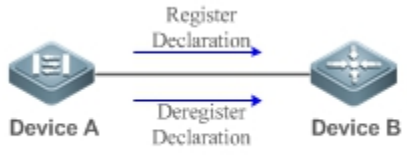
Basic Concepts

GVRP

GVRP is an application of GARP used to register and deregister VLAN attributes in the following modes:

- When a port receives a VLAN attribute declaration, the port will register the VLAN attributes contained in the declaration (that is, the port will join the VLAN).
- When a port receives a VLAN attribute revocation declaration, the port will deregister the VLAN attributes contained in the declaration (that is, the port will exit the VLAN).

Figure 10-3



Dynamic VLAN

A VLAN that can be dynamically created and deleted without the need for manual configuration is called a dynamic VLAN.

You can manually convert a dynamic VLAN to a static VLAN, but not the way around.

A protocol state machine controls the joining of ports to dynamic VLANs created through GVRP. Only the Trunk ports that receive GVRP VLAN attribute declaration can join these VLANs. You cannot manually add ports to dynamic VLANs.

Message Types

(1) Join message

When a GARP application entity hopes other GARP entities to register its attributes, it will send a Join message. When a GARP entity receives a Join message from another entity or requires other entities to register its static attributes, it will send a Join message. There are two types of Join message: JoinEmpty and JoinIn.

- JoinEmpty message: Used to declare an unregistered attribute
- JoinIn message: Used to declare a registered attribute

(2) Leave message

When a GARP application entity hopes other GARP entities to deregister its attributes, it will send a Leave message. When a GARP entity receives a Leave message from another entity or requires other entities to deregister its statically deregistered attributes, it will send a Leave message. There are two types of Leave message: LeaveEmpty and LeaveIn.

- LeaveEmpty message: Used to deregister an unregistered attribute
- LeaveIn message: Used to deregister a registered attribute

(3) LeaveAll message

Each GARP application entity starts its LeaveAll timer during startup. When the timer times out, the entity sends a LeaveAll message to deregister all attributes to enable other GARP entities to reregister attributes. When the GARP application entity receives a LeaveAll

message from another entity, it also sends a LeaveAll message. The LeaveAll timer is restarted when a LeaveAll message is sent again to initiate a new cycle.

Timer Types

GARP defines four timers used to control GARP message sending.

(1) Hold timer

The Hold timer controls the sending of GARP messages (including Join and Leave messages). When a GARP application entity has its attributes changed or receives a GARP message from another entity, it starts the Hold timer. During the timeout period, the GARP application entity encapsulates all GARP messages to be sent into packets as few as possible, and sends the packets when the timer times out. This reduces the quantity of sent packets and saves bandwidth resources.

(2) Join timer

The Join timer controls the sending of Join messages. After a GARP application entity sends a Join message, it waits for one timeout interval of the Join timer to ensure that the Join message is reliably transmitted to another entity. If the GARP application entity receives a JoinIn message from another entity before the timer times out, it will not resend the Join message; otherwise, it will resend the Join message. Not each attribute has its own Join timer, but each GARP application entity has one Join timer.

(3) Leave timer

The Leave timer controls attribute deregistration. When a GARP application entity hopes other entities to deregister one of its attributes, it sends a Leave message. Other entities which receive the Leave message start the Leave timer. The attribute will be deregistered only if these entities receive no Join message mapped to the attribute during the timeout period.

(4) LeaveAll timer

Each GARP application entity starts its own LeaveAll timer upon startup. When the timer times out, the entity sends a LeaveAll message to enable other entities to reregister attributes. Then the LeaveAll timer is restarted to initiate a new cycle.

GVRP Advertising Modes

GVRP allows a switch to inform other interconnected devices of its VLANs and instruct the peer device to create specific VLANs and add the ports that transmit GVRP packets to corresponding VLANs.

Two GVRP advertising modes are available:

- Normal mode: A device externally advertises its VLAN information, including dynamic and static VLANs.
- Non-applicant mode: A device does not externally advertise its VLAN information.

GVRP Registration Modes

A GVRP registration mode specifies whether the switch that receives a GVRP packet processes the VLAN information in the packet, such as dynamically creating a new VLAN and adding the port that receives the packet to the VLAN.

Two GVRP registration modes are available:

- Normal mode: Process the VLAN information in the received GVRP packet.
- Disabled mode: No to process the VLAN information in the received GVRP packet.

Overview

Feature	Description

Intra-Topology VLAN Information Synchronization	Dynamically creates VLANs and adds/removes ports to/from VLANs, which reduces the manual configuration workload and the probability of VLAN disconnection due to missing configuration.

11.3.1 Intra-Topology VLAN Information Synchronization

Working Principle

GVRP is an application of GARP based on the GARP working mechanism. GVRP maintains the dynamic registration information of VLANs on a device and propagates the information to other devices. A GVRP-enabled device receives VLAN registration information from other devices and dynamically updates the local VLAN registration information. The device also propagates the local VLAN registration information to other devices so that all devices in a LAN maintain consistent VLAN information. The VLAN registration information propagated by GVRP includes the manually-configured static registration information on the local device and the dynamic registration information from other devices.

External VLAN Information Advertising

The Trunk port on a GVRP-enabled device periodically collects VLAN information within the port, including the VLANs that the Trunk port joins or exits. The collected VLAN information is encapsulated in a GVRP packet to be sent to the peer device. After the Trunk port on the peer device receives the packet, it resolves the VLAN information. Then corresponding VLANs will be dynamically created, and the Trunk port will join the created VLANs or exit other VLANs. For details about the VLAN information, see the above description of GVRP message types.

VLAN Registration and Deregistration

Upon receiving a GVRP packet, the switch determines whether to process the VLAN information in the packet according to the registration mode of the corresponding port. For details, see the above description of GVRP registration modes.

11.4 Configuration

Configuration	Description and Command	
Configuring Basic GVRP Features and VLAN Information Synchronization	(Mandatory) It is used to enable GVRP and dynamic VLAN creation.	
	gvrp enable	Enables GVRP.
	gvrp dynamic-vlan-creation enable	Enables dynamic VLAN creation.
	switchport mode trunk	Switches to Trunk port mode. GVRP take effects only in Trunk mode.
	switchport trunk allowed vlan all	Allows the traffic from all VLANs to pass through.
	gvrp applicant state	Configures the advertising mode of a port. The Normal mode indicates to advertise VLAN information externally by sending a GVRP packet. The Non-applicant mode indicates not to advertise VLAN information externally.

Configuration	Description and Command	
	gvrp registration mode	Configures the registration mode of a port. The Normal mode indicates to process the VLAN information in the received GVRP packet, such as dynamically creating VLANs and adding ports to VLANs. The Disabled mode indicates not to process the VLAN information in the received GVRP packet.
	(Optional) It is used to configure timers and the registration mode and advertising mode of a port.	
	gvrp timer	Configures timers.
Configuring GVRP PDUs Transparent Transmission	(Optional) It is used to configure GVRP PDUs transparent transmission.	
	bridge-frame forwarding protocol gvrp	Enables GVRP PDUs transparent transmission.
Configuring the GVRP PDUs Tunnel Feature	(Optional) It is used to configure the GVRP PDUs Tunnel feature.	
	l2protocol-tunnel gvrp	Enables the GVRP PDUs Tunnel feature in global configuration mode.
	l2protocol-tunnel gvrp enable	Enables the GVRP PDUs Tunnel feature in interface configuration mode.
	l2protocol-tunnel gvrp tunnel-dmac	Configures the transparent transmission address used by the GVRP PDUs Tunnel feature.

11.4.1 Configuring Basic GVRP Features and VLAN Information Synchronization

Configuration Effect

- Dynamically create/delete VLANs and add/remove ports to/from VLANs.
- Synchronize VLAN information between devices to ensure normal intra-topology communication.
- Reduce the manual configuration workload and simplify VLAN management.

Notes

- GVRP must be enabled on both connected devices. GVRP information is transmitted only by Trunk Links. The transmitted information contains the information of all VLANs on the current device, including dynamically learned VLANs and manually configured VLANs.
- If STP is enabled, only ports in Forwarding state participate in GVRP (such as receiving and sending GVRP PDUs) and have their VLAN information propagated by GVRP.
- All VLAN ports added by GVRP are tagged ports.
- The system does not save the VLAN information that is dynamically learned by GVRP. The information will be lost when the device is reset and cannot be saved manually.
- All devices that need to exchange GVRP information must maintain consistent GVRP timers (Join timer, Leave timer, and Leaveall timer).
- If STP is not enabled, all available ports can participate in GVRP. If Single Spanning Tree (SST) is enabled, only ports in Forwarding state in the SST Context participate in GVRP. If Multi Spanning Tree (MST) is enabled, GVRP can run in the Spanning Tree Context to

which VLAN1 belongs. You cannot specify other Spanning Tree Context for GVRP.

Configuration Steps

Enabling GVRP

- Mandatory.
- Only GVRP-enabled devices can process GVRP packets.
- After GVRP is enabled on a device, the device sends GVRP packets carrying VLAN information. If GVRP is disabled on the device, the device does not send GVRP packets carrying VLAN information or process received GVRP packets.

Command	**gvrp enable**
Parameter Description	N/A
Defaults	By default, GVRP is disabled.
Command Mode	Global configuration mode
Usage Guide	GVRP can be enabled only in global configuration mode. If GVRP is not enabled globally, you can still set other GVRP parameters, but the parameter settings take effect only when GVRP starts running.

Enabling Dynamic VLAN Creation

- Mandatory.
- After dynamic VLAN creation is enabled on a device, the device will dynamically create VLANs upon receiving GVRP Join messages.

The parameters of a dynamic VLAN created through GVRP cannot be modified manually.

Command	**gvrp dynamic-vlan-creation enable**
Parameter Description	N/A
Defaults	By default, dynamic VLAN creation is disabled.
Command Mode	Global configuration mode
Usage Guide	When a port receives a JoinIn or JoinEmpty message that indicates a non-existent VLAN on the local device, GVRP may create this VLAN, depending on the configuration of this command.

Configuring Timers

- Optional.
- There are three GVRP timers: Join timer, Leave timer, and Leaveall timer, which are used to control message sending intervals.
- The timer interval relationships are as follows: The interval of the Leave timer must be three times or more greater than that of the Join timer; the interval of the Leaveall timer must be greater than that of the Leave timer.
- The three timers are controlled by the GVRP state machine and can be triggered by each other.

Command	**gvrp timer** { **join** *timer-value* | **leave** *timer-value* | **leaveall** *timer-value* }
Parameter Description	*timer-value* : 1–2,147,483,647 ms
Defaults	The default interval of the Join timer is 200 ms, that of the Leave timer is 600 ms, and that of the Leaveall timer is

	10,000 ms.
Command Mode	Global configuration mode
Usage Guide	The interval of the Leave timer must be three times or more greater than that of the Join timer. The interval of the Leaveall timer must be greater than that of the Leave timer. The time unit is milliseconds. The following timer intervals are recommended in actual networking: Join timer: 6,000 ms (6s) Leave timer: 30,000 ms (30s) Leaveall timer: 120,000 ms (2 minutes) Ensure that the GVRP timer settings on all interconnected GVRP devices are consistent; otherwise, GVRP may work abnormally.

Configuring the Advertising Mode of a Port

- Optional.
- Two GVRP advertising modes are available: Normal (default) and Non-applicant.
- Normal mode: Indicates that a device externally advertises its VLAN information.
- Non-applicant mode: Indicates that a device does not externally advertise its VLAN information.

Command	**gvrp applicant state** { **normal** | **non-applicant** }
Parameter Description	**normal:** Indicates that a port externally advertises VLAN information. **non-applicant:** Indicates that a port does not externally advertise VLAN information.
Defaults	By default, ports are allowed to send GVRP notification.
Command Mode	Interface configuration mode
Usage Guide	This command is used to configure the GVRP advertising mode of a port.

Configuring the Registration Mode of a Port

- Optional.
- Two GVRP registration modes are available: Normal and Disabled.
- To enable dynamic VLAN registration on a port, run the **gvrp registration mode normal** command. To disable dynamic VLAN registration on a port, run the **gvrp register mode disable** command.
- If dynamic VLAN registration is enabled, dynamic VLANs will be created on the local device when the port receives a GVRP packet carrying VLAN information from the peer end. If dynamic VLAN registration is disabled, no dynamic VLAN will be created on the local device when the port receives a GVRP packet from the peer end.

The two registration modes do not affect the static VLANs on the port. The registration mode for manually-created static VLANs is always Fixed Registrar.

Command	**gvrp registration mode** { **normal** | **disabled** }
Parameter Description	**normal:** Indicates that the port is allowed to join a dynamic VLAN. **disabled:** Indicates that the port is not allowed to join a dynamic VLAN.
Defaults	If GVRP is enabled, the port in Trunk mode is enabled with dynamic VLAN registration by default.

Command Mode	Interface configuration mode
Usage Guide	This command is used to configure the GVRP registration mode of a port.

Switching to Trunk Port Mode

- Mandatory.
- GVRP takes effect only on ports in Trunk mode.

Verification

- Run the **show gvrp configuration** command to check the configuration.
- Check whether a dynamic VLAN is configured and the corresponding port joins the VLAN.

Configuration Example

Enabling GVRP in a Topology and Dynamically Maintaining VLANs and the VLAN-Port Relationship

Scenario Figure 10-4	
Configuration Steps	● On Switch A and Switch C, configure VLANs used for communication in the customer network. ● Enable the GVRP and dynamic VLAN creation features on Switch A, Switch B, and Switch C. ● Configure the ports connected between switches as Trunk ports, and ensure that the VLAN lists of Trunk ports include the communication VLANs. By default, a Trunk port allows the traffic from all VLANs to pass through. ● It is recommended that STP be enabled to avoid loops.
A	1. Create VLAN 1–200 used for communication in the customer network. A# configure terminal Enter configuration commands, one per line. End with CNTL/Z. A(config)# vlan range 1-200 2. Enable the GVRP and dynamic VLAN creation features. A(config)# gvrp enable A(config)# gvrp dynamic-vlan-creation enable 3. Configure the port connected to Switch B as a Trunk port. By default, a Trunk port allows the traffic from all VLANs to pass through. A(config)# interface gigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# switchport mode trunk 4. Configure the advertising mode and registration mode of the Trunk port. The Normal mode is used by default and does not need to be configured manually. A(config-if-GigabitEthernet 0/1)# gvrp applicant state normal A(config-if-GigabitEthernet 0/1)# gvrp registration mode normal

	A(config-if-GigabitEthernet 0/1)# end
C	● The configuration on Switch C is the same as that on Switch A.
B	1. Enable the GVRP and dynamic VLAN creation features. B# configure terminal B(config)# gvrp enable B(config)# gvrp dynamic-vlan-creation enable 2. Configure the ports connected to Switch A and Switch C as Trunk ports. B(config)# interface range GigabitEthernet 0/2-3 B(config-if-GigabitEthernet 0/2)# switchport mode trunk
Verification	Check whether the GVRP configuration on each device is correct. Check whether VLAN 2–100 are dynamically created on Switch B and whether Port G 0/2 and Port G 0/3 on Switch B join the dynamic VLANs.
A	A# show gvrp configuration Global GVRP Configuration: GVRP Feature:enabled GVRP dynamic VLAN creation:enabled Join Timers(ms):200 Leave Timers(ms):600 Leaveall Timers(ms):1000 Port based GVRP Configuration: PORT Applicant Status Registration Mode ---------- ---------- ---------- GigabitEthernet 0/1 normal normal
B	B# show gvrp configuration Global GVRP Configuration: GVRP Feature:enabled GVRP dynamic VLAN creation:enabled Join Timers(ms):200 Leave Timers(ms):600 Leaveall Timers(ms):1000 Port based GVRP Configuration: PORT Applicant Status Registration Mode ---------- ---------- ----------

	GigabitEthernet 0/2 normal normal GigabitEthernet 0/3 normal normal
C	C# show gvrp configuration Global GVRP Configuration: GVRP Feature:enabled GVRP dynamic VLAN creation:enabled Join Timers(ms):200 Leave Timers(ms):600 Leaveall Timers(ms):1000 Port based GVRP Configuration: PORT Applicant Status Registration Mode ---------- ---------- ---------- GigabitEthernet 0/1 normal normal

Common Errors

- The ports connected between devices are not in Trunk mode.
- The VLAN lists of the ports connected between devices do not include the VLANs used for communication in the customer network.
- The GVRP advertising modes and registration modes of Trunk ports are not set to Normal.

11.4.2 Enabling GVRP PDUs Transparent Transmission

Configuration Effect

Enable devices to transparently transmit GVRP PDU frames to realize normal inter-device GVRP calculation when GVRP is not enabled.

Notes

GVRP PDUs transparent transmission takes effect only when GVRP is disabled. After GVRP is enabled, devices will not transparently transmit GVRP PDU frames.

Configuration Steps

Configuring GVRP PDUs Transparent Transmission

- Optional.
- Perform this configuration when you need to enable devices to transparently transmit GVRP PDU frames when GVRP is disabled.

Command	**bridge-frame forwarding protocol gvrp**
Parameter Description	N/A
Defaults	By default, GVRP PDUs transparent transmission is disabled.

Command Mode	Global configuration mode
Usage Guide	In the IEEE 802.1Q standard, the destination MAC address 01-80-C2-00-00-06 for GVRP PDUs is reserved. Devices compliant with IEEE 802.1Q do not forward received GVRP PDU frames. However, in actual network deployment, devices may need to transparently transmit GVRP PDU frames to realize normal inter-device GVRP calculation when GVRP is not enabled. GVRP PDUs transparent transmission takes effect only when GVRP is disabled. After GVRP is enabled, devices will not transparently transmit GVRP PDU frames.

Verification

Run the **show run** command to check whether GVRP PDUs transparent transmission is enabled.

Configuration Example

Configuring GVRP PDUs Transparent Transmission

Scenario **Figure 10-5**	
	Enable GVRP on DEV A and DEV C. (DEV B is not enabled with GVRP.)
Configuration Steps	Configure GVRP PDUs transparent transmission on DEV B to realize normal GVRP calculation between DEV A and DEV C.
DEV B	FS(config)#bridge-frame forwarding protocol gvrp
Verification	Run the **show run** command to check whether GVRP PDUs transparent transmission is enabled.
DEV B	FS#show run Building configuration... Current configuration : 694 bytes bridge-frame forwarding protocol gvrp

11.4.3 Configuring the GVRP PDUs Tunnel Feature

Configuration Effect

Transparently transmit GVRP packets between customer networks through tunnels in SP networks without impact on the SP networks, and thereby separate the GVRP calculation in customer networks from that in SP networks.

Notes

The GVRP PDUs Tunnel feature takes effect after it is enabled in global configuration mode and interface configuration mode.

Configuration Steps

Configuring the GVRP PDUs Tunnel Feature

- (Optional) Perform this configuration when you need to separate GVRP calculation between customer networks and SP networks in a QinQ environment.
- Run the **l2protocol-tunnel gvrp** command in global configuration mode to enable the GVRP PDUs Tunnel feature.
- Run the **l2protocol-tunnel gvrp enable** command in interface configuration mode to enable the GVRP PDUs Tunnel feature.
- Run the **l2protocol-tunnel gvrp tunnel-dmac** *mac-address* command to configure the transparent transmission address used by the GVRP PDUs Tunnel feature.

Command	**l2protocol-tunnel gvrp**
Parameter Description	N/A
Defaults	By default, the GVRP PDUs Tunnel feature is disabled.
Command Mode	Global configuration mode
Usage Guide	The GVRP PDUs Tunnel feature takes effect after it is enabled in global configuration mode and interface configuration mode.

Command	**l2protocol-tunnel gvrp enable**
Parameter Description	N/A
Defaults	By default, the GVRP PDUs Tunnel feature is disabled.
Command Mode	Interface configuration mode
Usage Guide	The GVRP PDUs Tunnel feature takes effect after it is enabled in global configuration mode and interface configuration mode.

Command	**l2protocol-tunnel gvrp tunnel-dmac** *mac-address*
Parameter Description	*mac-address*: Indicates the GVRP address used by transparent transmission.
Defaults	The default address is 01d0.f800.0006.
Command Mode	Global configuration mode
Usage Guide	In GVRP PDUs Tunnel application, when a GVRP packet from a customer network enters the PE in an SP network, the destination MAC address of the packet is changed to a private address before the packet is forwarded in the SP network. When the packet reaches the peer PE, the destination MAC address is changed to a public address before the packet is sent to the customer network at the other end. In this way, the GVRP packet can be transparently transmitted across the SP network. The private address is the transparent transmission address used by the GVRP PDUs Tunnel feature.

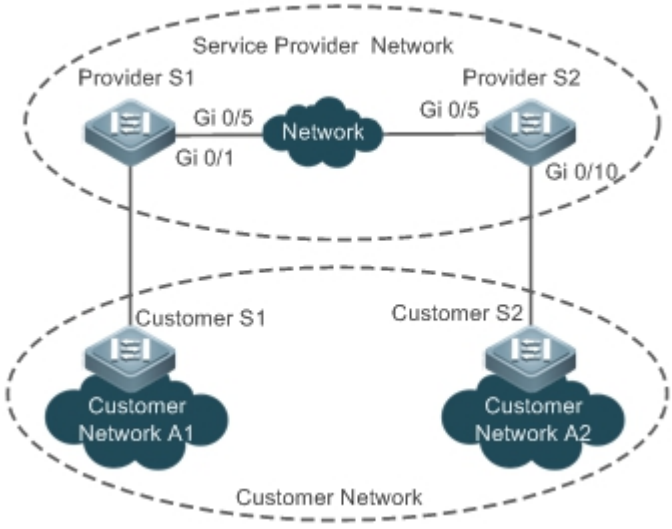
	Address range for transparent transmission of GVRP packets: 01d0.f800.0006, 011a.a900.0006 When no transparent transmission address is configured, the default address 01d0.f800.0006 is used.

Verification

Run the **show l2protocol-tunnel gvrp** command to check the GVRP PDUs Tunnel configuration.

Configuration Example

Configuring the GVRP PDUs Tunnel Feature

Scenario **Figure 10-6**	
Configuration Steps	• Enable basic QinQ on the PEs (Provider S1 and Provider S2) in the SP network to transmit data packets from the customer network through VLAN 200 in the SP network. • Enable GVRP transparent transmission on the PEs (Provider S1 and Provider S2) in the SP network to allow the SP network to tunnel GVRP packets from the customer network via the GVRP PDUs Tunnel feature.
Provider S1	Step 1: Create VLAN 200 of the SP network. FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#vlan 200 FS(config-vlan)#exit Step 2: Enable basic QinQ on the port connected to the customer network to tunnel data from the customer network through VLAN 200. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel FS(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 200 FS(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel allowed vlan add untagged 200 Step 3: Enable GVRP transparent transmission on the port connected to the customer network.

	FS(config-if-GigabitEthernet 0/1)#l2protocol-tunnel gvrp enable FS(config-if-GigabitEthernet 0/1)#exit Step 4: Enable GVRP transparent transmission globally. FS(config)#l2protocol-tunnel gvrp Step 5: Configure an uplink port. FS(config)# interface gigabitEthernet 0/5 FS(config-if-GigabitEthernet 0/5)#switchport mode uplink
Provider S2	The configuration on Provider S2 is similar to that on Provider S1.
Verification	● Check whether the GVRP PDUs Tunnel configuration is correct. ● Check whether the Tunnel port is configured correctly. Pay attention to the following: - The port type is dot1q-tunnel. - The outer tag VLAN is the Native VLAN and added to the VLAN list of the Tunnel port. - The ports on the PEs in the uplink direction are configured as Uplink ports.
Provider S1	1. Check whether the GVRP PDUs Tunnel configuration is correct. FS#show l2protocol-tunnel gvrp L2protocol-tunnel: Gvrp Enable L2protocol-tunnel destination mac address: 01d0.f800.0006 GigabitEthernet 0/1 l2protocol-tunnel gvrp enable 2. Check whether the QinQ configuration is correct. FS#show running-config interface GigabitEthernet 0/1 switchport mode dot1q-tunnel switchport dot1q-tunnel allowed vlan add untagged 200 switchport dot1q-tunnel native vlan 200 l2protocol-tunnel gvrp enable ! interface GigabitEthernet 0/5 switchport mode uplink
Provider S2	The verification on Provider S2 is the same as that on Provider S1.

Common Errors

In an SP network, transparent transmission addresses are not configured consistently, which affects the transmission of GVRP PDU frames.

11.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears port counters.	**clear gvrp statistics** { *interface-id* | **all** }

Displaying

Description	Command
Displays port counters.	**show gvrp statistics** { *interface-id* | **all** }
Displays the current GVRP status.	**show gvrp status**
Displays the current GVRP configuration.	**show gvrp configuration**
Displays the information of the GVRP PDUs Tunnel feature.	**show l2protocol-tunnel gvrp**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Enables GVRP event debugging.	**debug gvrp event**
Enables GVRP timer debugging.	**debug gvrp timer**

12 Configuring LLDP

12.1 Overview

The Link Layer Discovery Protocol (LLDP), defined in the IEEE 802.1AB standard, is used to discover the topology and identify topological changes. LLDP encapsulates local information of a device into LLDP data units (LLDPDUs) in the type/length/value (TLV) format and then sends the LLDPDUs to neighbors. It also stores LLDPDUs from neighbors in the management information base (MIB) to be accessed by the network management system (NMS).

With LLDP, the NMS can learn about topology, for example, which ports of a device are connected to other devices and whether the rates and duplex modes at both ends of a link are consistent. Administrators can quickly locate and rectify a fault based on the information.

A FS LLDP-compliant device is capable of discovering neighbors when the peer is either of the following:

- FS LLDP-compliant device
- Endpoint device that complies with the Link Layer Discovery Protocol-Media Endpoint Discovery (LLDP-MED)

Protocols and Standards

- IEEE 802.1AB 2005: Station and Media Access Control Connectivity Discovery
- ANSI/TIA-1057: Link Layer Discovery Protocol for Media Endpoint Devices

12.2 Applications

Application	Description
Displaying Topology	Multiple switches, a MED device, and an NMS are deployed in the network topology.
Conducting Error Detection	Two switches are directly connected and incorrect configuration will be displayed.

12.2.1 Displaying Topology

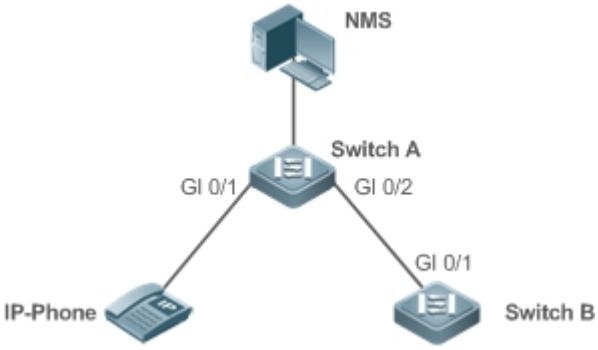
Scenario

Multiple switches, a MED device, and an NMS are deployed in the network topology.

As shown in the following figure, the LLDP function is enabled by default and no additional configuration is required.

- Switch A and Switch B discover that they are neighbors.
- Switch A discovers its neighbor MED device, that is, IP-Phone, through port GigabitEthernet 0/1.
- The NMS accesses MIB of switch A.

Figure 11-1



Remarks	FS Switch A, Switch B, and IP-Phone support LLDP and LLDP-MED. LLDP on switch ports works in TxRx mode. The LLDP transmission interval is 30 seconds and transmission delay is 2 seconds by default.

Deployment

- Run LLDP on a switch to implement neighbor discovery.
- Run the Simple Network Management Protocol (SNMP) on the switch so that the NMS acquires and sets LLDP-relevant information on the switch.

12.2.2 Conducting Error Detection

Scenario

Two switches are directly connected and incorrect configuration will be displayed.

As shown in the following figure, the LLDP function and LLDP error detection function are enabled by default, and no additional configuration is required.

- After you configure a virtual local area network (VLAN), port rate and duplex mode, link aggregation, and maximum transmission unit (MTU) of a port on Switch A, an error will be prompted if the configuration does not match that on Switch B, and vice versa.

Figure 11-2

Remarks	FS Switch A and Switch B support LLDP. LLDP on switch ports works in TxRx mode. The LLDP transmission interval is 30 seconds and transmission delay is 2 seconds by default.

Deployment

- Run LLDP on a switch to implement neighbor discovery and detect link fault.

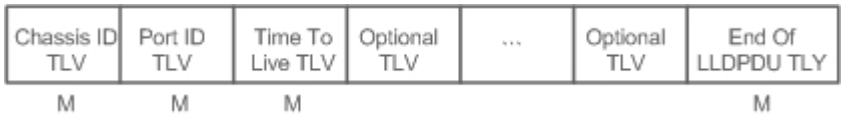
12.3 Features

Basic Concepts

LLDPDU

LLDPDU is a protocol data unit encapsulated into an LLDP packet. Each LLDPDU is a sequence of TLV structures. The TLV collection consists of three mandatory TLVs, a series of optional TLVs, and one End Of TLV. The following figure shows the format of an LLDPDU.

Figure 11-3 LLDPDU Format



In the preceding figure:

- M indicates a mandatory TLV.
- In an LLDPDU, Chassis ID TLV, Port ID TLV, Time To Live TLV, and End Of LLDPDU TLV are mandatory and TLVs of other TLVs are optional.

LLDP Encapsulation Format

LLDP packets can be encapsulated in two formats: Ethernet II and Subnetwork Access Protocols (SNAP).

The following figure shows the format of LLDP packets encapsulated in the Ethernet II format.

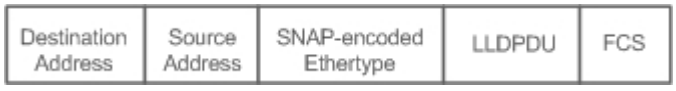
Figure 11-4 Ethernet II Format

In the preceding figure:

- Destination Address: Indicates the destination MAC address, which is the LLDP multicast address 01-80-C2-00-00-0E.
- Source Address: Indicates the source MAC address, which is the port MAC address.
- Ethertype: Indicates the Ethernet type, which is 0x88CC.
- LLDPDU: Indicates the LLDP protocol data unit.
- FCS: Indicates the frame check sequence.

Figure 11-5 shows the format of LLDP packets encapsulated in the SNAP format.

Figure 11-5 SNAP Format



In the preceding figure:

- Destination Address: Indicates the destination MAC address, which is the LLDP multicast address 01-80-C2-00-00-0E.
- Source Address: Indicates the source MAC address, which is the port MAC address.
- SNAP-encoded Ethertype: Indicates the Ethernet type of the SNMP encapsulation, which is AA-AA-03-00-00-00-88-CC.
- LLDPDU: Indicates the LLDP protocol data unit.
- FCS: Indicates the frame check sequence.

TLV

TLVs encapsulated into an LLDPDU can be classified into two types:

- Basic management TLVs
- Organizationally specific TLVs

Basic management TLVs are a collection of basic TLVs used for network management. Organizationally specific TLVs are defined by standard organizations and other institutions, for example, the IEEE 802.1 organization and IEEE 802.3 organization define their own TLV collections.

1. Basic management TLVs

The basic management TLV collection consists of two types of TLVs: mandatory TLVs and optional TLVs. A mandatory TLV must be contained in an LLDPDU for advertisement and an optional TLV is contained selectively.

The following table describes basic management TLVs.

TLV Type	Description	Mandatory/Optional
End Of LLDPDU TLV	Indicates the end of an LLDPDU, occupying two bytes.	Mandatory
Chassis ID TLV	Identifies a device with a MAC address.	Mandatory
Port ID TLV	Identifies a port sending LLDPDUs.	Fixed
Time To Live TLV	Indicates the time to live (TTL) of local information on a neighbor. When a device receives a TLV containing TTL 0, it deletes the neighbor information.	Mandatory
Port Description TLV	Indicates the descriptor of the port sending LLDPDUs.	Optional
System Name TLV	Describes the device name.	Optional
System Description TLV	Indicates the device description, including the hardware version, software version, and operating system information.	Optional
System Capabilities TLV	Describes main functions of the device, such as the bridge, routing, and relay functions.	Optional
Management Address TLV	Indicates the management address, which contains the interface ID and object identifier (OID).	Optional

FS LLDP-compliant switches support advertisement of basic management TLVs.

2. Organizationally specific TLVs

Different organizations, such as the IEEE 802.1, IEEE 802.3, IETF and device suppliers, define specific TLVs to advertise specific information about devices. The organizationally unique identifier (OUI) field in a TLV is used to distinguish different organizations.

- Organizationally specific TLVs are optional and are advertised in an LLDPDU selectively. Currently, there are three types of common organizationally specific TLVs: IEEE 802.1 organizationally specific TLVs, IEEE 802.3 organizationally specific TLVs, and LLDP-MED TLVs.

The following table describes IEEE 802.1 organizationally specific TLVs.

TLV Type	Description
Port VLAN ID TLV	Indicates the VLAN identifier of a port.
Port And Protocol VLAN ID TLV	Indicates the protocol VLAN identifier of a port.
VLAN Name TLV	Indicates the VLAN name of a port.
Protocol Identity TLV	Indicates the protocol type supported by a port.

FS LLDP-compliant switches do not send the Protocol Identity TLV but receive this TLV.

- IEEE 802.3 organizationally specific TLVs

The following table describes IEEE 802.3 organizationally specific TLVs.

TLV Type	Description
MAC/PHY Configuration//Status TLV	Indicates the rate and duplex mode of a port, and whether to support and enable auto-negotiation.
Power Via MDI TLV	Indicates the power supply capacity of a port.
Link Aggregation TLV	Indicates the link aggregation capacity of a port and the current aggregation state.
Maximum Frame Size TLV	Indicates the maximum size of the frame transmitted by a port.

FS LLDP-compliant devices support advertisement of IEEE 802.3 organizationally specific TLVs.

- LLDP-MED TLV

LLDP-MED is an extension to LLDP based on IEEE 802.1AB LLDP. It enables users to conveniently deploy the Voice Over IP (VoIP) network and detect faults. It provides applications including the network configuration policies, device discovery, PoE management, and inventory management, meeting requirements for low cost, effective management, and easy deployment.

The following table describes LLDP-MED TLVs.

TLV Type	Description
LLDP-MED Capabilities TLV	Indicates the type of the LLDP-MED TLV encapsulated into an LLDPDU and device type (network connectivity device or endpoint device), and whether to support LLDP-MED,.
Network Policy TLV	Advertises the port VLAN configuration, supported application type (such as voice or video services), and Layer-2 priority information.
Location Identification TLV	Locates and identifies an endpoint device.
Extended Power-via-MDI TLV	Provides more advanced power supply management.
Inventory – Hardware Revision TLV	Indicates hardware version of a MED device.
Inventory – Firmware Revision TLV	Indicates the firmware version of the MED device.
Inventory – Software Revision TLV	Indicates the software version of the MED device.
Inventory – Serial Number TLV	Indicates the serial number of the MED device.
Inventory – Manufacturer Name TLV	Indicates the name of the manufacturer of the MED device.
Inventory – Model Name TLV	Indicates the module name of the MED device.
Inventory – Asset ID TLV	Indicates the asset identifier of the MED device, used for inventory management and asset tracking.

FS LLDP-compliant FS devices support advertisement of LLDP-MED TLVs.

Overview

Feature	Description
LLDP Work Mode	Configures the mode of transmitting and receiving LLDP packets.
LLDP Transmission Mechanism	Enables directly connected LLDP-compliant devices to send LLDP packets to the peer.
LLDP Reception Mechanism	Enables directly connected LLDP-compliant devices to receive LLDP packets from the peer.

12.3.1 LLDP Work Mode

Configure the LLDP work mode so as to specify the LLDP packet transmission and reception mode.

Working Principle

LLDP provides three work modes:

- TxRx: Transmits and receives LLDPDUs.
- Rx Only: Only receives LLDPDUs.
- Tx Only: Only transmits LLDPDUs.

When the LLDP work mode is changed, the port initializes the protocol state machine. You can set a port initialization delay to prevent repeated initialization of a port due to frequent changes of the LLDP work mode.

Related Configuration

Configuring the LLDP Work Mode

The default LLDP work mode is TxRx.

You can run the **lldp mode** command to configure the LLDP work mode.

If the work mode is set to TxRx, the device can both transmit and receive LLDP packets. If the work mode is set to Rx Only, the device can only receive LLDP packets. If the work mode is set to Tx Only, the device can only transmit LLDP packets. If the work mode is disabled, the device cannot transmit or receive LLDP packets.

12.3.2 LLDP Transmission Mechanism

LLDP packets inform peers of their neighbors. When the LLDP transmission mode is cancelled or disabled, LLDP packets cannot be transmitted to neighbors.

Working Principle

LLDP periodically transmits LLDP packets when working in TxRx or Tx Only mode. When information about the local device changes, LLDP immediately transmits LLDP packets. You can configure a delay time to avoid frequent transmission of LLDP packets caused by frequent changes of local information.
LLDP provides two types of packets:

- Standard LLDP packet, which contains management and configuration information about the local device.
- Shutdown packet: When the LLDP work mode is disabled or the port is shut down, LLDP Shutdown packets will be transmitted. A Shutdown packet consists of the Chassis ID TLV, Port ID TLV, Time To Live TLV, and End OF LLDP TLV. TTL in the Time to Live TLV is 0. When a device receives an LLDP Shutdown packet, it considers that the neighbor information is invalid and immediately deletes it.

When the LLDP work mode is changed from disabled or Rx to TxRx or Tx, or when LLDP discovers a new neighbor (that is, a device receives a new LLDP packet and the neighbor information is not stored locally), the fast transmission mechanism is started so that the neighbor quickly learns the device information. The fast transmission mechanism enables a device to transmit multiple LLDP packets at an interval of 1 second.

Related Configuration

Configuring the LLDP Work Mode

The default work mode is TxRx.

Run the **lldp mode txrx** or **lldp mode tx** command to enable the LLDP packet transmission function. Run the **lldp mode rx** or **no lldp mode** command to disable the LLDP packet transmission function.

In order to enable LLDP packet reception, set the work mode to TxRx or Rx Only. If the work mode is set to Rx Only, the device can only receive LLDP packets.

Configuring the LLDP Transmission Delay

The default LLDP transmission delay is 2 seconds.

Run the **lldp timer tx-delay** command to change the LLDP transmission delay.

If the delay is set to a very small value, the frequent change of local information will cause frequent transmission of LLDP packets. If the delay is set to a very large value, no LLDP packet may be transmitted even if local information is changed.

Configuring the LLDP Transmission Interval

The default LLDP transmission interval is 30 seconds.

Run the **lldp timer tx-interval** command to change the LLDP transmission interval.

If the interval is set to a very small value, LLDP packets may be transmitted frequently. If the interval is set to a very large value, the peer may not discover the local device in time.

Configuring the TLVs to Be Advertised

By default, an interface is allowed to advertise TLVs of all types except Location Identification TLV.

Run the **lldp tlv-enable** command to change the TLVs to be advertised.

Configuring the LLDP Fast Transmission Count

By default, three LLDP packets are fast transmitted.

Run the **lldp fast-count** command to change the number of LLDP packets that are fast transmitted.

Increase or decrease the TLVs in LLDP.

12.3.3 LLDP Reception Mechanism

A device can discover the neighbor and determine whether to age the neighbor information according to received LLDP packets.

Working Principle

A device can receive LLDP packets when working in TxRx or Rx Only mode. After receiving an LLDP packet, a device conducts validity check. After the packet passes the check, the device checks whether the packet contains information about a new neighbor or about an existing neighbor and stores the neighbor information locally. The device sets the TTL of neighbor information according to the value of TTL TLV in the packet. If the value of TTL TLV is 0, the neighbor information is aged immediately.

Related Configuration

Configuring the LLDP Work Mode

The default LLDP work mode is TxRx.

Run the **lldp mode txrx** or **lldp mode rx** command to enable the LLDP packet reception function. Run the **lldp mode tx** or **no lldp**

mode command to disable the LLDP packet reception function.

In order to enable LLDP packet reception, set the work mode to TxRx or Rx Only. If the work mode is set to Tx Only, the device can only transmit LLDP packets.

12.4 Configuration

Configuration	Description and Command	
Configuring the LLDP Function	(Optional) It is used to enable or disable the LLDP function in global or interface configuration mode.	
	lldp enable	Enables the LLDP function.
	no lldp enable	Disables the LLDP function.
Configuring the LLDP Work Mode	(Optional) It is used to configure the LLDP work mode.	
	lldp mode {rx | tx | txrx }	Configures the LLDP work mode.
	no lldp mode	Shuts down the LLDP work mode.
Configuring the TLVs to Be Advertised	(Optional) It is used to configure the TLVs to be advertised.	
	lldp tlv-enable	Configures the TLVs to be advertised.
	no lldp tlv-enable	Cancels TLVs.
Configures the Management Address to Be Advertised	(Optional) It is used to configure the management address to be advertised in LLDP packets.	
	lldp management-address-tlv [*ip-address*]	Configures the management address to be advertised in LLDP packets.
	no lldp management-address-tlv	Cancels the management address.
Configuring the LLDP Fast Transmission Count	(Optional) It is used to configure the number of LLDP packets that are fast transmitted.	
	lldp fast-count *value*	Configures the LLDP fast transmission count.
	no lldp fast-count	Restores the default LLDP fast transmission count.
Configuring the TTL Multiplier and Transmission Interval	(Optional) It is used to configure the TTL multiplier and transmission interval.	
	lldp hold-multiplier *value*	Configures the TTL multiplier.
	no lldp hold-multiplier	Restores the default TTL multiplier.
	lldp timer tx-interval *seconds*	Configures the transmission interval.
	no lldp timer tx-interval	Restores the default transmission interval.
Configuring the Transmission Delay	(Optional) It is used to configure the delay time for LLDP packet transmission.	
	lldp timer tx-delay *seconds*	Configures the transmission delay.
	no lldp timer tx-delay	Restores the default transmission delay.
Configuring the Initialization Delay	(Optional) It is used to configure the delay time for LLDP to initialize on any interface.	
	lldp timer reinit-delay *seconds*	Configures the initialization delay.
	no lldp timer reinit-delay	Restores the default initialization delay.

Configuration	Description and Command	
Configuring the LLDP Trap Function	(Optional) It is used to configure the LLDP Trap function.	
	lldp notification remote-change enable	Enables the LLDP Trap function.
	no lldp notification remote-change enable	Disables the LLDP Trap function.
	lldp timer notification-interval	Configures the LLDP Trap transmission interval.
	no lldp timer notification-interval	Restores the default LLDP Trap transmission interval.
Configuring the LLDP Error Detection Function	(Optional) It is used to configure the LLDP error detection function.	
	lldp error-detect	Enables the LLDP error detection function.
	no lldp error-detect	Disables the LLDP error detection function.
Configuring the LLDP Encapsulation Format	(Optional) It is used to configure the LLDP encapsulation format.	
	lldp encapsulation snap	Sets the LLDP encapsulation format to SNAP.
	no lldp encapsulation snap	Sets the LLDP encapsulation format to Ethernet II.
Configuring the LLDP Network Policy	(Optional) It is used to configure the LLDP Network Policy.	
	lldp network-policy profile *profile-num*	Configures an LLDP Network Policy.
	no lldp network-policy profile *profile-num*	Deletes an LLDP Network Policy.
Configuring the Civic Address	(Optional) It is used to configure the civic address of a device.	
	{ **country** | **state** | **county** | **city** | **division** | **neighborhood** | **street-group** | **leading-street-dir** | **trailing-street-suffix** | **street-suffix** | **number** | **street-number-suffix** | **landmark** | **additional-location-information** | **name** | **postal-code** | **building** | **unit** | **floor** | **room** | **type-of-place** | **postal-community-name** | **post-office-box** | **additional-code** } *ca-word*	Configures the civic address of a device.
	no { **country** | **state** | **county** | **city** | **division** | **neighborhood** | **street-group** | **leading-street-dir** | **trailing-street-suffix** | **street-suffix** | **number** | **street-number-suffix** | **landmark** | **additional-location-information** | **name** | **postal-code** | **building** | **unit** | **floor** | **room** | **type-of-place** | **postal-community-name** | **post-office-box** | **additional-code** } *ca-word*	Deletes civic address of a device.
Configuring the Emergency Telephone Number	(Optional) It is used to configure the emergency telephone number of a device.	
	lldp location elin identifier *id* **elin-location** *tel-number*	Configures the emergency telephone number of a device.

Configuration	Description and Command	
	no lldp location elin identifier *id*	Deletes the emergency telephone number of a device.
Configuring the Detection of Compatible Neighbors	(Optional) It is used to configure the detection of compatible neighbors function.	
	lldp compliance vendor	Enables the detection of compatible neighbors.
	no lldp compliance vendor	Disables the detection of compatible neighbors.

12.4.1 Configuring the LLDP Function

Configuration Effect

- Enable or disable the LLDP function.

Notes

- To make the LLDP function take effect on an interface, you need to enable the LLDP function globally and on the interface.

Configuration Steps

- Optional.
- Configure the LLDP function in global or interface configuration mode.

Verification

Display LLDP status

- Check whether the LLDP function is enabled in global configuration mode.
- Check whether the LLDP function is enabled in interface configuration mode.

Related Commands

Enabling the LLDP Function

Command	**lldp enable**
Parameter Description	N/A
Command Mode	Global configuration mode/Interface configuration mode
Usage Guide	The LLDP function takes effect on an interface only after it is enabled in global configuration mode and interface configuration mode.

Disabling the LLDP Function

Command	**no lldp enable**
Parameter Description	N/A
Command Mode	Global configuration mode/Interface configuration mode
Usage Guide	N/A

Configuration Example

Disabling the LLDP Function

Configuration Steps	Disable the LLDP function in global configuration mode.
	FS(config)#no lldp enable
Verification	Display global LLDP status.
	FS(config)#show lldp status Global status of LLDP: Disable

Common Errors

- If the LLDP function is enabled on an interface but disabled in global configuration mode, the LLDP function does not take effect on the interface.
- A port can learn a maximum of five neighbors.
- If a neighbor does not support LLDP but it is connected to an LLDP-supported device, a port may learn information about the device that is not directly connected to the port because the neighbor may forward LLDP packets.

12.4.2 Configuring the LLDP Work Mode

Configuration Effect

- If you set the LLDP work mode to TxRx, the interface can transmit and receive packets.
- If you set the LLDP work mode to Tx, the interface can only transmit packets but cannot receive packets.
- If you set the LLDP work mode to Rx, the interface can only receive packets but cannot transmit packets.
- If you disable the LLDP work mode, the interface can neither receive nor transmit packets.

Notes

- LLDP runs on physical ports (AP member ports for AP ports). Stacked ports and VSL ports do not support LLDP.

Configuration Steps

- Optional.
- Set the LLDP work mode to Tx or Rx as required.

Verification

Display LLDP status information on an interface

- Check whether the configuration takes effect.

Related Commands

Configuring the LLDP Work Mode

Command	**lldp mode** { **rx** | **tx** | **txrx** }
Parameter Description	**rx**: Only receives LLDPDUs. **tx**: Only transmits LLDPDUs. **txrx**: Transmits and receives LLDPDUs.
Command Mode	Interface configuration mode
Usage Guide	To make LLDP take effect on an interface, make sure to enable LLDP globally and set the LLDP work mode on the interface to Tx, Rx or TxRx.

Disabling the LLDP Work Mode

Command	**no lldp mode**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	After the LLDP work mode on an interface is disabled, the interface does not transmit or receive LLDP packets.

Configuration Example

Configuring the LLDP Work Mode

Configuration Steps	Set the LLDP work mode to Tx in interface configuration mode.
	FS(config)#interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)#lldp mode tx
Verification	Display LLDP status information on the interface.
	FS(config-if-GigabitEthernet 0/1)#show lldp status interface gigabitethernet 0/1 Port [GigabitEthernet 0/1] Port status of LLDP : Enable Port state : UP Port encapsulation : Ethernet II Operational mode : TxOnly Notification enable : NO Error detect enable : YES Number of neighbors : 0 Number of MED neighbors : 0

12.4.3 Configuring the TLVs to Be Advertised

Configuration Effect

- Configure the type of TLVs to be advertised to specify the LLDPDUs in LLDP packets.

Notes

- If you configure the **all** parameter for the basic management TLVs, IEEE 802.1 organizationally specific TLVs, and IEEE 802.3 organizationally specific TLVs, all optional TLVs of these types are advertised.
- If you configure the **all** parameter for the LLDP-MED TLVs, all LLDP-MED TLVs except Location Identification TLV are advertised.
- If you want to configure the LLDP-MED Capability TLV, configure the LLDP 802.3 MAC/PHY TLV first; If you want to cancel the LLDP 802.3 MAC/PHY TLV, cancel the LLDP-MED Capability TLV first.
- If you want to configure LLDP-MED TLVs, configure the LLDP-MED Capability TLV before configuring other types of LLDP-MED TLVs. If you want to cancel LLDP-MED TLVs, cancel the LLDP-MED Capability TLV before canceling other types of LLDP-MED TLVs If a device is connected to an IP-Phone that supports LLDP-MED, you can configure the Network Policy TLV to push policy configuration to the IP-Phone.
- If a device supports the DCBX function by default, ports of the device are not allowed to advertise IEEE 802.3 organizationally specific TLVs and LLDP-MED TLVs by default.

Configuration Steps

- Optional.
- Configure the type of TLVs to be advertised on an interface.

Verification

Display the configuration of TLVs to be advertised on an interface

- Check whether the configuration takes effect.

Related Commands

Configuring TLVs to Be Advertised

Command	**lldp tlv-enable** { **basic-tlv** { **all | port-description | system-capability | system-description | system-name** } **|dot1-tlv** { **all | port-vlan-id | protocol-vlan-id** [*vlan-id*] | **vlan-name** [*vlan-id*] } **|dot3-tlv** { **all | link-aggregation | mac-physic | max-frame-size | power } | med-tlv { all | capability | inventory | location** { **civic-location | elin** } **identifier** *id* **| network-policy profile** [*profile-num*] **| power-over-ethernet** } }
Parameter Description	**basic-tlv**: Indicates the basic management TLV. **port-description**: Indicates the Port Description TLV. **system-capability**: Indicates the System Capabilities TLV. **system-description**: Indicates the System Description TLV. **system-name**: Indicates the System Name TLV. **dot1-tlv**: Indicates the IEEE 802.1 organizationally specific TLVs. **port-vlan-id**: Indicates the Port VLAN ID TLV. **protocol-vlan-id**: Indicates the Port And Protocol VLAN ID TLV. *vlan-id*: Indicates the Port Protocol VLAN ID, ranging from 1 to 4,094.

	vlan-name: Indicates the VLAN Name TLV. *vlan-id*: Indicates the VLAN name, ranging from 1 to 4,094. **dot3-tlv**: Indicates the IEEE 802.3 organizationally specific TLVs. **link-aggregation**: Indicates the Link Aggregation TLV. **mac-physic**: Indicates the MAC/PHY Configuration/Status TLV. **max-frame-size**: Indicates the Maximum Frame Size TLV. **power**: Indicates the Power Via MDI TLV. **med-tlv**: Indicates the LLDP MED TLV. **capability**: Indicates the LLDP-MED Capabilities TLV. **Inventory**: Indicates the inventory management TLV, which contains the hardware version, firmware version, software version, SN, manufacturer name, module name, and asset identifier. **location**: Indicates the Location Identification TLV. **civic-location**: Indicates the civic address information and postal information. **elin**: Indicates the emergency telephone number. *id*: Indicates the policy ID, ranging from 1 to 1,024. **network-policy**: Indicates the Network Policy TLV. *profile-num*: Indicates the Network Policy ID, ranging from 1 to 1,024. **power-over-ethernet**: Indicates the Extended Power-via-MDI TLV.
Command Mode	Interface configuration mode
Usage Guide	N/A

Canceling TLVs

Command	**no lldp tlv-enable {basic-tlv** { **all | port-description | system-capability | system-description | system-name** } **| dot1-tlv** { **all | port-vlan-id | protocol-vlan-id | vlan-name** } **| dot3-tlv** { **all | link-aggregation | mac-physic | max-frame-size | power** } **| med-tlv { all | capability | inventory | location** { **civic-location | elin** } **identifier** *id* **| network-policy profile** [*profile-num*] **| power-over-ethernet** } }
Parameter Description	**basic-tlv**: Indicates the basic management TLV. **port-description**: Indicates the Port Description TLV. **system-capability**: Indicates the System Capabilities TLV. **system-description**: Indicates the System Description TLV. **system-name**: Indicates the System Name TLV. **dot1-tlv**: Indicates the IEEE 802.1 organizationally specific TLVs. **port-vlan-id**: Indicates the Port VLAN ID TLV. **protocol-vlan-id**: Indicates the Port And Protocol VLAN ID TLV. **vlan-name**: Indicates the VLAN Name TLV. **dot3-tlv**: Indicates the IEEE 802.3 organizationally specific TLVs. **link-aggregation**: Indicates the Link Aggregation TLV. **mac-physic**: Indicates the MAC/PHY Configuration/Status TLV. **max-frame-size**: Indicates the Maximum Frame Size TLV. **power**: Indicates the Power Via MDI TLV. **med-tlv**: Indicates the LLDP MED TLV. **capability**: Indicates the LLDP-MED Capabilities TLV. **Inventory**: Indicates the inventory management TLV, which contains the hardware version, firmware version, software

	version, SN, manufacturer name, module name, and asset identifier. **location**: Indicates the Location Identification TLV. **civic-location**: Indicates the civic address information and postal information. **elin**: Indicates the emergency telephone number. *id*: Indicates the policy ID, ranging from 1 to 1,024. **network-policy**: Indicates the Network Policy TLV. *profile-num*: Indicates the Network Policy ID, ranging from 1 to 1,024. **power-over-ethernet**: Indicates the Extended Power-via-MDI TLV.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring TLVs to Be Advertised

Configuration Steps	Cancel the advertisement of the IEEE 802.1 organizationally specific Port And Protocol VLAN ID TLV.
	FS(config)#interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)#no lldp tlv-enable dot1-tlv protocol-vlan-id
Verification	Display LLDP TLV configuration in interface configuration mode.
	FS(config-if-GigabitEthernet 0/1)#show lldp tlv-config interface gigabitethernet 0/1 LLDP tlv-config of port [GigabitEthernet 0/1] NAME STATUS DEFAULT ---------- ------ ------- Basic optional TLV: Port Description TLV YES YES System Name TLV YES YES System Description TLV YES YES System Capabilities TLV YES YES Management Address TLV YES YES IEEE 802.1 extend TLV: Port VLAN ID TLV YES YES Port And Protocol VLAN ID TLV NO YES VLAN Name TLV YES YES IEEE 802.3 extend TLV:

	MAC-Physic TLV YES YES Power via MDI TLV YES YES Link Aggregation TLV YES YES Maximum Frame Size TLV YES YES LLDP-MED extend TLV: Capabilities TLV YES YES Network Policy TLV YES YES Location Identification TLV NO NO Extended Power via MDI TLV YES YES Inventory TLV YES YES

12.4.4 Configures the Management Address to Be Advertised

Configuration Effect

- Configure the management address to be advertised in LLDP packets in interface configuration mode.
- After the management address to be advertised is cancelled, the management address in LLDP packets is subject to the default settings.

Notes

- LLDP runs on physical ports (AP member ports for AP ports). Stacked ports and VSL ports do not support LLDP.

Configuration Steps

- Optional.
- Configure the management address to be advertised in LLDP packets in interface configuration mode.

Verification

Display LLDP information on a local interface

- Check whether the configuration takes effect.

Related Commands

Configuring the Management Address to Be Advertised

Command	**lldp management-address-tlv** [*ip-address*]
Parameter Description	*ip-address*: Indicates the management address to be advertised in an LLDP packet.
Command Mode	Interface configuration mode
Usage Guide	A management address is advertised through LLDP packets by default. The management address is the IPv4 address of the minimum VLAN supported by the port. If no IPv4 address is configured for the VLAN, LLDP keeps searching for

	the qualified IP address. If no IPv4 address is found, LLDP searches for the IPv6 address of the minimum VLAN supported by the port. If no IPv6 address is found, the loopback address 127.0.0.1 is used as the management address.

Canceling the Management Address

Command	**no lldp management-address-tlv**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	A management address is advertised through LLDP packets by default. The management address is the IPv4 address of the minimum VLAN supported by the port. If no IPv4 address is configured for the VLAN, LLDP keeps searching for the qualified IP address. If no IPv4 address is found, LLDP searches for the IPv6 address of the minimum VLAN supported by the port. If no IPv6 address is found, the loopback address 127.0.0.1 is used as the management address.

Configuration Example

Configuring the Management Address to Be Advertised

Configuration Steps	Set the management address to 192.168.1.1 on an interface.
	FS(config)#interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)#lldp management-address-tlv 192.168.1.1
Verification	Display configuration on the interface.
	FS(config-if-GigabitEthernet 0/1)#show lldp local-information interface GigabitEthernet 0/1 Lldp local-information of port [GigabitEthernet 0/1] Port ID type : Interface name Port id : GigabitEthernet 0/1 Port description : GigabitEthernet 0/1 Management address subtype : ipv4 Management address : 192.168.1.1 Interface numbering subtype : ifIndex Interface number : 1 Object identifier : 802.1 organizationally information Port VLAN ID : 1

```
Port and protocol VLAN ID(PPVID)    : 1
        PPVID Supported                    : YES
        PPVID Enabled                      : NO
VLAN name of VLAN 1                        : VLAN0001
Protocol Identity                          :

802.3 organizationally information
Auto-negotiation supported                 : YES
Auto-negotiation enabled                   : YES
PMD auto-negotiation advertised            : 1000BASE-T full duplex mode, 100BASE-TX full duplex mode, 100BASE-TX
half duplex mode, 10BASE-T full duplex mode, 10BASE-T half duplex mode
Operational MAU type                       : speed(100)/duplex(Full)
PoE support                                : NO
Link aggregation supported                 : YES
Link aggregation enabled                   : NO
Aggregation port ID                        : 0
Maximum frame Size                         : 1500

LLDP-MED organizationally information
Power-via-MDI device type                  : PD
Power-via-MDI power source                 : Local
Power-via-MDI power priority               :
Power-via-MDI power value                  :
Model name                                 : Model name
```

12.4.5 Configuring the LLDP Fast Transmission Count

Configuration Effect

- Configure the number of LLDP packets that are fast transmitted.

Configuration Steps

- Optional.
- Configure the number of LLDP packets that are fast transmitted in global configuration mode.

Verification

Displaying the global LLDP status information

- Check whether the configuration takes effect.

Related Commands

Configuring the LLDP Fast Transmission Count

Command	**lldp fast-count** *value*
Parameter Description	*value*: Indicates the number of LLDP packets that are fast transmitted. The value ranges from 1 to 10. The default value is 3.
Command Mode	Global configuration mode
Usage Guide	N/A

Restoring the Default LLDP Fast Transmission Count

Command	**no lldp fast-count**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the LLDP Fast Transmission Count

Configuration Steps	Set the LLDP fast transmission count to 5 in global configuration mode.
	FS(config)#lldp fast-count 5
Verification	Display the global LLDP status information.
	FS(config)#show lldp status Global status of LLDP : Enable Neighbor information last changed time : Transmit interval : 30s Hold multiplier : 4 Reinit delay : 2s Transmit delay : 2s Notification interval : 5s Fast start counts : 5

12.4.6 Configuring the TTL Multiplier and Transmission Interval

Configuration Effect

- Configure the TTL multiplier.
- Configure the LLDP packet transmission interval.

Configuration Steps

- Optional.
- Perform the configuration in global configuration mode.

Verification

Display LLDP status information on an interface

- Check whether the configuration takes effect.

Related Commands

Configuring the TTL Multiplier

Command	**lldp hold-multiplier** *value*
Parameter Description	*value*: Indicates the TLL multiplier. The value ranges from 2 to 10. The default value is 4.
Command Mode	Global configuration mode
Usage Guide	In an LLDP packet. the value of Time To Live TLV is calculated based on the following formula: Time to Live TLV= TTL multiplier x Packet transmission interval + 1. Therefore, you can modify the Time to Live TLV in LLDP packets by configuring the TTL multiplier.

Restoring the Default TTL Multiplier

Command	**no lldp hold-multiplier**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	In an LLDP packet, the value of Time To Live TLV is calculated based on the following formula: Time to Live TLV = TTL multiplier x Packet transmission interval + 1. Therefore, you can modify the Time to Live TLV in LLDP packets by configuring the TTL multiplier.

Configuring the Transmission Interval

Command	**lldp timer tx-interval** *seconds*
Parameter Description	*seconds*: Indicates the LLDP packet transmission interval. The value ranges from 5 to 32,768.
Command Mode	Global configuration mode
Usage Guide	N/A

Restoring the Default Transmission Interval

Command	**no lldp timer tx-interval**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the TTL Multiplier and Transmission Interval

Configuration Steps	Set the TTL multiplier to 3 and the transmission interval to 20 seconds. The TTL of local device information on neighbors is 61 seconds.
	FS(config)#lldp hold-multiplier 3 FS(config)#lldp timer tx-interval 20
Verification	Display the global LLDP status information.
	FS(config)#lldp hold-multiplier 3 FS(config)#lldp timer tx-interval 20 FS(config)#show lldp status Global status of LLDP : Enable Neighbor information last changed time : Transmit interval : 20s Hold multiplier : 3 Reinit delay : 2s Transmit delay : 2s Notification interval : 5s Fast start counts : 3

12.4.7 Configuring the Transmission Delay

Configuration Effect

- Configure the delay time for LLDP packet transmission.

Configuration Steps

- Optional.
- Perform the configuration in global configuration mode.

Verification

Displaying the global LLDP status information

- Check whether the configuration takes effect.

Related Commands

Configuring the Transmission Delay

Command	**lldp timer tx-delay** *seconds*
Parameter Description	*seconds*: Indicates the transmission delay. The value ranges from 1 to 8,192.
Command Mode	Global configuration mode
Usage Guide	When local information of a device changes, the device immediately transmits LLDP packets to its neighbors. Configure the transmission delay to prevent frequent transmission of LLDP packets caused by frequent changes of local information.

Restoring the Default Transmission Delay

Command	**no lldp timer tx-delay**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	When local information of a device changes, the device immediately transmits LLDP packets to its neighbors. Configure the transmission delay to prevent frequent transmission of LLDP packets caused by frequent changes of local information.

Configuration Example

Configuring the Transmission Delay

Configuration Steps	Set the transmission delay to 3 seconds.
	FS(config)#lldp timer tx-delay 3
Verification	Display the global LLDP status information.
	FS(config)#show lldp status Global status of LLDP : Enable Neighbor information last changed time : Transmit interval : 30s Hold multiplier : 4 Reinit delay : 2s Transmit delay : 3s Notification interval : 5s

	Fast start counts : 3

12.4.8 Configuring the Initialization Delay

Configuration Effect

- Configure the delay time for LLDP to initialize on any interface.

Configuration Steps

- Optional.
- Configure the delay time for LLDP to initialize on any interface.

Verification

Display the global LLDP status information

- Check whether the configuration takes effect.

Related Commands

Configuring the Initialization Delay

Command	**lldp timer reinit-delay** *seconds*
Parameter Description	*seconds*: Indicates the initialization delay . The value ranges from 1 to 10 seconds.
Command Mode	Global configuration mode
Usage Guide	Configure the initialization delay to prevent frequent initialization of the state machine caused by frequent changes of the port work mode.

Restoring the Default Initialization Delay

Command	**no lldp timer reinit-delay**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure the initialization delay to prevent frequent initialization of the state machine caused by frequent changes of the port work mode.

Configuration Example

Configuring the Initialization Delay

Configuration Steps	Set the initialization delay to 3 seconds.
	FS(config)#lldp timer reinit-delay 3
Verification	Display the global LLDP status information.

	FS(config)#show lldp status Global status of LLDP : Enable Neighbor information last changed time : Transmit interval : 30s Hold multiplier : 4 Reinit delay : 3s Transmit delay : 2s Notification interval : 5s Fast start counts : 3

12.4.9 Configuring the LLDP Trap Function

Configuration Effect

- Configure the interval for transmitting LLDP Trap messages.

Configuration Steps

Enabling the LLDP Trap Function

- Optional.
- Perform the configuration in interface configuration mode.

Configuring the LLDP Trap Transmission Interval

- Optional.
- Perform the configuration in global configuration mode.

Verification

Display LLDP status information

- Check whether the LLDP Trap function is enabled.
- Check whether the interval configuration takes effect.

Related Commands

Enabling the LLDP Trap Function

Command	**lldp notification remote-change enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The LLDP Trap function enables a device to send its local LLDP information (such as neighbor discovery and communication link fault) to the NMS server so that administrators learn about the network performance

Disabling the LLDP Trap Function

Command	**no lldp notification remote-change enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The LLDP Trap function enables a device to send its local LLDP information (such as neighbor discovery and communication link fault) to the NMS server so that administrators learn about the network performance.

Configuring the LLDP Trap Transmission Interval

Command	**lldp timer notification-interval** *seconds*
Parameter Description	*seconds*: Indicates the interval for transmitting LLDP Trap messages. The value ranges from 5 to 3,600 seconds. The default value is 5 seconds.
Command Mode	Global configuration mode
Usage Guide	Configure the LLDP Trap transmission interval to prevent frequent transmission of LLDP Trap messages. LLDP changes detected within this interval will be transmitted to the NMS server.

Restoring the LLDP Trap Transmission Interval

Command	**no lldp timer notification-interval**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure the LLDP Trap transmission interval to prevent frequent transmission of LLDP Trap messages. LLDP changes detected within this interval will be transmitted to the NMS server.

Configuration Example

Enabling the LLDP Trap Function and Configuring the LLDP Trap Transmission Interval

Configuration Steps	Enable the LLDP Trap function and set the LLDP Trap transmission interval to 10 seconds.
	FS(config)#lldp timer notification-interval 10 FS(config)#interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)#lldp notification remote-change enable
Verification	Display LLDP status information.
	FS(config-if-GigabitEthernet 0/1)#show lldp status Global status of LLDP : Enable Neighbor information last changed time : Transmit interval : 30s

```
Hold multiplier                          : 4
Reinit delay                             : 2s
Transmit delay                           : 2s
Notification interval                    : 10s
Fast start counts                        : 3
-----------------------------------------------------------
Port [GigabitEthernet 0/1]
-----------------------------------------------------------
Port status of LLDP                      : Enable
Port state                               : UP
Port encapsulation                       : Ethernet II
Operational mode                         : RxAndTx
Notification enable                      : YES
Error detect enable                      : YES
Number of neighbors                      : 0
Number of MED neighbors                  : 0
```

12.4.10 Configuring the LLDP Error Detection Function

Configuration Effect

- Enable the LLDP error detection function. When LLDP detects an error, the error is logged.
- Configure the LLDP error detection function to detect VLAN configuration at both ends of a link, port status, aggregate port configuration, MTU configuration, and loops.

Notes

N/A

Configuration Steps

- Optional.
- Enable or disable the LLDP error detection function in interface configuration mode.

Verification

Display LLDP status information on an interface

- Check whether the configuration takes effect.

Related Commands

Enabling the LLDP Error Detection Function

Command	**lldp error-detect**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The LLDP error detection function relies on specific TLVs in LLDP packets exchanged between devices at both ends of a link. Therefore, a device needs to advertise correct TLVs to ensure the LLDP error detection function.

Disabling the LLDP Error Detection Function

Command	**no lldp error-detect**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The LLDP error detection function relies on specific TLVs in LLDP packets exchanged between devices at both ends of a link. Therefore, a device needs to advertise correct TLVs to ensure the LLDP error detection function.

Configuration Example

Enabling the LLDP Error Detection Function

Configuration Steps	Enable the LLDP error detection function on interface GigabitEthernet 0/1.
	FS(config)#interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)#lldp error-detect
Verification	Display LLDP status information on the interface.
	FS(config-if-GigabitEthernet 0/1)#show lldp status interface gigabitethernet 0/1 Port [GigabitEthernet 0/1] Port status of LLDP : Enable Port state : UP Port encapsulation : Ethernet II Operational mode : RxAndTx Notification enable : NO Error detect enable : YES Number of neighbors : 0 Number of MED neighbors : 0

12.4.11 Configuring the LLDP Encapsulation Format

Configuration Effect

- Configure the LLDP encapsulation format.

Configuration Steps

- Optional.
- Configure the LLDP encapsulation format on an interface.

Verification

Display LLDP status information of an interface

- Check whether the configuration takes effect.

Related Commands

Setting the LLDP Encapsulation Format to SNAP

Command	**lldp encapsulation snap**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The LLDP encapsulation format configuration on a device and its neighbors must be consistent.

Restoring the Default LLDP Encapsulation Format (Ethernet II)

Command	**No lldp encapsulation snap**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The LLDP encapsulation format configuration on a device and its neighbors must be consistent.

Configuration Example

Setting the LLDP Encapsulation Format to SNAP

Configuration Steps	Set the LLDP encapsulation format to SNAP.
	FS(config)#interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)#lldp encapsulation snap
Verification	Display LLDP status information on the interface.
	FS(config-if-GigabitEthernet 0/1)#show lldp status interface gigabitethernet 0/1

	Port [GigabitEthernet 0/1] Port status of LLDP : Enable Port state : UP Port encapsulation : Snap Operational mode : RxAndTx Notification enable : NO Error detect enable : YES Number of neighbors : 0 Number of MED neighbors : 0

12.4.12 Configuring the LLDP Network Policy

Configuration Effect

- Configure the LLDP Network Policy.
- If a device is connected to an IP-Phone that supports LLDP-MED, you can configure the Network Policy TLV to push policy configuration to the IP-Phone, , which enables the IP-Phone to change the tag and QoS of voice streams. In addition to the LLDP Network Policy, perform the following steps on the device: 1. Enable the Voice VLAN function and add the port connected to the IP-Phone to the Voice VLAN. 2. Configure the port connected to the IP-Phone as a QoS trusted port (the trusted DSCP mode is recommended). 3. If 802.1X authentication is also enabled on the port, configure a secure channel for the packets from the Voice VLAN. If the IP-Phone does not support LLDP-MED, enable the voice VLAN function and add the MAC address of the IP-Phone to the Voice VLAN OUI list manually.
- For the configuration of the QoS trust mode, see *Configuring IP QoS*; for the configuration of the Voice VLAN, see *Configuring Voice VLAN*; for the configuration of the secure channel, see *Configuring ACL*.

Configuration Steps

- Optional.
- Configure the LLDP Network Policy.

Verification

Displaying the LLDP network policy configuration.

- Check whether the configuration takes effect.

Related Commands

Configuring the LLDP Network Policy

Command	**lldp network-policy profile** *profile-num*
Parameter Description	*profile-num*: Indicates the ID of an LLDP Network Policy. The value ranges from 1 to 1,024.
Command Mode	Global configuration mode

Usage Guide	Run this command to enter the LLDP network policy mode after specifying a policy ID. After entering the LLDP network policy mode, run the **{ voice | voice-signaling } vlan** command to configure a specific network policy.

Deleting the LLDP Network Policy

Command	**no lldp network-policy profile** *profile-num*
Parameter Description	*profile-num*: Indicates the LLDP Network Policy ID. The value ranges from 1 to 1,024.
Command Mode	Interface configuration mode
Usage Guide	Run this command to enter the LLDP network policy mode after specifying a policy ID. After entering the LLDP network policy mode, run the **{ voice | voice-signaling } vlan** command to configure a specific network policy.

Configuration Example

Configuring the LLDP Network Policy

Configuration Steps	Set the Network Policy TLV to 1 for LLDP packets to be advertised by port GigabitEthernet 0/1 and set the VLAN ID of the Voice application to 3, COS to 4, and DSCP to 6.
	FS#config FS(config)#lldp network-policy profile 1 FS(config-lldp-network-policy)# voice vlan 3 cos 4 FS(config-lldp-network-policy)# voice vlan 3 dscp 6 FS(config-lldp-network-policy)#exit FS(config)# interface gigabitethernet 0/1 FS(config-if-GigabitEthernet 0/1)# lldp tlv-enable med-tlv network-policy profile 1
Verification	Display the LLDP network policy configuration on the local device.
	network-policy information: ---------- network policy profile :1 voice vlan 3 cos 4 voice vlan 3 dscp 6

12.4.13 Configuring the Civic Address

Configuration Effect

- Configure the civic address of a device.

Configuration Steps

- Optional.
- Perform this configuration in LLDP Civic Address configuration mode.

Verification

Display the LLDP civic address of the local device

- Check whether the configuration takes effect.

Related Commands

Configuring the Civic Address of a Device

Command	Configure the LLDP civic address. Use the **no** option to delete the address. { **country | state | county | city | division | neighborhood | street-group | leading-street-dir | trailing-street-suffix | street-suffix | number | street-number-suffix | landmark | additional-location-information | name | postal-code | building | unit | floor | room | type-of-place | postal-community-name | post-office-box | additional-code** } *ca-word*
Parameter Description	**country**: Indicates the country code, with two characters. CH indicates China. **state**: Indicates the CA type is 1. **county**: Indicates that the CA type is 2. **city**: Indicates that the CA type is 3. **division**: Indicates that the CA type is 4. **neighborhood**: Indicates that the CA type is 5. **street-group**: Indicates that the CA type is 6. **leading-street-dir**: Indicates that the CA type is 16. **trailing-street-suffix**: Indicates that the CA type is 17. **street-suffix**: Indicates that the CA type is 18. **number**: Indicates that the CA type is 19. **street-number-suffix**: Indicates that the CA type is 20. **landmark**: Indicates that the CA type is 21. **additional-location-information**: Indicates that the CA type is 22. **name**: Indicates that the CA type is 23. **postal-code**: Indicates that the CA type is 24. **building**: Indicates that the CA type is 25. **unit**: Indicates that the CA type is 26. **floor**: Indicates that the CA type is 27. **room**: Indicates that the CA type is 28. **type-of-place**: Indicates that the CA type is 29. **postal-community-name**: Indicates that the CA type is 30. **post-office-box**: Indicates that the CA type is 31. **additional-code**: Indicates that the CA type is 32. *ca-word*: Indicates the address.
Command Mode	LLDP Civic Address configuration mode
Usage Guide	After entering the LLDP Civic Address configuration mode, configure the LLDP civic address.

Deleting the Civic Address of a Device

Command	**no** { **country** | **state** | **county** | **city** | **division** | **neighborhood** | **street-group** | **leading-street-dir** | **trailing-street-suffix** | **street-suffix** | **number** | **street-number-suffix** | **landmark** | **additional-location-information** | **name** | **postal-code** | **building** | **unit** | **floor** | **room** | **type-of-place** | **postal-community-name** | **post-office-box** | **additional-code** }
Parameter Description	N/A
Command Mode	LLDP Civic Address configuration mode
Usage Guide	After entering the LLDP Civic Address configuration mode, configure the LLDP civic address.

Configuring the Device Type

Command	**device-type** *device-type*
Parameter Description	*device-type*: Indicates the device type. The value ranges from 0 to 2. The default value is 1. 0 indicates that the device type is DHCP server. 1 indicates that the device type is switch. 2 indicates that the device type is LLDP MED .
Command Mode	LLDP Civic Address configuration mode
Usage Guide	After entering the LLDP Civic Address configuration mode, configure the device type.

Restoring the Device Type

Command	**no device-type**
Parameter Description	N/A
Command Mode	LLDP Civic Address configuration mode
Usage Guide	After entering the LLDP Civic Address configuration mode, restore the default settings.

Configuration Example

Configuring the Civic Address of a Device

Configuration Steps	Set the address of port GigabitEthernet 0/1 as follows: set country to CH, city to Fuzhou, and postal code to 350000.
	FS#config FS(config)#lldp location civic-location identifier 1 FS(config-lldp-civic)# country CH FS(config-lldp-civic)# city Fuzhou FS(config-lldp-civic)# postal-code 350000
Verification	Display the LLDP civic address of port GigabitEthernet 0/1 1.

	civic location information: ---------- Identifier :1 country :CH device type :1 city :Fuzhou postal-code :350000

12.4.14 Configuring the Emergency Telephone Number

Configuration Effect

- Configure the emergency telephone number of a device.

Configuration Steps

- Optional.
- Perform this configuration in global configuration mode.

Verification

Display the emergency telephone number of the local device

- Check whether the configuration takes effect.

Related Commands

Configuring the Emergency Telephone Number of a Device

Command	**lldp location elin identifier** *id* **elin-location** *tel-number*
Parameter Description	*id*: Indicates the identifier of an emergency telephone number. The value ranges from 1 to 1,024. *tel-number*: Indicates emergency telephone number, containing 10-25 characters.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure the emergency telephone number.

Deleting the Emergency Telephone Number of a Device

Command	**no lldp location elin identifier** *id*
Parameter Description	*id*: Indicates the identifier of an emergency telephone number. The value ranges from 1 to 1,024.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Emergency Telephone Number of a Device

Configuration Steps	Set the emergency telephone number of port GigabitEthernet 0/1 to 085285555556.
	FS#config FS(config)#lldp location elin identifier 1 elin-location 085283671111
Verification	Display the emergency telephone number of port GigabitEthernet 0/1.
	elin location information: ---------- Identifier :1 elin number :085283671111

12.4.15 Configuring Compatible Neighbors Detection

Configuration Effect

- Enables compatible neighbors detection function.

Configuration Steps

- Optional.
- Perform this configuration in global configuration mode.

Verification

Display the LLDP information.

- Check whether the configuration takes effect.

Related Commands

Enabling compatible neighbors detection

Command	**lldp compliance vendor**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Disabling compatible neighbors detection

Command	**no lldp compliance vendor**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring Compatible Neighbors Detection

Configuration Steps	Configuring compatible neighbor detection.
	FS(config)# lldp compliance vendor
Verification	Display the LLDP information.
	FS(config)#show lldp status Global status of LLDP : Enable Global vendor compliance : YES

12.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears LLDP statistics.	**clear lldp statistics** [**interface** *interface-name*]
Clears LLDP neighbor information.	**clear lldp table** [**interface** *interface-name*]

Displaying

Description	Command
Displays LLDP information on the local device, which will be organized as TLVs and sent to neighbors.	**show lldp local-information** [**global** | **interface** interface-name]
Displays the LLDP civic address or emergency telephone number of a local device.	**show lldp location** { **civic-location** | **elin-location** } { **identifier** *id* | **interface** *interface-name* | **static** }
Displays LLDP information on a neighbor.	**show lldp neighbors [interface** *interface-name*] [**detail**]
Displays the LLDP network policy configuration of the local device.	**show lldp network-policy** { **profile** [*profile-num*] | **interface** *interface-name* }
Displays LLDP statistics.	**show lldp statistics** [**global** | **interface** *interface-name*]
Displays LLDP status information.	**show lldp status** [**interface** *interface-name*]
Displays the configuration of TLVs to be advertised by a port.	**show lldp tlv-config** [**interface** *interface-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs LLDP error processing.	**debug lldp error**

Debugs LLDP event processing.	**debug lldp event**
Debugs LLDP hot backup processing.	**debug lldp ha**
Debugs the LLDP packet reception.	**debug lldp packet**
Debugs the LLDP state machine.	**debug lldp stm**

13 Configuring QinQ

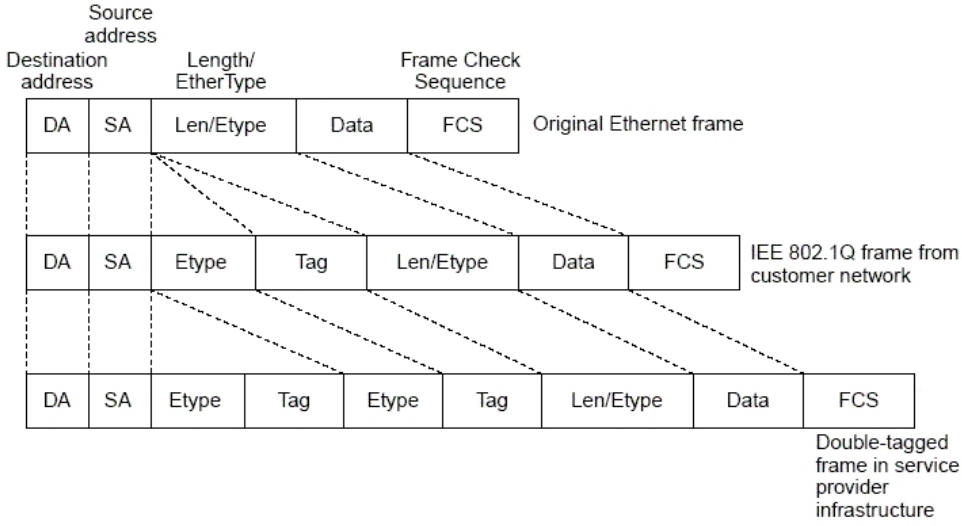
13.1 Overview

QinQ is used to insert a public virtual local area network (VLAN) tag into a packet with a private VLAN tag to allow the double-tagged packet to be transmitted over a service provider (SP) network.

Users on a metropolitan area network (MAN) must be separated by VLANs. IEEE 802.1Q supports only 4,094 VLANs, far from enough. Through the double-tag encapsulation provided by QinQ, a packet is transmitted over the SP network based on the unique outer VLAN tag assigned by the public network. In this way, private VLANs can be reused, which increases the number of available VLAN tags and provides a simple Layer-2 virtual private network (VPN) feature.

Figure 12-1 shows the double-tag insertion process. The entrance to an SP network is called a dot1q-tunnel port, or Tunnel port for short. All frames entering provider edges (PEs) are considered untagged. All tags, whether untagged frames or frames with customer VLAN tags, are encapsulated with the tags of the SP network. The VLAN ID of the SP network is the ID of the default VLAN for the Tunnel port.

Figure 12-1 Outer Tag Encapsulation



Protocols and Standards

- IEEE 802.1ad

13.2 Applications

Application	Description
Implementing Layer-2 VPN Through Port-Based Basic QinQ	Data is transmitted from Customer A and Customer B to the peer end without conflict on the SP network even if the data comes from the same VLAN.
Implementing Layer-2 VPN and Service Flow Management Through C-TAG-Based Selective QinQ	Outer tags are inserted into frames flexibly based on different customer VLANs to achieve Layer-2 VPN, segregate service flows (e.g., broadband Internet access and IPTV), and implement various QoS policies. Customer tag (C-TAG)-based QinQ is more flexible than port-based QinQ.
Implementing Layer-2 VPN and Service Flow Management Through ACL-Based Selective QinQ	The different service flows, such as broadband Internet access and IPTV, are segregated based on access control lists (ACLs). Different QoS policies are applied to service flows through selective QinQ.

Application	Description
Implementing VLAN Aggregation for Different Services Through VLAN Mapping	Different service flows (PC, IPTV, and VoIP) are transmitted through different VLANs. The VLANs are aggregated on a campus network so that only one VLAN is used to carry the same service flows, thus saving VLAN resources.
Implementing QinQ-Based Layer-2 Transparent Transmission	Customer Network A and Customer Network B in different areas can perform unified Multiple Spanning Tree Protocol (MSTP) calculation or VLAN deployment across the SP network without affecting the SP network.

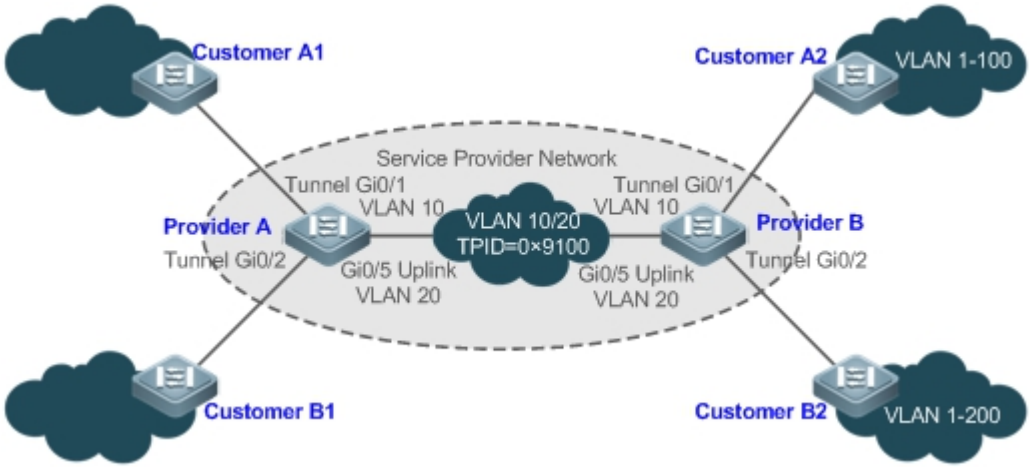
13.2.1 Implementing Layer-2 VPN Through Port-Based Basic QinQ

Scenario

An SP provides the VPN service to Customer A and Customer B.

- Customer A and Customer B belong to different VLANs on the SP network and achieve communication through respective SP VLANs.
- The VLANs of Customer A and Customer B are transparent to the SP network. The VLANs can be reused without conflicts.
- The Tunnel port encapsulates a native VLAN tag in each packet. Packets are transmitted through the native VLAN over the SP network without impact on the VLANs of Customer A and Customer B, thus implementing simple Layer-2 VPN.

Figure 12-2



Remarks	Customer A1 and Customer A2 are the customer edges (CEs) for Customer A network. Customer B1 and Customer B2 are the CEs for Customer B network. Provider A and Provider B are the PEs on the SP network. Customer A and Customer B access the SP network through Provider A and Provider B. The VLAN of Customer A ranges from 1 to100. The VLAN of Customer B ranges from 1 to 200.

Deployment

- Enable basic QinQ on PEs to implement Layer-2 VPN.
- The tag protocol identifiers (TPIDs) used by many switches (including FS switches) are set to 0x8100, but the switches of some vendors do not use 0x8100. In the latter case, you need to change the TPID value on the Uplink ports of PEs to the values of the TPIDs used by third-party switches.

- Configure priority replication and priority mapping for class of service (CoS) on the Tunnel ports of PEs, and configure different QoS policies for different service flows (for details, see *Configuring QoS*).

13.2.2 Implementing Layer-2 VPN and Service Flow Management Through C-TAG-Based Selective QinQ

Scenario

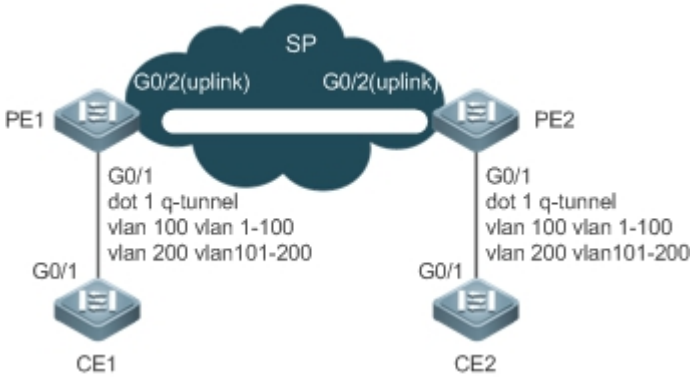
Basic QinQ encapsulates an outer tag of the native VLAN in a packet. That is, the encapsulation of outer tags depends on the native VLAN on Tunnel ports. Selective QinQ encapsulates an outer tag in a packet based on its inner tag to implement VPN transparent transmission and apply QoS policies flexibly.

- Broadband Internet access and IPTV are important services carried by MANs. The SPs manage different service flows through different VLANs and provides QoS policies for the VLANs or CoS. You can enable C-TAG-based QinQ on PEs to encapsulate outer VLAN tags in the service flows to achieve transparent transmission based on the QoS policies of the SP network.
- Important services and regular services are separated within different VLAN ranges. The customer can transmit service flows transparently over an SP network through C-TAG-based selective QinQ and ensure preferential transmission of important service flows by using the QoS policies of the SP network.

In Figure 12-3, the CEs are aggregated by the floor switches inside residential buildings. The broadband Internet access and IPTV services are segregated by VLANs with different QoS policies.

- The service flows of broadband Internet access and IPTV are transmitted transparently by different VLANs over the SP network.
- The SP network provides QoS policies based on VLANs or CoS. On the PEs, you can encapsulate an outer tag in the service flow based on its inner VLAN tag or set a CoS to ensure preferential transmission of service flows over the SP network.
- The CoS values of service packets can be changed through priority mapping or replication so that the QoS policies of the SP network are applied flexibly.

Figure 12-3



Remarks	CE 1 and CE 2 access the SP network through PE1 and PE2. On CE 1 and CE 2, the broadband Internet access flows are transmitted through VLAN 1–100, and IPTV flows are transmitted through VLAN 101–200. PE 1 and PE 2 are configured with Tunnel ports and VLAN mappings to segregate service flows.

Deployment

- Configure C-TAG-based selective QinQ on the ports (G0/1) of PE 1 and PE 2 connected to CE 1 and CE 2 respectively to realize the segregation and transparent transmission of service flows.
- If the SP network provides QoS policies based on VLANs or CoS, you can encapsulate an outer tag in the service flow based on its

inner tag or set a CoS through priority replication or mapping on PE 1 and PE 2 to ensure preferential transmission of service flows over the SP network.

13.2.3 Implementing Layer-2 VPN and Service Flow Management Through ACL-Based Selective QinQ

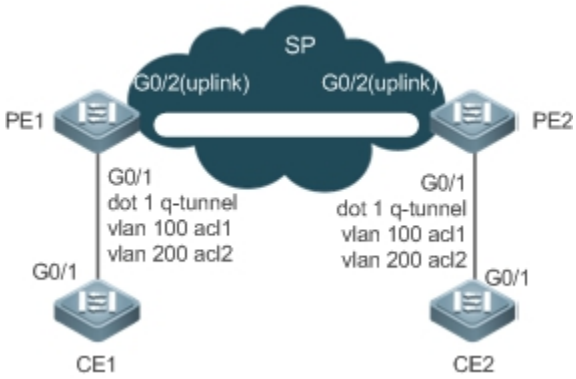
Scenario

The service flows from the customer network may be classified by MAC address, IP address, or protocol type, instead of by VLAN. The customer network may contain many low-end access devices unable to segregate service flows by VLAN IDs. In the preceding two situations, the packets from the customer network cannot be encapsulated with outer tags based on their inner tags to realize transparent transmission and implement QoS policies. Service flows may be classified by MAC address, IP address, or protocol type through ACLs. Selective QinQ uses ACLs to segregate service flows and add or modify outer tags in order to implement Layer-2 VPN and QoS policies based on different service flows.

In Figure 12-4, different VLANs are configured on PE 1 and PE 2 to transmit different service flows classified through ACLs. If the SP network provides QoS policies based on different services, certain services can be transmitted preferentially.

- Outer VLAN tags are encapsulated based on different service flows. The service flows of a customer network can be transmitted transparently, and its branch offices can access each other.
- The SP network provides QoS policies based on the VLAN tags or CoS values to ensure preferential transmission of certain service flows.

Figure 12-4



Remarks	CE 1 and CE 2 access the SP network through PE1 and PE2. PE 1 and PE 2 classify flows based on ACLs: ACL 1 matches the Point-to-Point Protocol over Ethernet (PPPoE) flows, and ACL 2 matches the IPTV flows. PE 1 and PE 2 are configured with Tunnel ports, as well as outer tag encapsulation policies applicable to service flows recognized by different ACLs.

Deployment

- Configure ACLs on PE 1 and PE 2 to segregate service flows.
- Configure ACL-based selective QinQ on the ports (G0/1) of PE 1 and PE 2 connected to CE 1 and CE 2 respectively to realize the segregation and transparent transmission of service flows.
- If the SP network provides QoS policies based on VLANs or CoS, you can encapsulate an outer tag in the service flow based on its inner tag or set a CoS through priority replication or mapping on PE 1 and PE 2 to ensure preferential transmission of service flows over the SP network.

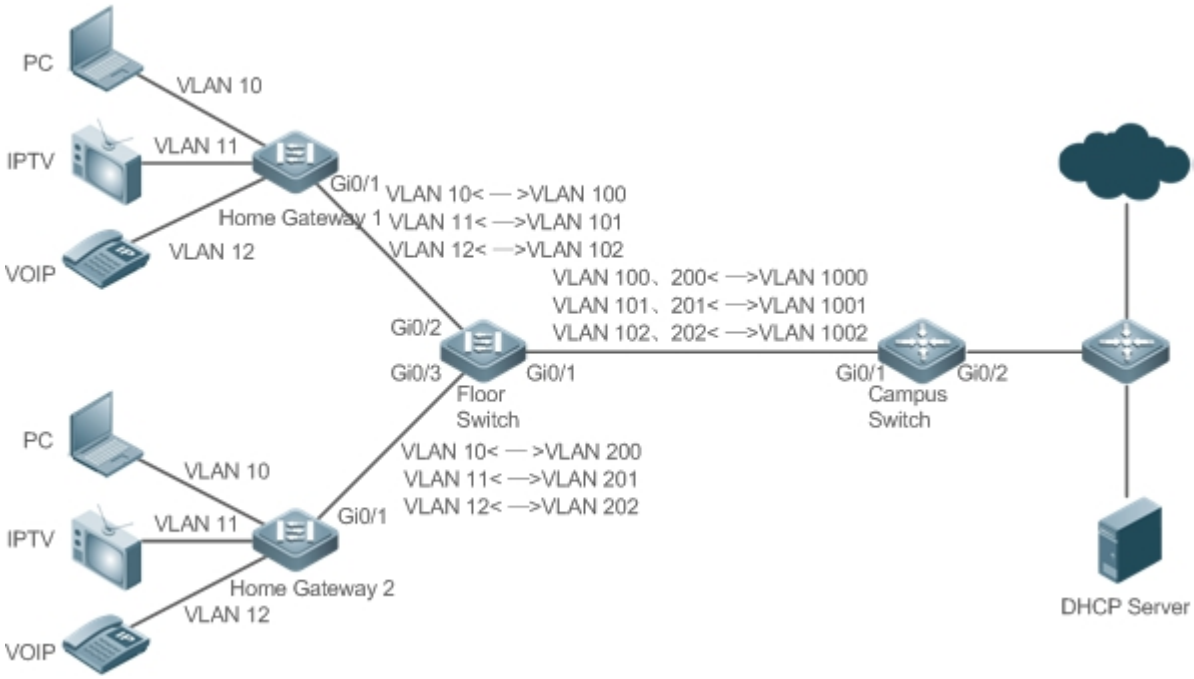
13.2.4 Implementing VLAN Aggregation for Different Services Through VLAN Mapping

Scenario

The different service flows of different users are segregated on a campus network.

- The different service flows are transmitted through different VLANs on the home gateway.
- The same service flows from different users are segregated on the floor switch.
- The same service flows from different users are sent by a campus switch through one single VLAN.

Figure 12-5



Remarks	PC, IPTV, and VoIP are different user services. Switch A and Switch B are the gateway devices of different users. Switch C is a floor switch. Switch D is a campus switch.

Deployment

- On the home gateway devices, configure VLANs for different services to segregate service flows. For example, configure VLAN 10 for the PC service, VLAN 11 for IPTV, and VLAN 12 for VoIP.
- On the ports of the floor switch (Switch D) connected to the home gateway devices, configure VLAN mapping to segregate the service flows of different users.
- On the campus switch, configure VLAN mapping to segregate the service flows.
- Through the preceding deployment, the different service flows of different users are segregated.

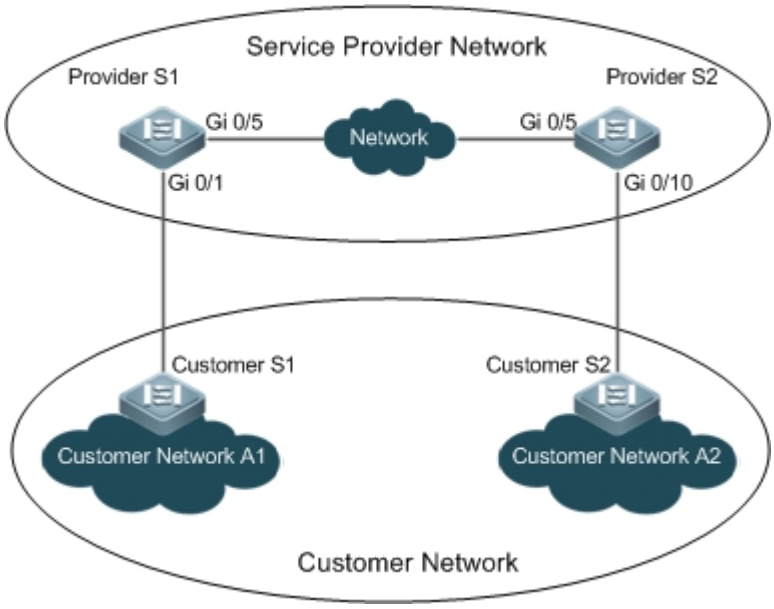
13.2.5 Implementing QinQ-Based Layer-2 Transparent Transmission

Scenario

The Layer-2 transparent transmission between customer networks has no impact on the SP network.

- The Layer-2 packets on customer networks are transparent to SP networks and can be transmitted between the customer networks without impact on the SP networks.

Figure 12-6



Remarks	Customer S1 and Customer S2 access the SP network through Provider S1 and Provider S2. Provider S1 and Provider S2 are enabled with Layer-2 transparent transmission globally, and the Gi 0/1 and Gi 0/10 ports are enabled with Layer-2 transparent transmission.

Deployment

- On the ports of the PEs (Provider S1 and Provider S2) connected to Customer S1 and Customer S2 respectively, configure Layer-2 transparent transmission between Customer Network A1 and Customer Network A2 without impact on the SP network.
- Configure STP transparent transmission based on user requirements to realize transparent transmission of bridge protocol data unit (BPDU) packets between Customer Network A1 and Customer Network A2 and to perform unified MSTP calculation across the SP network.
- Configure GARP VLAN Registration Protocol (GVRP) transparent transmission based on user requirements to realize transparent transmission of GVRP packets between Customer Network A1 and Customer Network A2 and dynamic VLAN configuration on the customer networks across the SP network.

13.3 Features

Basic Concepts

Basic QinQ

Configure basic QinQ on a Tunnel port and configure a native VLAN for the port. Packets entering the port are encapsulated with outer tags containing the native VLAN ID. Basic QinQ does not segregate service flows and cannot encapsulate packets flexibly based on VLANs.

Selective QinQ

Selective QinQ is classified into two types: selective QinQ based on C-TAGs and selective QinQ based on ACLs.

In C-TAG-based selective QinQ, outer tags are encapsulated in packets based on the inner tags to segregate service flows and realize

transparent transmission.

In ACL-based selective QinQ, outer tags are encapsulated in packets based on the ACLs to segregate service flows.

TPID

An Ethernet frame tag consists of four fields: TPID, User Priority, Canonical Format Indicator (CFI), and VLAN ID.

By default, the TPID is 0x8100 according to IEEE802.1Q. On the switches of some vendors, the TPID is set to 0x9100 or other values. The TPID configuration aims to ensure that the TPIDs of packets to be forwarded are compatible with the TPIDs supported by third-party switches.

Priority Mapping and Priority Replication

The default value of User Priority in Ethernet frame tags is 0, indicating regular flows. You can set this field to ensure preferential transmission of certain packets. You can specify User Priority by setting the value of CoS in a QoS policy.

Priority replication: If the SP network provides a QoS policy corresponding to a specified CoS in the inner tag, you can replicate the CoS of the inner tag to the outer tag to enable transparent transmission based on the QoS policy provided by the SP network.

Priority mapping: If the SP network provides various QoS policies corresponding to specified CoS values for different service flows, you can map the CoS value of the inner tag to the CoS value of the outer tag to ensure preferential transmission of service flows based on the QoS policies provided by the SP network.

Layer-2 Transparent Transmission

STP and GVRP packets may affect the topology of the SP network. If you want to unify the topology of two customer networks separated by the SP network without affecting the SP network topology, transmit the STP and GVRP packets from the customer networks over the SP network transparently.

Overview

Feature	Description
Basic QinQ	Configures the Tunnel port and specifies whether packets sent from the port are tagged.
Selective QinQ	Encapsulates different outer tags in data flows based on ACLs.
VLAN Mapping	Replaces the inner tags of packets with outer tags, and then restores the outer tags to inner tags based on the same rules.
TPID Configuration	By default, the TPID is 0x8100 according to IEEE802.1Q. On the switches of some vendors, the TPIDs of outer tags are set to 0x9100 or other values. The TPID configuration aims to ensure that the TPIDs of packets to be forwarded are compatible with the TPIDs supported by third-party switches.
MAC Address Replication	In ACL-based selective QinQ, the VLAN IDs for the MAC addresses that switches learn belong to the native VLAN. If VLAN conversion is implemented based on ACLs, upon receiving packets from the peer end, the local end may fail to query MAC addresses, causing a flood. To address this problem, MAC address replication is provided to replicate the MAC addresses of the native VLAN to the VLAN where the outer tag is located.
Layer-2 Transparent Transmission	Transmits Layer-2 packets between customer networks without impact on SP networks.
Priority Replication	If the SP network provides a QoS policy corresponding to a specified CoS value in the inner tag, you can replicate the CoS of the inner tag to the outer tag to enable transparent transmission based on the QoS policy provided by the SP network.
Priority Mapping	If the SP network provides various QoS policies corresponding to specified CoS values for different service flows, you can map the CoS value of the inner tag to the CoS value of the outer tag to ensure preferential transmission of service flows based on the QoS policies provided by the SP network.

13.3.1 Basic QinQ

Basic QinQ can be used to implement simple Layer-2 VPN, but it lacks flexibility in encapsulating outer tags.

Working Principle

After a Tunnel port receives a packet, the switch adds the outer tag containing the default VLAN ID to the packet. If the received packet already carries a VLAN tag, it is encapsulated as a double-tagged packet. If it does not have a VLAN tag, it is added with the VLAN tag containing the default VLAN ID.

13.3.2 Selective QinQ

Selective QinQ adds different outer tags to data flows flexibly.

Working Principle

Selective QinQ can be used to encapsulate different outer tags based on inner tags, MAC addresses, protocol numbers, source addresses, destination addresses, priorities, or the port numbers of applications. In this way, packets of different users, services, and priorities are encapsulated with different outer VLAN tags.

You can configure the following selective QinQ policies:

- Add an outer VLAN tag based on the inner VLAN tag.
- Modify an outer VLAN tag based on the outer VLAN tag.
- Modify an outer VLAN tag based on the inner VLAN tag.
- Modify an outer VLAN tag based on the inner and outer VLAN tags.
- Add an outer VLAN tag based on the ACL.
- Modify an outer VLAN tag based on the ACL.
- Modify an inner VLAN tag based on the ACL.

13.3.3 VLAN Mapping

Working Principle

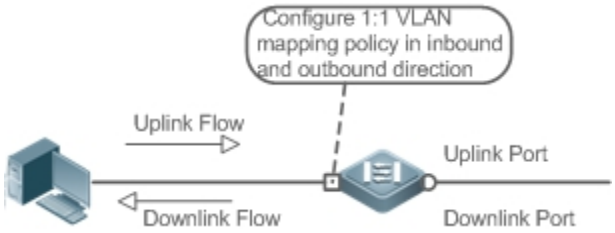
The inner tag of a packet is replaced by an outer tag to allow the packet to be transmitted based on the public network topology. When the packet is transmitted to the customer network, the outer tag is restored to the original inner tag based on the same rule. VLAN mapping supports the following mapping rule:

- 1:1 VLAN mapping: Changes a VLAN ID to the specified VLAN ID.

↘ 1:1 VLAN Mapping Mode 1

1:1 VLAN mapping is mainly applied on floor switches to use different VLANs to carry the same services from different users, as shown in Figure 12-7.

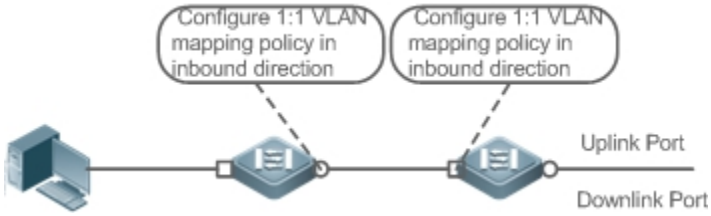
Figure 12-7



- Configure the Downlink port with a VLAN mapping policy in the inbound direction to map the inner tag of the uplink flow to an outer tag.
- Configure the Uplink port with a VLAN mapping policy in the outbound direction to map the outer tag of the downlink flow to the original inner tag.

1:1 VLAN Mapping Mode 2

Figure 12-8



- Configure the Downlink port with a VLAN mapping policy in the inbound direction to map the inner tag of the uplink flow to an outer tag.
- For downstream data flows, Configure the Uplink port with a VLAN mapping policy in the inbound direction to the outer tag of the downlink flow to the original inner tag.

13.3.4 TPID Configuration

Working Principle

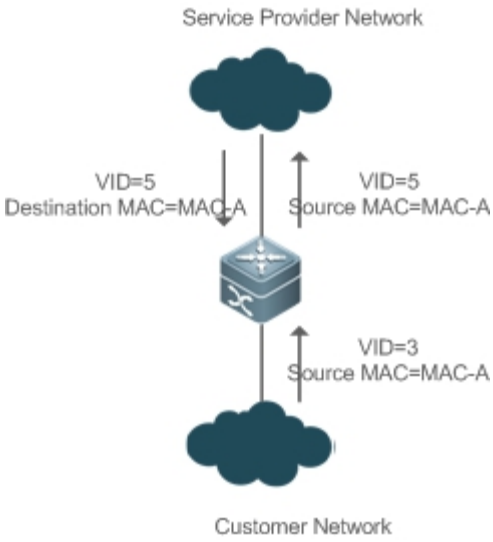
An Ethernet frame tag consists of four fields, namely, TPID, User Priority, CFI, and VLAN ID. By default, the TPID is 0x8100 according to IEEE802.1Q. On the switches of some vendors, the TPIDs of outer tags are set to 0x9100 or other values. The TPID configuration feature allows you to configure TPIDs on ports, which will replace the TPIDs of the outer VLAN tags in packets with the configured TPIDs to realize TPID compatibility.

13.3.5 MAC Address Replication

Working Principle

In ACL-based selective QinQ, the MAC address learned by a switch belongs to the native VLAN. The Tunnel port tags the packet with the specified outer VLAN ID based on the selective QinQ policy. Upon receiving a reply packet containing the same outer VLAN tag, the Tunnel port fails to find the MAC address in the outer VLAN as it is in the native VLAN, causing a flood.

Figure 12-9



As in Figure 12-9, the customer network is connected to the Tunnel port of the switch. Configured with native VLAN 4, the Tunnel port tags the packet whose source MAC address is A with outer VLAN 5. Upon receiving a packet with inner tag VLAN 3 and source MAC address A, the switch tags the packet with outer VLAN 5. Because the port is configured with native VLAN 4, MAC address A is learned by VLAN 4. Upon receiving the reply packet, the switch looks for MAC address A on VLAN 5 because the outer tag of the packet contains VLAN ID 5. However, MAC address A is not learned by VLAN 5, causing floods.

You can configure the Tunnel port to replicate the MAC address of the native VLAN to the outer VLAN to avoid continuous flooding of the packets from the SP network. You can also configure the Tunnel port to replicate the MAC address of the outer VLAN for the outer tag to the native VLAN to avoid continuous flooding of the packets from the customer network.

13.3.6 Layer-2 Transparent Transmission

Working Principle

The Layer-2 transparent transmission feature is designed to realize the transmission of Layer-2 packets between customer networks without impact on SP networks. When a Layer-2 packet from a customer network enters a PE, the PE changes the destination MAC address of the packet to a private address before forwarding the packet. The peer PE changes the destination MAC address to a public address to send the packet to the customer network at the other end, realizing transparent transmission on the SP network.

13.3.7 Priority Replication

Working Principle

If the SP network provides a QoS policy corresponding to a specified User Priority (CoS) in the inner tag, you can replicate the CoS of the inner tag to the outer tag to enable transparent transmission based on the QoS policy provided by the SP network.

13.3.8 Priority Mapping

Working Principle

If the SP network provides various QoS policies corresponding to specified CoS values for different service flows, you can map the CoS value of the inner tag to the CoS value of the outer tag to ensure preferential transmission of service flows based on the QoS policies provided by the SP network.

13.4 Configuration

Configuration	Description and Command	
Configuring QinQ	Mandatory.	
	switchport mode dot1q-tunnel	Configures a Tunnel port.
	switchport dot1q-tunnel allowed vlan { [**add**] **tagged** *vlist* | [**add**] **untagged** *vlist* | **remove** *vlist* }	Adds the VLANs to the Tunnel port in tagged or untagged mode.
	switchport dot1q-tunnel native vlan *VID*	Configures the default VLAN for the Tunnel port.
Configuring C-TAG-Based Selective QinQ	(Mandatory) It is used to configure C-TAG-based selective QinQ based on basic QinQ. Selective QinQ prevails over basic QinQ.	
	dot1q outer-vid *VID* **register inner-vid** *v_list*	Configures the policy to add the VLAN IDs of outer tags based on inner tags.
Configuring ACL-Based Selective QinQ	(Mandatory) It is used to configure ACL-based selective QinQ based on basic QinQ. Selective QinQ prevails over basic QinQ.	
	traffic-redirect access-group *acl* **nested-vlan** *VID* **in**	Configures the policy to add the VLAN IDs of outer tags based on ACLs.
Configuring VLAN Mapping	(Mandatory) It is used to enable VLAN mapping.	
	vlan-mapping-in vlan *cvlan* **remark** *svlan*	Configures 1:1 VLAN mapping in the inbound direction. This feature changes the inner VLAN ID of the packet entering a port to a specified outer VLAN ID.
	vlan-mapping-out vlan *svlan* **remark** *cvlan*	Configures 1:1 VLAN mapping in the outbound direction. This feature changes the outer VLAN ID of the packet exiting a port to a specified inner VLAN ID.
Configuring TPIDs	(Optional) It is used to realize TPID compatibility.	
	frame-tag tpid *tpid*	Configures the TPID of a frame tag. If you want to set it to 0x9100, configure the **frame-tag tpid 9100** command. By default, the TPID is in hexadecimal format. You need to configure this feature on an egress port.
Configuring MAC Address Replication	(Optional) It is used to configure MAC address replication to prevent floods.	
	mac-address-mapping *x* **source-vlan** *src-vlan-list* **destination-vlan** *dst-vlan-id*	Replicates the dynamic MAC address of the source VLAN to the destination VLAN.
Configuring an Inner/Outer VLAN Tag Modification Policy	(Optional) It is used to adjust the outer and inner VLAN tags of the packets transmitted over SP networks based on network topologies.	
	dot1q relay-vid *VID* **translate local-vid** *v_list*	Configures the policy to change the VLAN IDs of outer tags based on the outer tags.
	dot1q relay-vid *VID* **translate inner-vid** *v_list*	Configures the policy to change the VLAN IDs of outer tags based on inner tags.
	dot1q new-outer-vlan *VID* **translate old-outer-vlan vid inner-vlan** *v_list*	Configures the policy to change the VLAN IDs of outer tags based on outer and inner tags.

Configuration	Description and Command	
	traffic-redirect access-group *acl* **outer-vlan** *VID* **in**	Configures the policy to change the VLAN IDs of outer tags based on an ACL.
	traffic-redirect access-group *acl* **inner-vlan** *VID* **out**	Configures the policy to change the VLAN IDs of inner tags based on an ACL.
Configuring Priority Mapping and Priority Replication	(Optional) It is used to apply the QoS policy provided by the SP network by priority replication.	
	inner-priority-trust enable	Replicates the value of the User Priority field in the inner tag (C-TAG) to the User Priority field of the outer tag (S-TAG).
	(Optional) It is used to apply the QoS policy provided by the SP network by priority mapping.	
	dot1q-Tunnel cos *inner-cos-value* **remark-cos** *outer-cos-value*	Sets the value of the User Priority field in the outer tag (S-TAG) based on the User Priority field of the inner tag (C-TAG).
Configuring Layer-2 Transparent Transmission	(Optional) It is used to transmit MSTP and GVRP packets transparently based on the customer network topology without affecting the SP network topology.	
	l2protocol-tunnel stp	Enables STP transparent transmission in global configuration mode.
	l2protocol-tunnel stp enable	Enables STP transparent transmission in interface configuration mode.
	l2protocol-tunnel gvrp	Enables GVRP transparent transmission in global configuration mode.
	l2protocol-tunnel gvrp enable	Enables GVRP transparent transmission in interface configuration mode.
	l2protocol-tunnel{**STP**|**GVRP**}**tunnel-dmac** *mac-address*	Configures a transparent transmission address.

Pay attention to the following limitations when you configure QinQ:

Do not configure a routed port as the Tunnel port.

Do not enable 802.1X on the Tunnel port.

Do not enable the port security function on the Tunnel port.

When the Tunnel port is configured as the source port of the remote switched port analyzer (RSPAN), the packets whose outer tags contain VLAN IDs consistent with the RSPAN VLAN IDs are monitored.

If you want to match the ACL applied to the Tunnel port with the VLAN IDs of inner tags, use the inner keyword.

Configure the egress port of the customer network connected to the SP network as an Uplink port. If you configure the TPID of the outer tag on a QinQ-enabled port, set the TPID of the outer tag on the Uplink port to the same value.

By default, the maximum transmission unit (MTU) on a port is 1,500 bytes. After added with an outer VLAN tag, a packet is four bytes longer. It is recommended to increase the port MTU on the SP networks to at least 1,504 bytes.

After a switch port is enabled with QinQ, you must enable SVGL sharing before enabling IGMP snooping. Otherwise, IGMP snooping will not work on the QinQ-enabled port.

If a packet matches two or more ACL-based selective QinQ policies without priority, only one policy is executed. It is recommended to specify the priority.

13.4.1 Configuring QinQ

Configuration Effect

- Implement Layer-2 VPN based on a port-based QinQ policy.

Notes

- It is not recommended to configure the native VLAN of the Trunk port on the PE as its default VLAN, because the Trunk port strips off the tags containing the native VLAN IDs when sending packets.

Configuration Steps

Configuring the Tunnel port

- (Mandatory) Configure the Tunnel port in interface configuration mode.
- Run the **switchport mode dot1q-tunnel** command in interface configuration mode to configure the Tunnel port.

Command	**switchport mode dot1q-tunnel**
Parameter Description	N/A
Defaults	By default, no Tunnel port is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Native VLAN

- Mandatory.
- Configure the native VLAN for the Tunnel port.
- After you configure the native VLAN, add it to the VLAN list of the Tunnel port in untagged mode.
- Run the **switchport dot1q-tunnel native vlan** *VID* command in interface configuration mode to configure the default VLAN for the Tunnel port.
- If the native VLAN is added to the VLAN list in untagged mode, the outgoing packets on the Tunnel port are not tagged. If the native VLAN is added to the VLAN list in tagged mode, the outgoing packets on the Tunnel port are tagged with the native VLAN ID. To ensure the uplink and downlink transmission, add the native VLAN to the VLAN list in untagged mode.

Command	**switchport dot1q-tunnel native vlan** *VID*
Parameter Description	*VID*: Indicates the ID of the native VLAN. The value ranges from 1 to 4,094. The default value is 1.
Defaults	By default, the native VLAN is VLAN 1.
Command Mode	Interface configuration mode
Usage Guide	Use this command to configure the VLAN of the SP network.

Adding the VLANs on the Tunnel port

- Mandatory.
- After you configure the native VLAN, add it to the VLAN list of the Tunnel port in untagged mode.

- If port-based QinQ is enabled, you do not need to add the VLANs of the customer network to the VLAN list of the Tunnel port.
- If selective QinQ is enabled, add the VLANs of the customer network to the VLAN list of the Tunnel port in tagged or untagged mode based on requirements.
- Run the **switchport dot1q-tunnel allowed vlan** { [**add**] **tagged** *vlist* | [**add**] **untagged** *vlist* | **remove** *vlist* } command in interface configuration mode to add VLANs to the VLAN list of the Tunnel port. Upon receiving packets from corresponding VLANs, the Tunnel port adds or removes tags based on the settings.

Command	**switchport dot1q-tunnel allowed vlan** { [**add**] **tagged** *vlist* | [**add**] **untagged** *vlist* | **remove** *vlist* }
Parameter Description	*v_list*: Indicates the list of the VLANs on the Tunnel port.
Defaults	By default, VLAN 1 is added to the VLAN list of the Tunnel port in untagged mode. Other VLANs are not added.
Command Mode	Interface configuration mode
Usage Guide	Use this command to add or remove VLANs on the Tunnel port and specify whether the outgoing packets are tagged or untagged. If basic QinQ is enabled, add the native VLAN to the VLAN list of the Tunnel port in untagged mode.

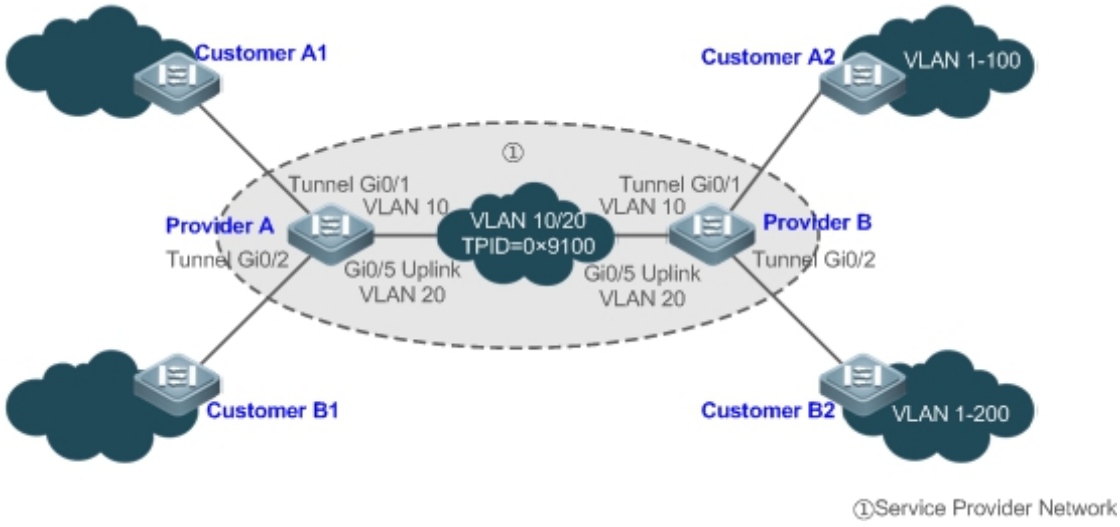
Verification

Check the Tunnel port configuration.

- Check whether the Tunnel port is configured properly on a switch.

Configuration Example

Configuring Basic QinQ to Implement Layer-2 VPN

Scenario **Figure 12-10**	
Configuration Steps	- Configure Tunnel ports on the PEs and connect the CEs to the Tunnel ports. - Configure the native VLANs for the Tunnel ports and add the native VLANs to the VLAN lists of the Tunnel ports respectively in untagged mode. - Configure VLANs on the customer networks based on requirements. QinQ-enabled switches encapsulate outer tags in packets for transmission over the SP network. Therefore, you do not need to configure customer VLANs on the PEs. The TPID is 0x8100 by default according to IEEE802.1Q. On some third-party switches, the TPID is set to a different value. If such switches are deployed, set the TPIDs on the ports connected to the third-party switches to realize

	TPID compatibility. If the PEs are connected through Trunk ports or Hybrid ports, do not configure the native VLANs for the Trunk ports or Hybrid ports as the default VLANs for the Tunnel ports. The Trunk ports or Hybrid ports strip off the VLAN tags containing the Native VLAN IDs when sending packets.
Provider A	Step 1: Create VLAN 10 and VLAN 20 on the SP network to segregate the data of Customer A and Customer B. ProviderA#configure terminal Enter configuration commands, one per line. End with CNTL/Z. ProviderA(config)#vlan 10 ProviderA(config-vlan)#exit ProviderA(config)#vlan 20 ProviderA(config-vlan)#exit Step 2: Enable basic QinQ on the port connected to the network of Customer A to use VLAN 10 for tunneling. ProviderA(config)#interface gigabitEthernet 0/1 ProviderA(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel ProviderA(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 10 ProviderA(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel allowed vlan add untagged 10 Step 3: Enable basic QinQ on the port connected to the network of Customer B to use VLAN 20 for tunneling. ProviderA(config)#interface gigabitEthernet 0/2 ProviderA(config-if-GigabitEthernet 0/2)#switchport mode dot1q-tunnel ProviderA(config-if-GigabitEthernet 0/2)#switchport dot1q-tunnel native vlan 20 ProviderA(config-if-GigabitEthernet 0/2)#switchport dot1q-tunnel allowed vlan add untagged 20 Step 4: Configure an Uplink port. ProviderA(config)# interface gigabitEthernet 0/5 ProviderA(config-if-GigabitEthernet 0/5)#switchport mode uplink Step 5: Change the TPID of the outgoing packets on the Uplink port to a value (for example, 0x9100) recognizable by third-party switches. ProviderA(config-if-GigabitEthernet 0/5)#frame-tag tpid 9100 Step 6: Configure Provider B by performing the same steps.
Verification	Customer A1 sends a packet containing VLAN ID 100 destined to Customer A2. The packet through Provider A is tagged with the outer tag specified by the Tunnel port. The packet that reaches Customer A2 carries the original VLAN ID 100. Check whether the Tunnel port is configured correctly. Check whether the TPID is configured correctly.
Provider A	ProviderA#show running-config interface GigabitEthernet 0/1 switchport mode dot1q-tunnel switchport dot1q-tunnel allowed vlan add untagged 10

	switchport dot1q-tunnel native vlan 10 spanning-tree bpdufilter enable ! interface GigabitEthernet 0/2 switchport mode dot1q-tunnel switchport dot1q-tunnel allowed vlan add untagged 20 switchport dot1q-tunnel native vlan 20 spanning-tree bpdufilter enable ! interface GigabitEthernet 0/5 switchport mode uplink frame-tag tpid 0x9100 ProviderA#show interfaces dot1q-tunnel ========Interface Gi0/1======== Native vlan: 10 Allowed vlan list:1,10, Tagged vlan list: ========Interface Gi0/2======== Native vlan: 20 Allowed vlan list:1,20, Tagged vlan list: ProviderA#show frame-tag tpid Ports Tpid ------- ---------- Gi0/5 0x9100
Provider B	Check Provider B by performing the same steps.

Common Errors

- The native VLAN is not added to the VLAN list of the Tunnel port in untagged mode.
- No TPID is configured on the port connected to the third-party switch on which TPID is not 0x8100. As a result, packets cannot be recognized by the third-party switch.

13.4.2 Configuring C-TAG-Based Selective QinQ

Configuration Effect

- Encapsulate outer VLAN tags (S-TAGs) in packets based on inner tags to ensure preferential transmission and management of Layer-2 VPN and service flows.

Notes

- C-TAG-based selective QinQ must be configured based on basic QinQ.
- Some selective QinQ policies are not supported on some products due to limitations of chips.
- If you need to continue to adopt the VLAN tag priority specified by the customer network, you can configure priority replication to configure an outer tag the same as the inner tag.
- If the SP network requires the transmission of packets based on the priority of the outer tag, you need to configure priority replication to set the CoS of the outer tag to the specified value.

Configuration Steps

Configuring a Policy to Add the VLAN IDs of Outer Tags Based on Inner Tags

- Mandatory.
- Upon receiving a packet, the Tunnel port adds the VLAN ID of the outer tag based on the VLAN ID of the inner tag. This function enables the Tunnel port to add the VLAN ID of the inner tag to the outer tag and adds the port to the VLAN in untagged mode. In this way, the outgoing packets carry the original inner tags.

The ACL-based QinQ policy prevails over the port-based and C-TAG-based QinQ policy.

When a member port is added to or removed from an aggregate port (AP), the QinQ policy configured on the AP port will be deleted. You need to configure the policy again. It is recommended that you configure a selective QinQ policy on the AP port after you configure its member ports.

You must configure the Tunnel port and the port connected to the public network to permit packets with specified VLAN IDs (including the native VLAN ID) in the outer tag to pass through.

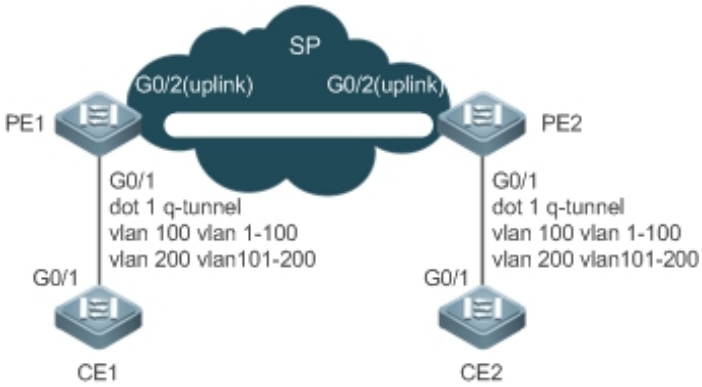
Command	**dot1q outer-vid** *VID* **register inner-vid** *v_list*
Parameter Description	N/A
Defaults	By default, no policy is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

- Check whether the users within the VLANs can communicate with each other.
- Check whether Layer-2 VPN is implemented.
- Check whether different service traffic is transmitted based on the selective QinQ policy, such as outer tag insertion, priority replication, and priority mapping.

Configuration Example

Implementing Layer-2 VPN and Service Flow Management Through C-TAG-Based Selective QinQ

Scenario **Figure 12-11**	
Configuration Steps	● Configure the ports on PE 1 and PE 2 connected to CE 1 and CE 2 as Tunnel ports. ● Configure a selective QinQ policy to add an outer tag to the packet based on its inner tag. ● If the SP network provides a VLAN-based QoS policy, the policy enables the port to add the outer tags with the corresponding VLAN ID to the specified service flow packets. ● If the SP network provides a CoS-based QoS policy and the CoS value is the same as that of the inner tag, you can configure priority mapping to replicate the CoS value of the inner tag to the outer VLAN tag so that the packet is transmitted based on the priority policy for the inner tag. ● If the SP network provides a CoS-based QoS policy, you can configure priority mapping to set the CoS value of the outer VLAN tag to a specified value so that the packet is transmitted based on the priority policy.
PE1	Step 1: Configure the VLAN for transparent transmission. PE1#configure terminal Enter configuration commands, one per line. End with CNTL/Z. PE1(config)#vlan 100 PE1(config-vlan)#exit PE1(config)#vlan 200 PE1(config-vlan)#exit Step 2: On the Downlink port of the access switch, configure a selective QinQ policy to add outer tags based on inner tags. Configure port Gi 0/1 as a Tunnel port. PE1(config)#interface gigabitEthernet 0/1 PE1(config-if)# switchport mode dot1q-tunnel Add VLAN 101 and VLAN 201 of the SP to the VLAN list of the Tunnel port and configure the Tunnel port to strip off the outer tag from incoming packets. PE1(config-if)# switchport dot1q-tunnel allowed vlan add untagged 100,200 Configure the Tunnel port to add outer tag VLAN 100 to incoming data frames containing inner tag VLAN 1–100. PE1(config-if)# dot1q outer-vid 100 register inner-vid 1-100 Configure the Tunnel port to add outer tag VLAN 200 to incoming data frames containing inner tag VLAN 101-200. PE1(config-if)# dot1q outer-vid 200 register inner-vid 101-200

	Step 3: Configure the port that accesses the SP network as an Uplink port. PE1(config)# interface gigabitEthernet 0/2 PE1(config-if-GigabitEthernet 0/2)#switchport mode uplink
PE2	● Perform the same configuration on PE 2.
Verification	Verify the configuration by checking whether: ● The Downlink port is configured as a Tunnel port. ● The VLAN specified by the outer tag is added to the VLAN list of the Tunnel port. ● The selective QinQ policy on the Tunnel port is correct. ● The Uplink port is configured correctly. Step 1: Check whether the VLAN mapping policy is correct.
PE1	PE1#show running-config interface gigabitEthernet 0/1 interface GigabitEthernet 0/1 switchport mode dot1q-tunnel switchport dot1q-tunnel allowed vlan add untagged 100,200 dot1q outer-vid 100 register inner-vid 1-200 dot1q outer-vid 200 register inner-vid 101-200 spanning-tree bpdufilter enable ! Step 2: Check the C-TAG-based selective QinQ policy. Check whether the mapping relationship between the inner and outer VLAN tags is correct. PE1#show registration-table Ports Type Outer-VID Inner-VID-list ------ ---------- ---------- ---------- *Gi0/1 Add-outer 100 1-200* *Gi0/1 Add-outer 200 101-200*

13.4.3 Configuring ACL-Based Selective QinQ

Configuration Effect

- Encapsulate outer VLAN tags (S-TAGs) in packets based on the ACL-based flow classification to allow the SP network to manage different services.

Notes

- ACL-based selective QinQ must be configured based on basic QinQ.
- Some selective QinQ policies are not supported on some products due to limitations of chips.
- If you need to continue to adopt the VLAN tag priority specified by the customer network, you can configure priority replication to configure an outer tag the same as the inner tag.

- If the SP network requires the transmission of packets based on the priority of the outer tag, you need to configure priority replication to set the CoS of the outer tag to the specified value.

The ACL-based QinQ policy prevails over the port-based and C-TAG-based QinQ policy.

When an ACL is deleted, the related policy will be automatically deleted.

Upon receiving a packet with two or more tags, the Tunnel port cannot add an outer tag to the packet based on the ACL-based selective QinQ policy.

If a packet matches two or more ACL-based selective QinQ policies without priority, only one policy is executed. It is recommended to specify the priority.

You must configure the Tunnel port and the port connected to the public network to permit packets with specified VLAN IDs (including the native VLAN ID) in the outer tag to pass through.

Configuration Steps

Configuring a Policy to Add the VLAN IDs of Outer Tags Based on ACLs

- Mandatory.
- The Tunnel port adds outer tags with different VLAN IDs to incoming packets based on the packet content.

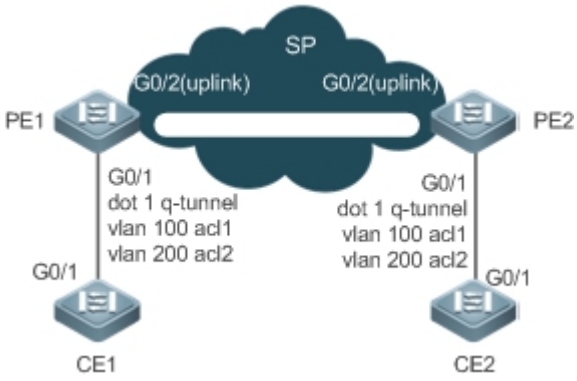
Command	**traffic-redirect access-group** *acl* **nested-vlan** *VID* **in**
Parameter Description	N/A
Defaults	By default, no policy is added.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

- Check whether the users of the same service in different branch offices can communicate with each other and whether specified service data is transmitted preferentially through virtual private LAN segment (VPLS) configuration.
- Check whether Layer-2 VPN is implemented.
- Check whether different service traffic is transmitted based on the selective QinQ policy, such as outer tag insertion, priority replication, and priority mapping.

Configuration Example

Implementing Layer-2 VPN and Service Flow Management Through ACL-Based Selective QinQ

Scenario **Figure 12-12**	
Configuration Steps	● Configure the ports on PE 1 and PE 2 connected to CE 1 and CE 2 as Tunnel ports. ● Configure ACL policies on PE 1 and PE 2 to segregate the service flows from the customer network. ● On the Tunnel ports, configure a selective QinQ policy to add an outer tag to the packet based on ACL policies. ● If the SP network provides a VLAN-based QoS policy, the policy enables the port to add the corresponding VLAN ID to the outer tags of the specified service flow. ● If the SP network provides a CoS-based QoS policy, you can configure priority mapping to set the CoS value of the outer VLAN tag to a specified value so that the packet is transmitted based on the priority policy.
PE 1	Step 1: Create an ACL to permit flows of PPPoE type (0x8863/0x8864) and IPoE type (0x0800) to pass through. PE1#configure terminal Enter configuration commands, one per line. End with CNTL/Z. PE1(config)# expert access-list extended acl1 PE1(config-exp-nacl)# permit 0x8863 any any PE1(config-exp-nacl)# permit 0x8864 any any PE1(config-exp-nacl)#exit PE1(config)# expert access-list extended acl2 PE1(config-exp-nacl)#permit 0x0800 any any Step 2: Configure VLAN 100 and VLAN 200 on the SP network to segregate data. PE#configure terminal Enter configuration commands, one per line. End with CNTL/Z. PE1(config)#vlan 100 PE1(config-vlan)#exit PE1(config)#vlan 200 PE1(config-vlan)#exit Step 3: On the Downlink port of the access switch, configure a selective QinQ policy to add outer VLAN tags based on ACLs. Configure port Gi 0/1 as a Tunnel port. PE1(config)#interface gigabitEthernet 0/1

	PE1(config-if)# switchport mode dot1q-tunnel Add VLAN 100 and VLAN 200 of the SP to the VLAN list of the Tunnel port and configure the Tunnel port to strip off the outer tag from incoming packets. PE1(config-if)#switchport dot1q-tunnel allowed vlan add untagged 100,200 Configure the Tunnel port to add outer tag VLAN 100 to the incoming data frames which match ACL 1. PE1(config-if)# traffic-redirect access-group acl1 nested-vlan 100 in Configure the Tunnel port to add outer tag VLAN 200 to the incoming data frames which match ACL 2. PE1(config-if)# traffic-redirect access-group acl1 nested-vlan 200 in Step 4: Configure the port connected to the SP network as an Uplink port. PE1(config)# interface gigabitEthernet 0/2 PE1(config-if-GigabitEthernet 0/2)#switchport mode uplink
Verification	Check whether the users of the same service in different branch offices can communicate with each other and whether specified service data is transmitted preferentially. ● Check whether Layer-2 VPN is implemented. ● Check whether the ACL is correct. ● Check whether the service priority is correct. ● Check whether the Downlink port is configured as a Tunnel port, whether the outer tag VLAN is added to the VLAN list of the Tunnel port, and whether the mapping policy on the Tunnel port is correct.
PE1	Step 1: Check whether the Tunnel port is configured correctly. FS#show running-config interface gigabitEthernet 0/1 interface GigabitEthernet 0/1 switchport mode dot1q-tunnel switchport dot1q-tunnel allowed vlan add untagged 100,200 traffic-redirect access-group acl1 nested-vlan 100 in traffic-redirect access-group acl2 nested-vlan 200 in spanning-tree bpdufilter enable ! Step 2: Check the ACL-based selective QinQ policy. Check whether the mapping relationship between the inner and outer VLAN tags is correct. PE1#show traffic-redirect Ports Type VID Match-filter ------- ---------- ----- ---------- Gi0/1 Nested-vid 101 acl1 Gi0/1 Nested-vid 201 acl2

Common Errors

- No ACL policy is configured.
- ACL policies are used to segregate flows based on MAC addresses. Packet floods will occur if MAC address replication is not configured.

13.4.4 Configuring VLAN Mapping

Configuration Effect

- Replace the inner tags of the packets with the outer tags to allow the packets to be transmitted based on the VLAN planning on the SP network.

Notes

- VLAN mapping can be configured only on Access ports, Trunk ports, Hybrid ports, or Uplink ports.

After VLAN mapping is configured, the VLAN IDs of the packets sent to the CPU are changed to the specified VLAN ID.

It is not recommended to configure VLAN mapping and selective QinQ on one port.

Configuration Steps

Configuring 1:1 VLAN Mapping

- Mandatory if the 1:1 mode is used. Configure a 1:1 VLAN mapping rule.
- Run the **vlan-mapping-in vlan** *CVID* **remark** *SVID* command or the **vlan-mapping-out vlan** *SVID* **remark CVID** command on a Trunk port or an Uplink port to enable 1:1 VLAN mapping.

Command	**vlan-mapping-in vlan** *src-vlan-list* **remark** *dest-vlan*
Parameter Description	*src-vlan-list*: Indicates the customer VLAN. *dest-vlan:* Indicates the service VLAN, which is the VLAN where the SP network is located.
Defaults	
Command Mode	Interface configuration mode
Usage Guide	Use this command to configure 1:1 VLAN mapping in the inbound direction.

Command	**vlan-mapping-out vlan** *src-vlan* **remark** *dest-vlan*
Parameter Description	*src-vlan*: Indicates the service VLAN, which is the VLAN where the SP network is located. *dest-vlan*: Indicates the customer VLAN.
Defaults	
Command Mode	Interface configuration mode
Usage Guide	Use this command to configure 1:1 VLAN mapping in the outbound direction.

Verification

Check whether VLAN mapping is configured correctly.

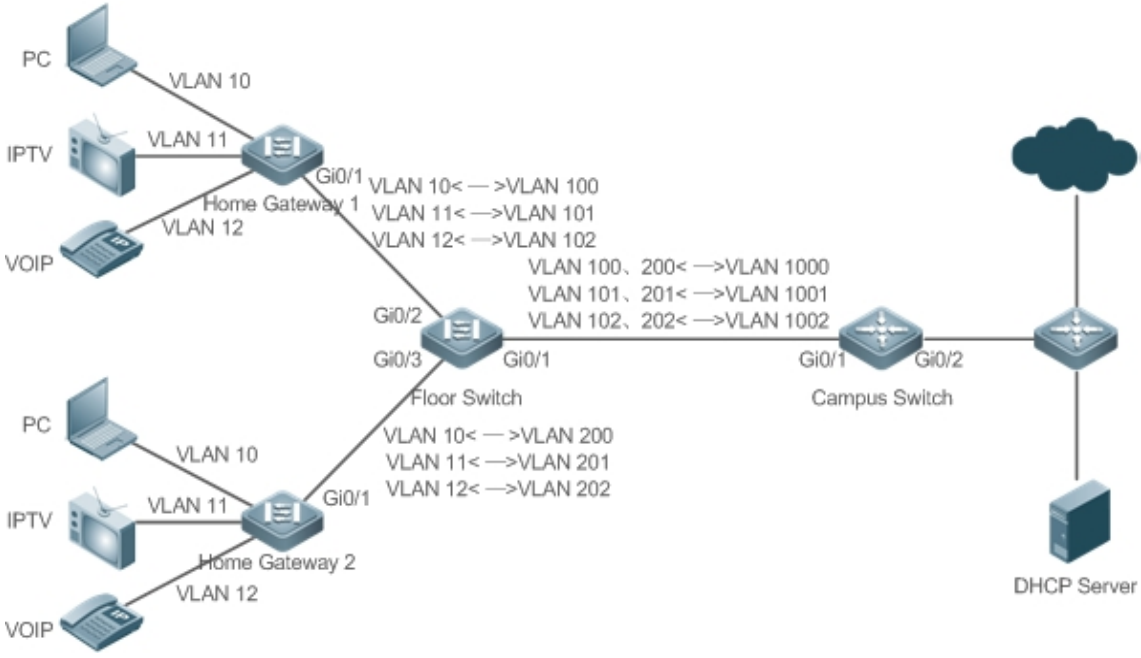
- Run the **show interfaces**[*intf-id*] **vlan-mapping** command to display the VLAN mapping.

Configuration Example

Implementing VLAN Aggregation for Different Services Through VLAN Mapping

Scenario

Figure 12-13



Configuration Steps

- Configure Home Gateway 1 and Home Gateway 2.

Step 1: On the home gateways, configure the original VLANs for different services.

```
FS#configure terminal
Enter configuration commands, one per line.    End with CNTL/Z.
FS(config)#vlan range 10-12
FS(config-vlan-range)#exit
```

Step 2: Configure the attributes of the ports connected to PC, IPTV, and VoIP. Assume that the connected ports are Gi 0/2, Gi 0/3, and Gi 0/4 respectively.

```
FS(config)#interface gigabitEthernet 0/2
FS(config-if-GigabitEthernet 0/2)#switchport access vlan 10
FS(config-if-GigabitEthernet 0/2)#exit
FS(config)#interface gigabitEthernet 0/3
FS(config-if-GigabitEthernet 0/3)#switchport access vlan 11
FS(config-if-GigabitEthernet 0/3)#exit
FS(config)#interface gigabitEthernet 0/4
FS(config-if-GigabitEthernet 0/4)#switchport access vlan 12
FS(config-if-GigabitEthernet 0/4)#exit
```

Step 3: Configure an Uplink port.

```
FS(config)# interface gigabitEthernet 0/1
FS(config-if-GigabitEthernet 0/1)#switchport mode uplink
```

- Configure a floor switch with 1:1 VLAN mapping policies.

Step 1: On the home gateways, configure the original VLANs and mapped VLANs for different services.

```
FS#configure terminal
Enter configuration commands, one per line.    End with CNTL/Z.
FS(config)#vlan range 10-12
FS(config-vlan-range)#exit
FS(config)#vlan range 100-102
FS(config-vlan-range)#exit
FS(config)#vlan range 200-202
FS(config-vlan-range)#exit
```

Step 2: On the Downlink port of Home Gateway 1, configure 1:1 VLAN mapping policies in the inbound and outbound directions.

```
FS(config)#interface gigabitEthernet 0/2
FS(config-if-GigabitEthernet 0/2)#switchport mode uplink
FS(config-if-GigabitEthernet 0/2)#vlan-mapping-in vlan 10 remark 100
FS(config-if-GigabitEthernet 0/2)#vlan-mapping-in vlan 11 remark 101
FS(config-if-GigabitEthernet 0/2)#vlan-mapping-in vlan 12 remark 102
FS(config-if-GigabitEthernet 0/2)#vlan-mapping-out vlan 100 remark 10
FS(config-if-GigabitEthernet 0/2)#vlan-mapping-out vlan 101 remark 11
FS(config-if-GigabitEthernet 0/2)#vlan-mapping-out vlan 102 remark 12
```

Step 3: On the Downlink port of Home Gateway 2, configure 1:1 VLAN mapping policies in the inbound and outbound directions.

```
FS(config)#interface gigabitEthernet 0/3
FS(config-if-GigabitEthernet 0/3)#switchport mode uplink
FS(config-if-GigabitEthernet 0/3)#vlan-mapping-in vlan 10 remark 200
FS(config-if-GigabitEthernet 0/3)#vlan-mapping-in vlan 11 remark 201
FS(config-if-GigabitEthernet 0/3)#vlan-mapping-in vlan 12 remark 202
FS(config-if-GigabitEthernet 0/3)#vlan-mapping-out vlan 200 remark 10
FS(config-if-GigabitEthernet 0/3)#vlan-mapping-out vlan 201 remark 11
FS(config-if-GigabitEthernet 0/3)#vlan-mapping-out vlan 202 remark 12
```

Step 4: Configure an Uplink port.

```
FS(config)# interface gigabitEthernet 0/1
FS(config-if-GigabitEthernet 0/1)#switchport mode uplink
```

Verification

Display the 1:1 VLAN mapping policies configured on the floor switch.

```
FS#show interfaces vlan-mapping
Ports          type    Status      Service-Vlan    Customer-Vlan-list
```

```
-----------   -----   ---------   -----------      ----------------
Gi0/2         in      active      100              10
Gi0/2         in      active      101              11
Gi0/2         in      active      102              12
Gi0/2         out     active      100              10
Gi0/2         out     active      101              11
Gi0/2         out     active      102              12
Gi0/3         in      active      200              10
Gi0/3         in      active      201              11
Gi0/3         in      active      202              12
Gi0/3         out     active      200              10
Gi0/3         out     active      201              11
Gi0/3         out     active      202              12
```

13.4.5 Configuring TPIDs

Configuration Effect

Configure the TPIDs in the tags on SP network devices to realize TPID compatibility.

Notes

If a PE connected to a third-party switch on which the TPID is not 0x8100, you need to configure the TPID on the port of the PE connected to the third-party switch.

Do not set the TPIDs to any of the following values: 0x0806 (ARP), 0x0200 (PUP), 0x8035 (RARP), 0x0800 (IP), 0x86DD (IPv6), 0x8863/0x8864 (PPPoE), 0x8137 (IPX/SPX), 0x8000 (IS-IS), 0x8809 (LACP), 0x888E (802.1X), 0x88A7 (clusters), and 0x0789 (reserved by FS Networks).

Configuration Steps

- If a PE connected to a third-party switch on which the TPID is not 0x8100, you need to configure the TPID on the port of the PE connected to the third-party switch.
- TPIDs can be configured in interface configuration mode and global configuration mode. The following example adopts interface configuration mode.

Configure the **frame-tag tpid** *0x9100* command in interface configuration mode to change the TPID to 0x9100. For details about the TPID value, see section 1.4.5.

Command	**frame-tag tpid** *tpid*
Parameter Description	*tpid:* Indicates the new value of the TPID.
Defaults	The default value of the TPID is 0x8100.
Command Mode	Interface configuration mode
Usage Guide	If a PE is connected to a third-party switch on which the TPID is not 0x8100, use this command to configure the TPID on

	the port connected to the third-party switch.

Verification

Check whether the TPID is configured.

Configuration Example

Configuring the TPID on a port

Configuration Steps	Configure the TPID on a port. FS(config)# interface gigabitethernet *0/1* FS(config-if)# frame-tag tpid *9100*
Verification	Display the TPID on the port. FS# show frame-tag tpid interfaces gigabitethernet *0/1* Port tpid ------- ---------- Gi0/1 0x9100

13.4.6 Configuring MAC Address Replication

Configuration Effect

- Replicate the dynamic address learned on a port from one VLAN to another.
- Avoid packet floods when service flows are segregated through MAC-based ACLs.

Notes

After MAC address replication is disabled, the system will delete all the learned MAC address entries from the destination VLAN.

MAC address replication can be configured on a port only once. If you need to modify the configuration, delete the current configuration and configure it again.

VLAN MAC address replication cannot be used together with VLAN sharing, and the MAC addresses cannot be replicated to dynamic VLANs.

Up to eight destination VLANs can be configured on each port. MAC address replication takes effect even if the port does not belong to the specified destination VLAN.

MAC address replication cannot be configured on the Host and Promiscuous ports, monitoring ports, and port security-/802.1X-enabled ports.

Only dynamic addresses can be replicated. Address replication is disabled when the address table is full. If source addresses already exist before replication is enabled, corresponding MAC addresses will not be replicated.

Replicated addresses have a higher priority than dynamic addresses but have a lower priority than other types of addresses.

When a MAC address ages, the replicated MAC address will also age. When the MAC address is deleted, the replicated address will be deleted automatically.

Hot backup is not supported. After primary/secondary switchover occurs, it is recommended that you disable MAC address replication and then enable it again.

The MAC address entries obtained through MAC address replication cannot be deleted manually. If you need to delete these entries, disable MAC address replication.

Configuration Steps

Configuring MAC Address Replication

- Perform this configuration to replicate MAC addresses from one VLAN to another to avoid packet floods.
- Run the **mac-address-mapping** <1-8> **source-vlan** *src-vlan-list* **destination-vlan** *dst-vlan-id* command on a Trunk port to enable MAC address replication. *src-vlan-list* and *dst-vlan-id* specify the VLAN range.

Command	**mac-address-mapping** *x* **source-vlan** *src-vlan-list* **destination-vlan** *dst-vlan-id*
Parameter Description	*x:* Indicates the index number for MAC address replication. The value ranges from 1 to 8. *src-vlan-list*: Indicates the source VLAN list. *dst-vlan-id:* Indicates the destination VLAN list.
Defaults	By default, MAC address replication is disabled.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

- Check whether the MAC address of the specified VLAN is replicated to another VLAN.

Configuration Example

Configuring MAC Address Replication

Configuration Steps	● Configure MAC address replication. FS(config)# interface gigabitethernet 0/1 FS(config-if)# switchport mode trunk FS(config-if)#mac-address-mapping 1 source-vlan 1-3 destination-vlan 5
Verification	● Check whether the configuration takes effect on the port. ● Send a packet from the source VLAN and check whether the source MAC address of the packet is replicated to the destination VLAN. FS# show interfaces mac-address-mapping Ports destination-VID Source-VID-list ---------- Gi0/1 5 1-3

Common Errors

- See "Notes".

13.4.7 Configuring an Inner/Outer VLAN Tag Modification Policy

Configuration Effect

- Modify outer or inner tags based on the actual networking requirements.

Notes

- The ACL-based QinQ policy prevails over the port-based and C-TAG-based QinQ policy.
- When an ACL is deleted, the related policy will be automatically deleted.
- Tag modification policies take effect only on Access ports, Trunk ports, Hybrid ports, and Uplink ports.
- Tag modification policies are mainly used to modify inner and outer tags on the SP network.
- If a packet matches two or more ACL-based selective QinQ policies without priority, only one policy is executed. It is recommended to specify the priority.

Configuration Steps

Configuring the Policy to Change the VLAN IDs of Outer Tags Based on Inner Tags

- Optional.
- Perform this configuration to change the VLAN IDs of outer tags based on the VLAN IDs of inner tags.
- You can change the VLAN IDs of the outer tags in the packets that enter Access ports, Trunk ports, Hybrid ports, and Uplink ports based on the VLAN IDs of the inner tags in these packets.

Command	**dot1q relay-vid** *VID* **translate inner-vid** *v_list*
Parameter Description	*VID*: Indicates the modified VLAN ID of the outer tag. *v_list:* Indicates the VLAN ID of the inner tag.
Defaults	By default, no policy is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Policy to Change the VLAN IDs of Outer Tags Based on the VLAN IDs of Outer and Inner Tags

- Optional.
- Perform this configuration to change the VLAN IDs of outer tags based on the VLAN IDs of inner and outer tags.
- You can change the VLAN IDs of the outer tags in the packets that enter Access ports, Trunk ports, Hybrid ports, and Uplink ports based on the VLAN IDs of the inner and outer tags in these packets.

Command	**dot1q new-outer-vlan** *new-vid* **translate old-outer-vlan** *vid* **inner-vlan** *v_list*
Parameter Description	*new-vid*: Indicates the modified VLAN ID of the outer tag. vid: Indicates the original VLAN ID of the outer tag. *v_list:* Indicates the VLAN ID of the inner tag.
Defaults	By default, no policy is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Policy to Change the VLAN IDs of Outer Tags Based on the Outer Tags

- Optional.
- Perform this configuration to change the VLAN IDs of outer tags based on these VLAN IDs.
- You can change the VLAN IDs of the outer tags in the packets that enter Access ports, Trunk ports, Hybrid ports, and Uplink ports based on these VLAN IDs.

Command	**dot1q relay-vid** *VID* **translate local-vid** *v_list*
Parameter Description	*VID*: Indicates the modified VLAN ID of the outer tag. *v_list:* Indicates the original VLAN ID of the outer tag.
Defaults	By default, no policy is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring a Policy to Change the VLAN IDs of Inner Tags Based on ACLs

- Optional.
- You can change the VLAN IDs of the inner tags in the packets that exit Access ports, Trunk ports, Hybrid ports, and Uplink ports based on the packet content.
- Before you configure such a policy, configure an ACL.

Command	**traffic-redirect access-group** *acl* **inner-vlan** *vid* **out**
Parameter Description	*acl*: Indicates the ACL. *vid:* Indicates the modified VLAN ID of the inner tag.
Defaults	By default, no policy is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring a Policy to Change the VLAN IDs of Outer Tags Based on ACLs

- Optional.
- You can change the VLAN IDs of the outer tags in the packets that exit Access ports, Trunk ports, Hybrid ports, and Uplink ports based on the packet content.
- Before you configure such a policy, configure an ACL.

Command	**traffic-redirect access-group** *acl* **outer-vlan** *vid* **in**
Parameter Description	*acl*: Indicates the ACL. *vid:* Indicates the modified VLAN ID of the outer tag.
Defaults	By default, no policy is configured.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

Check whether the configuration takes effect and whether the port modifies the tags in received packets based on the policy.

Configuration Example

Configuring the Policy to Change the VLAN IDs of Outer Tags Based on the Outer Tags

Configuration Steps	● Configure inner/outer tag modification policies on a port based on the actual networking requirements. ● The following example shows how to change VLAN IDs of outer tags based on outer tags and ACLs respectively. For details about other policies, see the description above. Configure a policy to change outer VLAN tags based on the outer VLAN tags. FS(config)# interface gigabitEthernet 0/1 FS(config-if)# switchport mode trunk FS(config-if)# dot1q relay-vid 100 translate local-vid 10-20 Configure a policy to change outer VLAN tags based on ACLs. FS# configure terminal FS(config)# ip access-list standard 2 FS(config-acl-std)# permit host 1.1.1.1 FS(config-acl-std)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if)# switchport mode trunk FS(config-if)# traffic-redirect access-group 2 outer-vlan 3 in
Verification	● Check whether the configuration takes effect on the port. ● Check whether the port changes the VLAN IDs of the outer tags in received packets based on the configured policy.

13.4.8 Configuring Priority Mapping and Priority Replication

Configuration Effect

- If an SP network provides a QoS policy based on the User Priority field of the inner tag, configure priority replication to apply the QoS policy to the outer tag.
- If an SP network provides a QoS policy based on the User Priority field of the inner tag, configure priority mapping to apply the User Priority field provided by the SP network to the outer tag.

Notes

Only a Tunnel port can be configured with priority replication, which has a higher priority than trusted QoS but lower than ACL-based QoS.

Priority replication and priority mapping cannot be both enabled on one port.

Only a Tunnel port can be configured with priority mapping, which prevails over QoS.

The configuration of priority mapping does not take effect if no trust mode is configured (trust none) or the trust mode is not matched with priority mapping.

Configuration Steps

- Only a Tunnel port can be configured with priority mapping or priority replication.
- Configure priority replication to apply the inner tag-based QoS policy provided by the SP network.
- Configure priority mapping to configure the User Priority field of the outer VLAN tag based on the inner tag and apply the QoS policy flexibly.
- To enable priority replication, run the **inner-priority-trust enable** command on the Tunnel port.
- To enable priority mapping, run the **dot1q-Tunnel cos** *inner-cos-value* **remark-cos** *outer-cos-value* command on the Tunnel port.

inner-cos-value and *outer-cos-value* range from 0 to 7.

The following priority mapping is used when no priority mapping is configured:

```
inner pri    0    1    2    3    4    5    6    7
-----------------------------------------------------------------------
outer pri    0    1    2    3    4    5    6    7
```

Command	**inner-priority-trust enable**
Parameter Description	N/A
Defaults	By default, priority replication is disabled.
Command Mode	Interface configuration mode
Usage Guide	N/A

Command	**dot1q-Tunnel cos** *inner-cos-value* **remark-cos** *outer-cos-value*
Parameter Description	*inner-cos-value*: Indicates the CoS value of the inner tag. *outer-cos-value*: Indicates the CoS value of the outer tag.
Defaults	By default, priority mapping is disabled.
Command Mode	Interface configuration mode
Usage Guide	N/A

Verification

- Run the **show inner-priority-trust interfaces** *type intf-id* command and the **show interfaces** *type intf-id* **remark** command to check whether priority mapping or priority replication takes effect.

Configuration Example

Configuring Priority Mapping and Priority Replication

Configuration Steps	● To maintain the packet priority, you need to replicate the priority of the inner tag in a packet to the outer tag on the Tunnel port. ● To flexibly control the packet priority on the Tunnel port, you can add outer tags of different priorities to packets based on the priorities of the inner tags in the packets. Configure priority replication. FS(config)# interface gigabitethernet *0/1* FS(config-if)#mls qos trust cos FS(config-if)# inner-priority-trust enable FS(config)# end Configure priority mapping. FS(config)# interface gigabitethernet 0/2 FS(config-if)#dot1q-Tunnel cos 3 remark-cos 5
Verification	● Display the priority configuration on the port. Check whether priority replication is enabled on the Tunnel port. FS# show inner-priority-trust interfaces gigabitethernet *0/1* Port inner-priority-trust ------ ---------- Gi0/1 enable Display the priority mapping configured on the Tunnel port. FS# show interfaces gigabitethernet 0/1 remark Ports Type From value To value ---------- ---------- ---------- -------- Gi0/1 Cos-To-Cos 3 5

Common Errors

See "Notes".

13.4.9 Configuring Layer-2 Transparent Transmission

Configuration Effect

Transmit Layer-2 packets transparently without impact on the SP network and the customer network.

Notes

If STP is not enabled, you need to run the **bridge-frame forwarding protocol bpdu** command to enable STP transparent transmission.

Transparent transmission enabled on a port takes effect only after enabled globally. When transparent transmission takes effect on the port, the port does not participate in related protocol calculation. If the port receives a packet whose destination MAC address

is the special broadcast address, it determines that a networking error occurs and discards the packet.

Configuration Steps

Configuring STP Transparent Transmission

- Mandatory if you need to transparently transmit BPDU packets through STP.
- Enable STP transparent transmission in global configuration mode and interface configuration mode.
- Run the **l2protocol-tunnel stp** command in global configuration mode to enable STP transparent transmission.
- Run the **l2protocol-tunnel stp enable** command in interface configuration mode to enable STP transparent transmission.

Command	**l2protocol-tunnel stp**
Parameter Description	N/A
Defaults	By default, STP transparent transmission is disabled.
Command Mode	Global configuration mode
Usage Guide	N/A

Command	**l2protocol-tunnel stp enable**
Parameter Description	N/A
Defaults	By default, STP transparent transmission is disabled.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring GVRP Transparent Transmission

- Mandatory if you need to transparently transmit GVRP packets.
- Enable GVRP transparent transmission in global configuration mode and interface configuration mode.
- Run the **l2protocol-tunnel gvrp** command in global configuration mode to enable GVRP transparent transmission.
- Run the **l2protocol-tunnel gvrp enable** command in interface configuration mode to enable GVRP transparent transmission.

Command	**l2protocol-tunnel gvrp**
Parameter Description	N/A
Defaults	By default, GVRP transparent transmission is disabled.
Command Mode	Global configuration mode
Usage Guide	N/A

Command	**l2protocol-tunnel gvrp enable**
Parameter Description	N/A
Defaults	By default, GVRP transparent transmission is disabled.

Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring a Transparent Transmission Address

- Optional.
- Configure a transparent transmission address.

Command	**l2protocol-tunnel** { **stp** | **gvrp** } **tunnel-dmac** *mac-address*
Parameter Description	*mac-address*: Indicates the address used to transparently transmit packets.
Defaults	By default, the first three bytes of the transparent transmission address is 01d0f8, and the last three bytes are 000005 and 000006 for STP and GVTP respectively.
Command Mode	Interface configuration mode
Usage Guide	The following addresses are available for STP: 01d0.f800.0005, 011a.a900.0005, 010f.e200.0003, 0100.0ccd.cdd0, 0100.0ccd.cdd1, and 0100.0ccd.cdd2. The following addresses are available for GVRP: 01d0.f800.0006 and 011a.a900.0006. When no transparent transmission address is configured, the default settings are used.

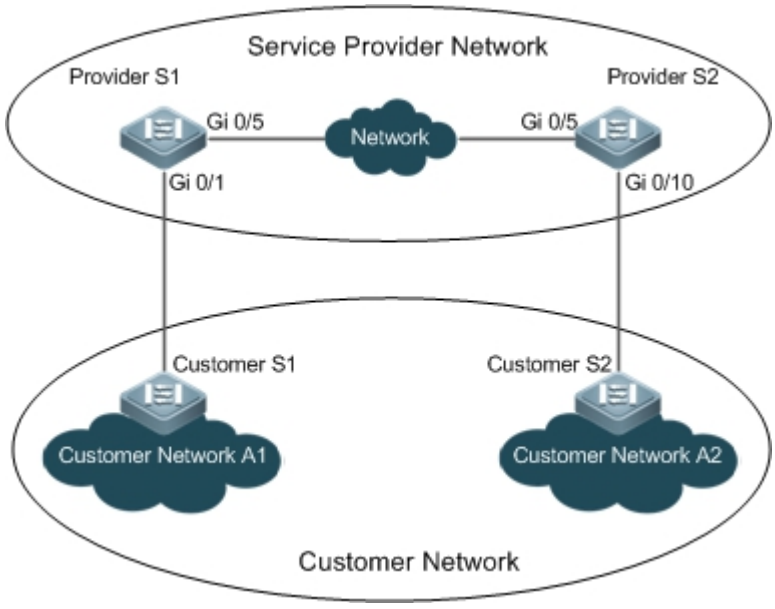
Verification

Run the **show l2protocol-tunnel stp** command and the **show l2protocol-tunnel gvrp** command to check whether the transparent transmission address is configured correctly.

Configuration Example

The following example shows how to configure STP transparent transmission.

Configuring STP Transparent Transmission

Scenario **Figure 12-14**	
Configuration	On the PEs (Provider S1 and Provider S2), enable STP transparent transmission in global configuration mode and

Steps	interface configuration mode. ● Before you enable STP transparent transmission, enable STP in global configuration mode to allow the switches to forward STP packets.
Provider S1	Step 1: Enable STP. bridge-frame forwarding protocol bpdu Step 2: Configure the VLAN for transparent transmission. ProviderS1#configure terminal Enter configuration commands, one per line. End with CNTL/Z. ProviderS1(config)#vlan 200 ProviderS1(config-vlan)#exit Step 3: Enable basic QinQ on the port connected to the customer network and use VLAN 200 for tunneling. ProviderS1(config)#interface gigabitEthernet 0/1 ProviderS1(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel ProviderS1(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 200 Step 4: Enable STP transparent transmission on the port connected to the customer network. ProviderS1(config-if-GigabitEthernet 0/1)#l2protocol-tunnel stp enable ProviderS1(config-if-GigabitEthernet 0/1)#exit Step 5: Enable STP transparent transmission in global configuration mode. ProviderS1(config)#l2protocol-tunnel stp Step 4: Configure an Uplink port. ProviderS1(config)# interface gigabitEthernet 0/5 ProviderS1(config-if-GigabitEthernet 0/5)#switchport mode uplink
Provider S2	Configure Provider S2 by performing the same steps.
Verification	Step 1: Check whether STP transparent transmission is enabled in global configuration mode and interface configuration mode. ProviderS1#show l2protocol-tunnel stp L2protocol-tunnel: Stp Enable GigabitEthernet 0/1 l2protocol-tunnel stp enable Step 2: Verify the configuration by checking whether: ● The port type is dot1q-tunnel. ● The outer tag VLAN is consistent with the native VLAN and added to the VLAN list of the Tunnel port. ● The port that accesses the SP network is configured as an Uplink port. ProviderS1#show running-config interface GigabitEthernet 0/1 switchport mode dot1q-tunnel

	switchport dot1q-tunnel allowed vlan add untagged 200 switchport dot1q-tunnel native vlan 200 l2protocol-tunnel stp enable spanning-tree bpdufilter enable ! interface GigabitEthernet 0/5 switchport mode uplink

Common Errors

- STP is not enabled in global configuration mode.
- Transparent transmission is not enabled in global configuration mode and interface configuration mode.

14 Configuring ERPS

14.1 Overview

Ethernet Ring Protection Switching (ERPS), also known as G.8032, is a ring protection protocol developed by the International Telecommunication Union (ITU). It is a data link layer protocol designed for Ethernet rings. ERPS prevents broadcast storms caused by data loops in an idle Ethernet ring and can rapidly recover the communication between nodes in the event that a link is disconnected in the Ethernet ring.

The Spanning Tree Protocol (STP) is another technique used to solve the Layer-2 loop problem. STP is at the mature application stage but requires a relatively long (seconds) convergence time compared to ERPS. ERPS reaches a Layer-2 convergence speed of less than 50 ms, faster than that of STP.

Scenario

- ITU-T G.8032/Y.1344: Ethernet ring protection switching

14.2 Applications

Application	Description
Single-Ring Protection	Only one ring exists in a network topology.
Tangent-Ring Protection	Two rings in a network topology share one device.
Intersecting-Ring Protection	Two or more rings in a network topology share one link.

14.2.1 Single-Ring Protection

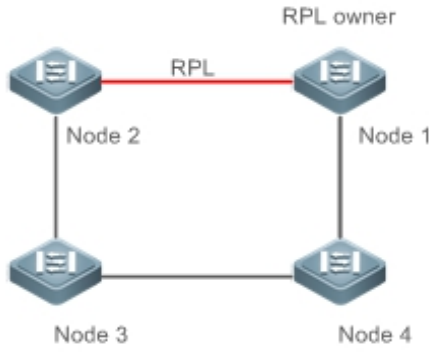
Scenario

Only one ring in a network topology needs to be protected.

In Figure 13-1, the network topology has only one ring, only one ring protection link (RPL) owner node, and only one RPL. All nodes must belong to the same ring automatic protection switching (R-APS) virtual local area network (VLAN).

- All devices in the ring network must support ERPS.
- Each link between devices must be a direct link without any intermediate device.

Figure 13-1



Remarks	The four devices in the ring network are aggregation switches.

Deployment

- All nodes in the physical topology are connected in ring mode.
- ERPS blocks the RPL to prevent loops. In Figure 13-1, the link between Node 1 and Node 2 is an RPL.
- ERPS is used to detect failures on each link between adjacent nodes.

14.3 Monitoring

Displaying

Description	Command
Displays whether the specified port is a Tunnel port.	**show dot1q-tunnel** [**interfaces** *intf-id*]
Displays the configuration of the Tunnel port.	**show interfaces dot1q-tunnel**
Displays the C-TAG-based selective QinQ policies on the Tunnel port.	**show registration-table** [**interfaces** *intf-id*]
Displays the C-TAG-based selective QinQ policies on the Access port, Trunk port or Hybrid port.	**show translation-table** [**interfaces** *intf-id*]
Displays VLAN mapping on ports.	**show interfaces** [*intf-id*] **vlan-mapping**
Displays the ACL-based selective QinQ policies.	**show traffic-redirect** [**interfaces** *intf-id*]
Displays the TPID configuration on ports.	**show frame-tag tpid interfaces** [*intf-id*]
Displays the configuration of priority replication.	**show inner-priority-trust**
Displays the configuration of priority mapping.	**show interface intf-name remark**
Displays the configuration of MAC address replication.	**show mac-address-mapping**
Displays the configuration of Layer-2 transparent transmission.	**show l2protocol-tunnel** { **gvrp** | **stp** }

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs QinQ.	**debug bridge qinq**

12.5.1 Tangent-Ring Protection

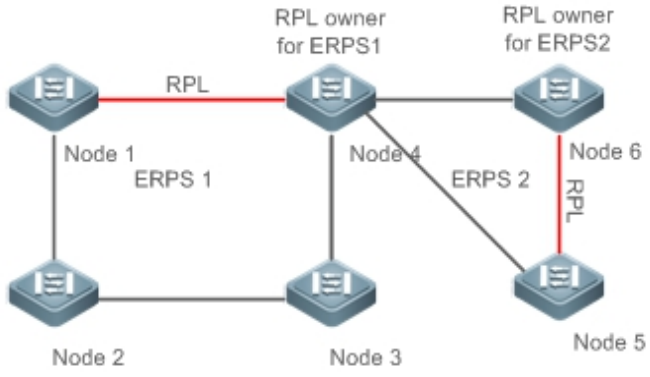
Scenario

The two rings in a network topology that share one device need to be protected.

In Figure 13-2, the two rings in the network topology share one device. Each ring has only one PRL owner node and only one RPL. The two rings belong to different R-APS VLANs.

- All devices in the ring network must support ERPS.
- Each link between devices must be a direct link without any intermediate device.

Figure 13-2



Remarks	The devices in the ring network are aggregation switches.

Deployment

- All nodes in the physical topology are connected in ring mode.
- ERPS blocks the RPL of each ring to prevent loops.
- ERPS is used to detect failures on each link between adjacent nodes.

12.5.2 Intersecting-Ring Protection

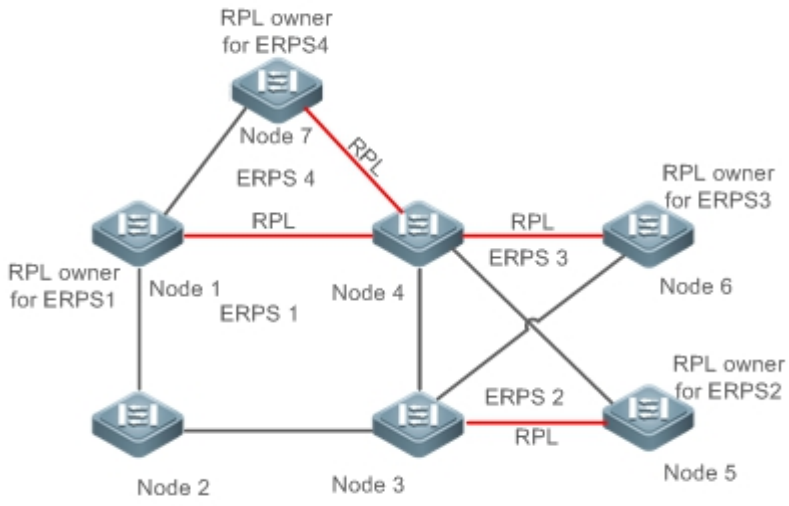
Scenario

Two or more rings in a network topology share one link. (Each link between intersecting nodes must be a direct link without any intermediate node.)

In Figure 13-3, four rings exist in the network topology. Each ring has only one PRL owner node and only one RPL. The four rings belong to different R-APS VLANs.

- All devices in the ring network must support ERPS.
- Each link between devices must be a direct link without any intermediate device.

Figure 13-3



Remarks	The devices in the ring network are aggregation switches.

Deployment

- All nodes in the physical topology are connected in ring mode.
- ERPS blocks the RPL of each ring to prevent loops.
- ERPS is used to detect failures on each link between adjacent nodes.

14.4 Features

Basic Concepts

Ethernet Ring

Ethernet rings are classified into common Ethernet rings and Ethernet subrings.

- **Common Ethernet ring:** Is an Ethernet topology with ring connection.
- **Ethernet subring:** An open topology that is mounted on other rings or networks through intersecting nodes and forms a closed topology with the channel between the intersecting nodes belonging to other rings or networks.

An Ethernet ring (a common Ethernet ring or an Ethernet subring) can be in one of the following states:

- **Idle state:** The physical links in the entire ring network are reachable.
- **Protection state:** A physical link in the ring network is disconnected.

Link and Channel

- **RPL:** An Ethernet ring (a common Ethernet ring or an Ethernet subring) has only one RPL. When an Ethernet ring is idle, the RPL is blocked and does not forward data packets to prevent loops. In Figure 13-2, the link between Node 1 and Node 4 is the RPL of ERPS 1, and Node 4 blocks the RPL port (the port mapped to the RPL). The link between Node 4 and Node 5 is the RPL of ERPS 2, and Node 5 blocks the RPL port.

- **Subring link:** Belongs to a subring in intersecting rings and is controlled by the subring. In Figure 13-3, ERPS 1 is a common Ethernet ring, and ERPS 2 is an Ethernet subring. The link between Node 4 and Node 5 and the link between Node 3 and Node 5 belong to ERPS 2. The other links belong to ERPS 1.

The link between Node 3 and Node 4 belongs to ERPS 1 rather than ERPS 2, and the link is not controlled by ERPS 2.

- **R-APS virtual channel:** Transmits ERPS packets of subrings between intersecting nodes in intersecting rings, but it does not belong to the subring. In Figure14-3, Node 1 blocks the RPL, and the packets of subring ERPS 2 are transmitted through the direct link between Node 3 and Node 4 in Ethernet ring ERPS 1. The direct link between Node 3 and Node 4 is the R-APS virtual channel of ERPS 2.

Node

Each device in an Ethernet ring is a node.

ERPS has the following node roles for a specific Ethernet ring:

- **RPL owner node:** A node that is adjacent to an RPL and is used to block the RPL to prevent loops when the Ethernet ring is free of faults. An Ethernet ring (a common Ethernet ring or an Ethernet subring) has only one RPL owner node. In Figure 13-2, Node 1 functions as the RPL owner node of Ethernet ring ERPS 1, and Node 6 functions as the RPL owner node of Ethernet subring ERPS 2.
- **Non-RPL owner node:** Any other node than the RPL owner node in an Ethernet ring. In Figure 13-2, nodes except Node 1 and Node 6 are non-RPL owner nodes of their respective rings.

ERPS has the following roles globally (not for a specific Ethernet ring):

- **Intersecting node:** A node that belongs to multiple intersecting Ethernet rings. In Figure 13-3, Node 3 and Node 4 are intersecting nodes.
- **Non-intersecting node:** A node that belongs to only one intersecting Ethernet ring. In Figure 13-3, Node 2 is a non-intersecting node.

VLAN

ERPS supports two types of VLAN: R-APS VLAN and data VLAN.

- **R-APS VLAN:** A VLAN for transmitting ERPS packets. On a device, the ports accessing an ERPS ring belong to the R-APS VLAN, and only such ports can join the R-APS VLAN. R-APS VLANs of different ERPS rings must be different. IP address configuration is prohibited on the R-APS VLAN ports.
- **Data VLAN:** A VLAN for transmitting data packets. Both ERPS ports and non-ERPS ports can be assigned to a data VLAN.

R-APS VLANs of different ERPS rings must be configured differently to differentiate packets of different ERPS rings; otherwise, ERPS may be abnormal.

ERPS Packet

ERPS packets (also called R-APS packets) are classified into Signal Fail (SF) packets, No Request (NR) packets, No Request, RPL Blocked (NR, RB) packets, and flush packets.

- SF packet: When the link of a node is down, the node sends SF packets to notify other nodes of its link failure.
- NR packet: When the failed link is restored, the node sends an NR packet to notify the RPL owner node of its link recovery.
- (RR, RB) packet: When all nodes in an ERPS ring function properly, the RPL owner node sends (RR, RB) packets periodically.
- Flush packet: In an intersecting ring, when a topology change occurs in a subring, the intersecting nodes send flush packets to

notify other devices in the Ethernet ring to which the subring is connected.

ERPS Timer

ERPS timers include the Holdoff timer, Guard timer, and WTR timer.

- **Holdoff timer:** Is used to minimize frequent ERPS topology switching due to intermittent link failures. After you configure the Holdoff timer, ERPS performs topology switching only if the link failure still persists after the timer times out.
- **Guard timer:** Is used to prevent a device from receiving expired R-APS messages. When the device detects that a link failure is cleared, it sends link recovery packets and starts the Guard timer. During the period before timer expiration, all packets except flush packets indicating a subring topology change will be discarded.
- **Wait-to-restore (WTR) timer:** Is effective only for RPL owner devices to avoid ring status misjudgment. When an RPL owner device detects that a failure is cleared, it does perform topology switching immediately but only if the Ethernet ring is recovered after the WTR timer times out. If a ring failure is detected again before timer expiration, the RPL owner device cancels the timer and does not perform topology switching.

Overview

Feature	Description
Ring Protection	Prevents broadcast storms caused by data loops and can rapidly recover the communication between nodes in the event that a link is disconnected in the Ethernet ring.
Load Balancing	Configures multiple Ethernet subrings in one ring network and forwards the traffic of different VLANs through different Ethernet subrings to balance load.

14.4.1 Ring Protection

Ring protection prevents broadcast storms caused by data loops and can rapidly recover the communication between nodes in the event that a link is disconnected in the Ethernet ring.

Working Principle

Normal Status

- All nodes in the physical topology are connected in ring mode.
- ERPS blocks the RPL to prevent loops.
- ERPS is used to detect failures on each link between adjacent nodes.

Link Failure

- A node adjacent to a failed node detects the failure.
- The nodes adjacent to a failed link block the failed link and send SF packets to notify other nodes in the same ring.
- The R-APS (SF) packet triggers the RPL owner node to unblock the RPL port. All nodes update their MAC address entries and ARP/ND entries and the ring enters the protection state.

Link Recovery

- When a failed link is restored, adjacent nodes still block the link and send NR packets indicating that no local failure exists.
- When the RPL owner node receives the first R-APS (NR) packet, it starts the WTR timer.

- When the timer times out, the RPL owner node blocks the RPL and sends an (NR, RB) packet.
- After receiving the (NR, RB) packet, other nodes update their MAC address entries and ARP/ND entries, and the node that sends the NR packet stops periodic packet transmission and unblocks the port.
- The ring network is restored to the normal state.

Related Configuration

Configuring the R-APS VLAN

By default, no R-APS VLAN is configured.

Run the **erps raps-vlan** command to configure the R-APS VLAN (management VLAN) of an ERPS ring to transmit ERPS packets.

Configuring an ERPS Ring

Run the **rpl-port** command in R-APS VLAN mode to configure the ERPS ring mapped to an R-APS VLAN.

Configuring an RPL and an RPL Owner Node

Run the **rpl-port** command in R-APS VLAN mode to specify an RPL and an RPL owner node.

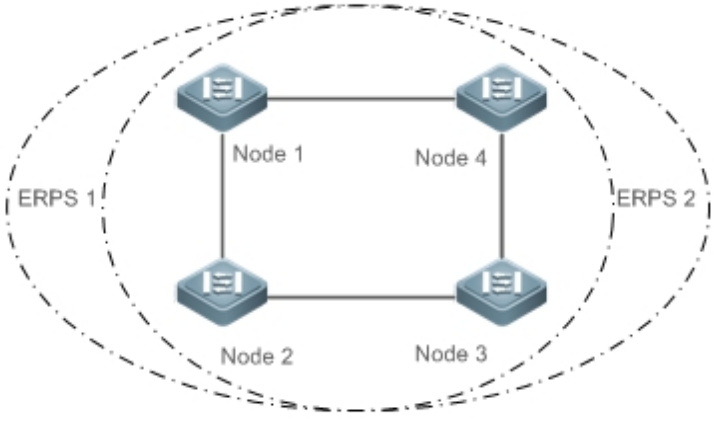
14.4.2 Load Balancing

You can configure multiple Ethernet subrings in one physical ring network and forward the traffic of different VLANs through different Ethernet subrings to balance load.

Working Principle

The multiple VLANs in a ring network can have their respective traffic forwarded by different paths through ERPS to balance load.

Figure 13-4 Single-Ring Load Balancing



In a physical ring network, multiple Ethernet rings can be configured to forward traffic of different VLANs (called protected VLANs) by different topologies to realize load balancing.

In Figure 13-4, two Ethernet rings are configured with different protected VLANs in the physical ring network. Node 1 is the RPL owner node of ERPS 1 and Node 3 is RPL owner node of ERPS 2. With such configurations, data of different VLANs can be transmitted by different links to realize single-ring load balancing.

Related Configuration

Configuring the Protected VLAN of an Ethernet Ring

Run the **protected-instance** command in R-APS VLAN mode to configure a protected VLAN set to realize load balancing.

14.5 Configuration

Configuration	Description and Command	
Single-Ring Configuration (Basic Function)	(Mandatory) Perform this configuration in global configuration mode.	
	erps enable	Enables ERPS.
	erps raps-vlan	Configures the R-APS VLAN of an Ethernet ring.
	(Mandatory) Perform this configuration in R-APS VLAN mode.	
	ring-port	Configures an ERPS ring.
	rpl-port	Configures the RPL owner node.
	state enable	Enables the specified R-APS ring.
Tangent-Ring Configuration	Tangent-ring configuration is based on single-ring configuration.	
Intersecting-Ring Configuration	(Optional) Perform this configuration in R-APS VLAN mode based on single-ring configuration.	
	associate sub-ring raps-vlan	Associates Ethernet subrings.
	sub-ring tc-propagation enable	Enables subring topology change notification.
Load Balancing Configuration	(Optional) Perform this configuration in R-APS VLAN mode based on single-ring configuration.	
	protected-instance	Configures the protected VLAN of an Ethernet ring.
ERPS Configuration Modification	(Optional) Perform this configuration in R-APS VLAN mode based on single-ring configuration.	
	timer	Modifies timer parameters.

14.5.1 Single-Ring Configuration (Basic Function)

Configuration Effect

- The single-ring scenario is the basic scenario of ERPS.
- Build an ERPS single-ring topology to realize data link redundancy.
- In an ERPS ring network, quickly switch services from a failed link to a normal link.

Notes

- Only one RPL owner node and only one RPL can be configured in one ERPS ring.
- All nodes in one ERPS ring must belong to the same R-APS VLAN.
- Only trunk ports can join an ERPS ring, and the trunk attributes cannot be modified after the port joins the ring.
- The ports in an ERPS ring do not participate in STP calculation regardless of whether the ERPS ring is enabled or not. When you configure an ERPS ring, ensure that loops will not occur when STP calculation is disabled on ports in the ring.

- ERPS does not use the same ports as RERP and REUP.

Configuration Steps

Configuring the R-APS VLAN of an Ethernet Ring

- (Mandatory) Perform this configuration in global configuration mode.
- Configure the same R-APS VLAN on all switches in the ERPS ring to transmit ERPS packets.

Configuring ERPS Ring Ports

- (Mandatory) Perform this configuration in R-APS VLAN mode.
- Configure the ports that form the ERPS ring as ERPS ring ports.

Configuring an RPL Owner Port

- (Mandatory) Perform this configuration in R-APS VLAN mode.
- Configure a single device in each ERPS ring as an RPL owner node, which will control the port to be blocked.

Enabling the Specified R-APS Ring

- (Mandatory) Perform this configuration in R-APS VLAN mode.
- Enable the specified R-APS ring in the same R-APS VLAN on each switch.

Enabling ERPS Globally

- (Mandatory) Perform this configuration in global configuration mode.
- Enable ERPS globally on each switch in the ERPS ring.

Verification

- Run the **show erps** command one each node to check the configuration.

Related Commands

Configuring the R-APS VLAN of an Ethernet Ring

Command	**erps raps-vlan** *vlan-id*
Parameter Description	*vlan-id*: R-APS VLAN ID
Command Mode	Global configuration mode
Usage Guide	ERPS takes effect in a ring only after ERPS is enabled globally and for the ring respectively.

Configuring an ERPS Ring

Command	**ring-port west** {*interface-name1* | **virtual-channel** } **east** { *interface-name2* | **virtual-channel**}
Parameter Description	*interface-name1*: Indicates the name of the West port. *interface-name2*: Indicates the name of the East port. **virtual-channel:** Assigns a port to a virtual link.

Command Mode	R-APS VLAN mode
Usage Guide	The R-APS VLAN must be the unused VLAN on a device. VLAN 1 cannot be configured as the R-APS VLAN. In an Ethernet ring, different devices must be configured with the same R-APS VLAN. If you need to transparently transmit ERPS packets on a device not configured with ERPS, ensure that only the two ports on the device connected to the ERPS ring allow packets from the R-APS VLAN of the ERPS ring to pass through. Otherwise, packets from other VLANs may be transparently transmitted to the R-APS VLAN, causing impact on the ERPS ring.

Configuring an RPL Owner Port

Command	**rpl-port** { **west** | **east** } **rpl-owner**
Parameter Description	**west**: Specifies the West port as an RPL owner port. **east**: Specifies the East port as an RPL owner port.
Command Mode	R-APS VLAN mode
Usage Guide	Each ring can be configured with only one RPL and only one RPL owner node.

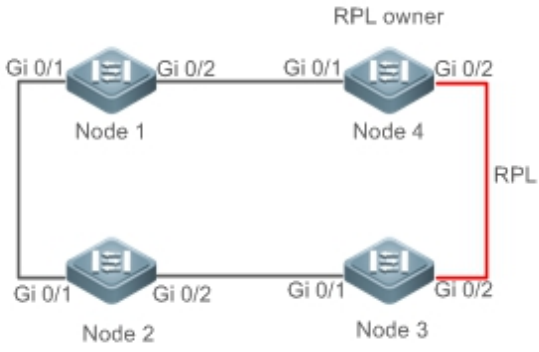
Enabling the Specified R-APS Ring

Command	**state enable**
Parameter Description	N/A
Command Mode	R-APS VLAN mode
Usage Guide	ERPS takes effect in a ring only after ERPS is enabled globally and for the ring respectively.

Enabling ERPS Globally

Command	**erps enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	ERPS takes effect in a ring only after ERPS is enabled globally and for the ring respectively.

Configuration Example

Scenario	
Configuration Steps	● Configure the R-APS VLAN in privileged mode. ● Configure the link mode of ports in the Ethernet ring. ● Enter R-APS VLAN mode and configure the ports to be added to the Ethernet ring and participate in ERPS calculation. ● Specify the RPL owner port. ● Enable ERPS in the specified ring. ● Enable ERPS globally.
Node 1	# Enter privileged mode. FS# configure terminal # Configure the R-APS VLAN. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Enable ERPS in the specified ring. FS(config-erps 4093)# state enable # Enable ERPS globally. FS(config-erps 4093)# exit

	FS(config)# erps enable
Node 2	The configuration on Node 2 is the same as that on Node 1.
Node 3	The configuration on Node 3 is the same as that on Node 1.
Node 4	# Enter privileged mode. FS# configure terminal # Configure the R-APS VLAN. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Specify the RPL owner port. FS(config-erps 4093)# rpl-port east rpl-owner # Enable ERPS in the specified ring. FS(config-erps 4093)# state enable FS(config-erps 4093)# exit # Enable ERPS globally. FS(config)# erps enable
Verification	Run the **show erps** command one each node to check the configuration. The configuration on Node 1 and Node 4 is used as an example.
Node 1	FS# show erps ERPS Information Global Status : Enabled Link monitored by : Not Oam ----------

	R-APS VLAN : 4093 Ring Status : Enabled West Port : Gi 0/1 (Forwardin) East Port : Gi 0/2 (Forwardin) RPL Port : None Protected VLANs : ALL RPL Owner : Enabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 2 minutes Current Ring State : Idle Associate R-APS VLAN :
Node 4	FS# show erps ERPS Information Global Status : Enabled Link monitored by : Not Oam ---------- R-APS VLAN : 4093 Ring Status : Enabled West Port : Gi 0/1 (Forwardin) East Port : Gi 0/2 (Blocking) RPL Port : East Port Protected VLANs : ALL RPL Owner : Enabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 2 minutes Current Ring State : Idle Associate R-APS VLAN :

Common Errors

- The R-APS ring has been enabled but ERPS is not enabled globally, so ERPS still does not take effect.
- Multiple RPL owner nodes are configured in one ring.
- Different R-APS VLANs are configured for the nodes in one ring.

14.5.2 Tangent-Ring Configuration

Configuration Effect

- Configure a tangent ring that consists of two ERPS rings sharing one device to realize data link redundancy.
- Quickly switch services from a failed link in one ERPS ring to a normal link.

Notes

- The tangent-ring configuration is basically the same as the single-ring configuration. You only need to associate the two ERPS rings on the tangent node.
- Only one RPL owner node and only one RPL can be configured in each ERPS ring.
- All nodes in one ERPS ring must belong to the same R-APS VLAN.
- Only trunk ports can join an ERPS ring, and the trunk attributes cannot be modified after the port joins the ring.
- The ports in an ERPS ring do not participate in STP calculation regardless of whether the ERPS ring is enabled or not. When you configure an ERPS ring, ensure that loops will not occur when STP calculation is disabled on ports in the ring.
- ERPS does not use the same ports as RERP and REUP.

Configuration Steps

- The tangent-ring configuration is basically the same as the single-ring configuration. You only need to associate the two ERPS rings on the tangent node.

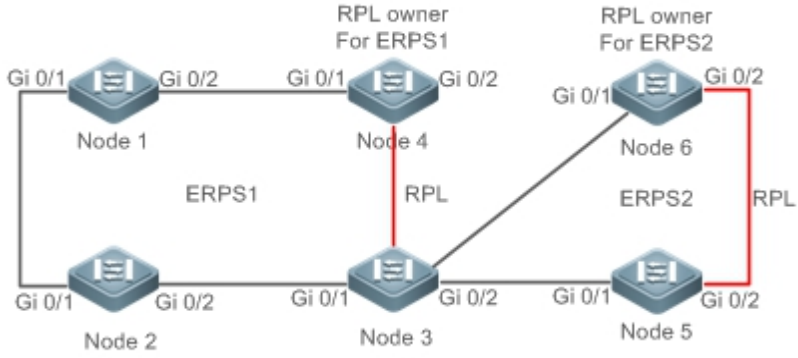
Verification

- Run the **show erps** command one each node to check the configuration.

Related Commands

- See the commands in section 14.4.1 "Single-Ring Configuration (Basic Function)."

Configuration Example

Scenario	
Configuration Steps	• Configure the R-APS VLAN in privileged mode. • Configure the link mode of ports in the Ethernet ring.

	● Enter R-APS VLAN mode and configure the ports to be added to the Ethernet ring and participate in ERPS calculation. ● Specify the RPL owner port. ● Enable ERPS in the specified ring. ● Enable ERPS globally.
Node 1	# Enter privileged mode. FS# configure terminal # Configure R-APS VLAN 4093. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Enable ERPS in the specified ring. FS(config-erps 4093)# state enable FS(config-erps 4093)# exit # Enable ERPS globally. FS(config)# erps enable
Node 2	The configuration on Node 2 is the same as that on Node 1.
Node 3	FS# configure terminal # Configure R-APS VLAN 4093. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit

	FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 FS(config-erps 4093)# state enable FS(config-erps 4093)# exit # Configure R-APS VLAN 100. FS(config)# erps raps-vlan 100 FS(config-erps 100)# exit FS(config)# interface gigabitEthernet 0/3 FS(config-if-gigabitEthernet 0/3)# switchport mode trunk FS(config-if-gigabitEthernet 0/3)# exit FS(config)# interface gigabitEthernet 0/4 FS(config-if-gigabitEthernet 0/4)# switchport mode trunk FS(config-if-gigabitEthernet 0/4)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 100 FS(config-erps 100)# ring-port west gigabitEthernet 0/3 east gigabitEthernet 0/4 FS(config-erps 100)# state enable FS(config-erps 4093)# exit FS(config)# erps enable
Node 4	FS# configure terminal # Configure R-APS VLAN 4093. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit

	# Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Specify the RPL owner port. FS(config-erps 4093)# rpl-port east rpl-owner FS(config-erps 4093)# state enable FS(config-erps 4093)# exit FS(config)# erps enable
Node 5	FS# configure terminal # Configure R-APS VLAN 100. FS(config)# erps raps-vlan 100 FS(config-erps 100)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 100 FS(config-erps 100)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 FS(config-erps 100)# state enable FS(config-erps 100)# exit FS(config)# erps enable
Node 6	FS# configure terminal # Configure R-APS VLAN 100. FS(config)# erps raps-vlan 100 FS(config-erps 100)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit

	FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 100 FS(config-erps 100)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Specify the RPL owner port. FS(config-erps 100)# rpl-port east rpl-owner FS(config-erps 100)# state enable FS(config)# erps enable
Verification	Run the **show erps** command one each node to check the configuration. The configuration on Node 3 is used as an example.
	FS# show erps ERPS Information Global Status : Enabled Link monitored by : Not Oam ---------- R-APS VLAN : 100 Ring Status : Enabled West Port : Gi 0/3 (Forwarding) East Port : Gi 0/4 (Forwarding) RPL Port : None Protected VLANs : ALL RPL Owner : Disabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 2 minutes Current Ring State : Idle Associate R-APS VLAN : ---------- R-APS VLAN : 4093 Ring Status : Enabled West Port : Gi 0/1 (Forwarding) East Port : Gi 0/2 (Forwarding)

	RPL Port : East Port Protected VLANs : ALL RPL Owner : Disabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 2 minutes Current Ring State : Idle Associate R-APS VLAN :

Common Errors

- The R-APS ring has been enabled but ERPS is not enabled globally, so ERPS still does not take effect.
- Multiple RPL owner nodes are configured in one ring.
- Different R-APS VLANs are configured for the nodes in one ring.

14.5.3 Intersecting-Ring Configuration

Configuration Effect

- Configure multiple ERPS rings to share links, thus realizing data link redundancy.
- Quickly switch services from a failed link in one ERPS ring to a normal link.

Notes

- Only one RPL owner node and only one RPL can be configured in each ERPS ring.
- All nodes in one ERPS ring must belong to the same R-APS VLAN.
- All nodes in the Ethernet ring must be associated with their respective subrings.
- Only trunk ports can join an ERPS ring, and the trunk attributes cannot be modified after the port joins the ring.
- The ports in an ERPS ring do not participate in STP calculation regardless of whether the ERPS ring is enabled or not. When you configure an ERPS ring, ensure that loops will not occur when STP calculation is disabled on ports in the ring.
- ERPS does not use the same ports as RERP and REUP.

Configuration Steps

Perform the following configuration after you complete the single-ring configuration described above:

Enabling Subring Topology Change Notification

- (Optional) Perform this configuration in R-APS VLAN mode.
- Enable subring topology change notification on intersecting nodes.
- If the link between intersecting nodes is faulty or blocked in the event of a subring topology change, the intersecting nodes will send packets to instruct the nodes in other Ethernet rings associated with the subring to update the topology.

Associating Ethernet Subrings

- (Optional) Perform this configuration in R-APS VLAN mode.
- Associate nodes in the main ring with Ethernet subrings.
- After nodes are associated with Ethernet subrings, ERPS packets of the subrings can be transmitted to other Ethernet rings.

Verification

- Run the **show erps** command one each node to check the configuration.

Related Commands

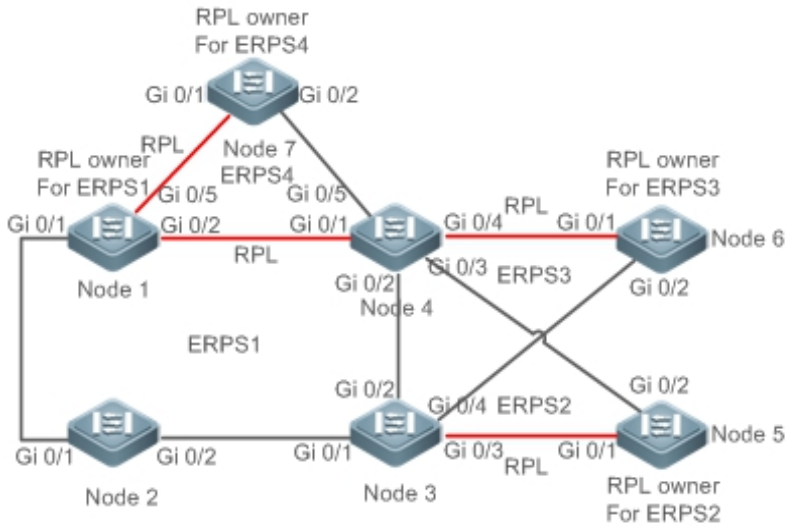
Enabling Subring Topology Change Notification

Command	**sub-ring tc-propagation enable**
Parameter Description	N/A
Command Mode	R-APS VLAN mode
Usage Guide	Run this command only on intersecting nodes.

Associating Ethernet Subrings

Command	**associate sub-ring raps-vlan** *vlan-list*
Parameter Description	*vlan-list*: Indicates the R-APS VLANs of subrings.
Command Mode	R-APS VLAN mode
Usage Guide	Run this command on all nodes in the Ethernet ring to allow its subrings to transmit ERPS packets to the Ethernet ring. After nodes are associated with subrings, ERPS packets of the subrings can be transmitted to other Ethernet rings. You can also use the command provided by the VLAN module to configure VLAN and its member ports to allow ERPS packets of subrings to be transmitted to other Ethernet rings while avoiding information leakage to user networks.

Configuration Example

Scenario	
Configuration Steps	● Configure the R-APS VLAN in privileged mode. ● Configure the link mode of ports in the Ethernet ring. ● Enter R-APS VLAN mode and configure the ports to be added to the Ethernet ring and participate in ERPS calculation. ● Specify the RPL owner port. ● Enable ERPS in the specified ring. ● Associate nodes in the Ethernet ring with subrings. ● Enable subring topology change notification on intersecting nodes. ● Enable ERPS globally.
Node 1	# Enter privileged mode. FS# configure terminal # Configure R-APS VLAN 4093. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation.

	FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Specify the port and RPL owner node for the RPL. FS(config-erps 4093)# rpl-port east rpl-owner # Enable ERPS in the specified ring. FS(config-erps 4093)# state enable # Enable ERPS globally. FS(config-erps 4093)# exit FS(config)# erps enable # Configure the R-APS VLAN of the subring ERPS 4. FS(config)# erps raps-vlan 300 FS(config-erps 300)# exit # Configure the link mode of ports in ERPS 4. FS(config)# interface gigabitEthernet 0/5 FS(config-if-gigabitEthernet 0/5)# switchport mode trunk FS(config-if-gigabitEthernet 0/5)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 300 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 300)# ring-port west gigabitEthernet 0/5 east virtual-channel # Enable ERPS in ERPS 4. FS(config-erps 300)# state enable # Associate ERPS 1 with ERPS 2, ERPS 3, and ERPS 4. FS(config-erps 300)# exit FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# associate sub-ring raps-vlan 100,200,300
Node 2	# Enter privileged mode. FS# configure terminal # Configure R-APS VLAN 4093. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit

	FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Enable ERPS in the specified ring. FS(config-erps 4093)# state enable # Enable ERPS globally. FS(config-erps 4093)# exit FS(config)# erps enable # Associate ERPS 1 with ERPS 2, ERPS 3, and ERPS 4. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# associate sub-ring raps-vlan 100,200,300
Node 3	# Perform the following configuration on Node 3 based on the configuration on Node 2: # Enter privileged mode. FS# configure terminal # Configure the R-APS VLAN of the subring ERPS 2. FS(config)# erps raps-vlan 100 FS(config-erps 100)# exit # Configure the link mode of ports in ERPS 2. FS(config)# interface gigabitEthernet 0/3 FS(config-if-gigabitEthernet 0/3)# switchport mode trunk FS(config-if-gigabitEthernet 0/3)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 100 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 100)# ring-port west virtual-channel east gigabitEthernet 0/3 # Enable ERPS in ERPS 2. FS(config-erps 100)# state enable # Configure the R-APS VLAN of the subring ERPS 3. FS(config)# erps raps-vlan 200 FS(config-erps 200)# exit

	# Configure the link mode of ports in ERPS 3. FS(config)# interface gigabitEthernet 0/4 FS(config-if-gigabitEthernet 0/4)# switchport mode trunk FS(config-if-gigabitEthernet 0/4)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 200 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 200)# ring-port west virtual-channel east gigabitEthernet 0/4 # Enable ERPS in ERPS 2. FS(config-erps 200)# state enable # Associate the Ethernet subrings ERPS 2, ERPS 3, and ERPS 4. FS(config-erps 200)# exit FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# associate sub-ring raps-vlan 100,200,300
Node 4	# Perform the following configuration on Node 4 based on the configuration on Node 2. # Enter privileged mode. FS# configure terminal # Configure the R-APS VLAN of the subring ERPS 2. FS(config)# erps raps-vlan 100 FS(config-erps 100)# exit # Configure the link mode of ports in ERPS 2. FS(config)# interface gigabitEthernet 0/3 FS(config-if-gigabitEthernet 0/3)# switchport mode trunk FS(config-if-gigabitEthernet 0/3)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 100 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 100)# ring-port west virtual-channel east gigabitEthernet 0/3 # Enable ERPS in ERPS 2. FS(config-erps 100)# state enable # Configure the R-APS VLAN of the subring ERPS 3. FS(config)# erps raps-vlan 200 FS(config-erps 200)# exit # Configure the link mode of ports in ERPS 3.

	FS(config)# interface gigabitEthernet 0/4 FS(config-if-gigabitEthernet 0/4)# switchport mode trunk FS(config-if-gigabitEthernet 0/4)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 200 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 200)# ring-port west virtual-channel east gigabitEthernet 0/4 # Enable ERPS in ERPS 3. FS(config-erps 200)# state enable # Configure the R-APS VLAN of the subring ERPS 4. FS(config-erps 200)# exit FS(config)# erps raps-vlan 300 FS(config-erps 300)# exit # Configure the link mode of ports in ERPS 4. FS(config)# interface gigabitEthernet 0/5 FS(config-if-gigabitEthernet 0/5)# switchport mode trunk FS(config-if-gigabitEthernet 0/5)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 300 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 300)# ring-port west virtual-channel east gigabitEthernet 0/5 # Enable ERPS in ERPS 4. FS(config-erps 300)# state enable # Associate the Ethernet subrings ERPS 2, ERPS 3, and ERPS 4. FS(config-erps 300)# exit FS(config)# erps raps-vlan 4093 FS(config-erps4093)# associate sub-ring raps-vlan 100,200,300
Node 5	# Enter privileged mode. FS# configure terminal # Configure the R-APS VLAN. FS(config)# erps raps-vlan 100 FS(config-erps 100)# end # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/1

	FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 100 # Configure the ports to be added to the Ethernet ring and participate in ERPS calculation. FS(config-erps 100)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Specify the port and RPL owner node for the RPL. FS(config-erps 100)# rpl-port east rpl-owner # Enable ERPS in the specified ring. FS(config-erps 100)# state enable # Enable ERPS globally. FS(config-erps 100)# exit FS(config)# erps enable
Node 6	# The configuration on Node 6 is basically the same as that on Node 5, except that you need to change the R-APS VLAN to VLAN 200.
Node 7	# The configuration on Node 7 is basically the same as that on Node 5, except that you need to change the R-APS VLAN to VLAN 300.
Verification	Run the **show erps** command one each node to check the configuration. The configuration on Node 3 is used as an example.
	FS# show erps ERPS Information Global Status : Enabled Link monitored by : Not Oam ---------- R-APS VLAN : 100 Ring Status : Enabled West Port : Virtual Channel East Port : Gi 0/3 (Forwarding) RPL Port : None Protected VLANs : ALL RPL Owner : Disabled Holdoff Time : 0 milliseconds

```
Guard Time                          : 500 milliseconds
WTR Time                            : 2 minutes
Current Ring State                  : Idle
Associate R-APS VLAN                :
-----------------------------------------
R-APS VLAN                          : 200
Ring Status                         : Enabled
West Port                           : Virtual Channel
East Port                           : Gi 0/4 (Forwarding)
RPL Port                            : None
Protected VLANs                     : ALL
RPL Owner                           : Disabled
Holdoff Time                        : 0 milliseconds
Guard Time                          : 500 milliseconds
WTR Time                            : 2 minutes
Current Ring State                  : Idle
Associate R-APS VLAN                :
-----------------------------------------
R-APS VLAN                          : 4093
Ring Status                         : Enabled
West Port                           : Gi 0/1 (Forwarding)
East Port                           : Gi 0/2 (Blocking)
RPL Port                            : East Port
Protected VLANs                     : ALL
RPL Owner                           : Disabled
Holdoff Time                        : 0 milliseconds
Guard Time                          : 500 milliseconds
WTR Time                            : 2 minutes
Current Ring State                  : Idle
Associate R-APS VLAN                : 100,200,300
```

Common Errors

- The R-APS ring has been enabled but ERPS is not enabled globally, so ERPS still does not take effect.
- Multiple RPL owner nodes are configured in one ERPS ring.
- Different R-APS VLANs are configured for the nodes in one ERPS ring.

- The nodes in the man ring are not associated with Ethernet subrings.

14.5.4 Load Balancing Configuration

Configuration Effect

- Control the direction of data flows in an ERPS ring to realize load balancing.
- When a link in the ring network enabled with load balancing fails, the traffic can be quickly switched to a normal link.

Notes

- Before you configure load balancing, configure the VLAN-instance relationship in MST configuration mode.
- When you configure load balancing, add all data VLANs of the devices to the ERPS protected VLAN list; otherwise, any unprotected VLAN will cause loops.
- Only trunk ports can join an ERPS ring, and the trunk attributes cannot be modified after the port joins the ring.
- The ports in an ERPS ring do not participate in STP calculation regardless of whether the ERPS ring is enabled or not. When you configure an ERPS ring, ensure that loops will not occur when STP calculation is disabled on ports in the ring.
- ERPS does not use the same ports as RERP and REUP.

Configuration Steps

Perform the following configuration after you complete the single-ring configuration described above:

Configuring the Protected VLAN of an Ethernet Ring

- (Optional) Perform this configuration in global configuration mode.
- When you configure load balancing for an Ethernet ring, you must specify the protected VLAN.

Verification

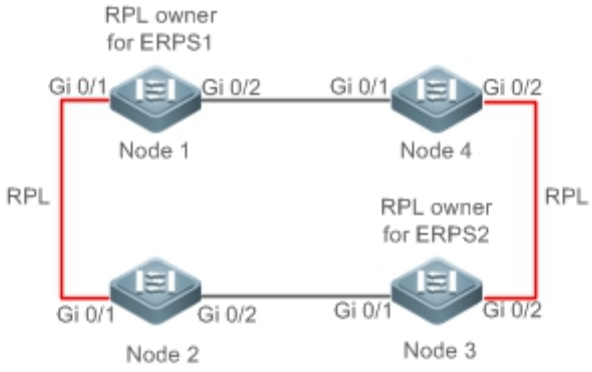
- Run the **show erps** command one each node to check the configuration.

Related Commands

Configuring the Protected VLAN of an Ethernet Ring

Command	**protected-instance** *instance-id-list*
Parameter Description	*instance-id-list*: Indicates the instance protected by the Ethernet ring.
Command Mode	R-APS VLAN mode
Usage Guide	The protected instance of the Ethernet ring is the protected VLAN.

Configuration Example

Scenario	
Configuration Steps	● Configure the R-APS VLAN in privileged mode. ● Configure the link mode of ports in the Ethernet ring. ● Configure the protected VLAN of the Ethernet ring. ● Enter R-APS VLAN mode and configure the ports to be added to the Ethernet ring and participate in ERPS calculation. ● Specify the RPL owner port. ● Enable ERPS in the specified ring. ● Enable ERPS globally.
Node 1	# Enter privileged mode. FS# configure terminal # Configure the Ethernet subring ERPS 1 as follows: # Configure the link mode of ports in ERPS 1. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# switchport mode trunk FS(config-if-gigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# switchport mode trunk FS(config-if-gigabitEthernet 0/2)# exit # Configure the protected VLAN, RPL owner port, and RPL of ERPS 1. FS(config)# spanning-tree mst configuration FS(config-mst)# instance 1 vlan 1-2000 FS(config-mst)# exit FS(config)# erps raps-vlan 100 FS(config-erps 100)# protected-instance 1 FS(config-erps 100)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 FS(config-erps 100)# rpl-port west rpl-owner # Configure the Ethernet subring ERPS 2 as follows:

	# Configure the ports to be added to ERPS 2 and participate in ERPS calculation. FS(config)# spanning-tree mst configuration FS(config-mst)# instance 2 vlan 2001-4094 FS(config-mst)# exit FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# protected-instance 2 FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/2 # Enable ERPS in ERPS 2 and globally respectively. FS(config-erps 4093)# state enable FS(config-erps 4093)# exit FS(config)# erps enable
Node 2	# The configuration on Node 2 is the same as that on Node 1, except that RPL configuration is not required on Node 2.
Node 3	# The configuration on Node 3 is the same as that on Node 1, except that RPL configuration is not required on Node 3. # Configure the RPL of ERPS 2 on Node 3. The RPL of ERPS 1 does not need to be configured on Node 3. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# rpl-port east rpl-owner
Node 4	The configuration on Node 4 is the same as that on Node 2.
Verification	Run the **show erps** command one each node to check the configuration. The configuration on Node 1 is used as an example.
Node 1	FS# show erps ERPS Information Global Status : Enabled Link monitored by : Not Oam ---------- R-APS VLAN : 200 Ring Status : Enabled West Port : Gi 0/1 (Blocking) East Port : Gi 0/2 (Forwarding) RPL Port : West Port Protected VLANs : 1-2000 RPL Owner : Enabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 2 minutes

```
Current Ring State          : Idle
Associate R-APS VLAN        :
-----------------------------------------
R-APS VLAN                  : 4093
Ring Status                 : Enabled
West Port                   : Gi 0/1 (Forwarding)
East Port                   : Gi 0/2 (Blocking)
RPL Port                    : West Port
Protected VLANs             : 2001-4094
RPL Owner                   : Enabled
Holdoff Time                : 0 milliseconds
Guard Time                  : 500 milliseconds
WTR Time                    : 2 minutes
Current Ring State          : Idle
Associate R-APS VLAN        :
```

Common Errors

- The R-APS ring has been enabled but ERPS is not enabled globally, so ERPS still does not take effect.
- Multiple RPL owner nodes are configured in one ERPS ring.
- Different R-APS VLANs are configured for the nodes in one ERPS ring.

14.5.5 ERPS Configuration Modification

Configuration Effect

- Switch configuration smoothly when the ERPS ring topology is changed.

Notes

- When you modify the ERPS configuration on a device, to avoid loops, first run the **shutdown** command to shut down an ERPS port in the ring. When the configuration is completed, run the **no shutdown** command to restart the port.
- All nodes in one ERPS ring must belong to the same R-APS VLAN.
- If you only need to modify the ERPS timers, skip this section.

Configuration Steps

Run the **shutdown** command to shut down an ERPS port and disable ERPS. Then modify the ERPS configuration according to section 14.4.1 "Single-Ring Configuration (Basic Function)" and complete the following settings, which are optional.

Configuring the Holdoff Timer, Guard Timer, and WRT Timer

- Optional.

- Perform this configuration in R-APS VLAN mode based on the actual application requirements.

Verification

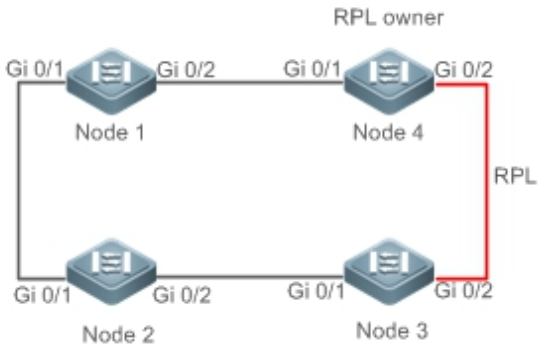
- Run the **show erps** command one each node to check the configuration.

Related Commands

Configuring the Holdoff Timer, Guard Timer, and WRT Timer

Command	**timer** { **holdoff-time** *interval1* | **guard-time** *interval2* | **wtr-time** *interval3* }
Parameter Description	*interval1:* Indicates the Holdoff timer interval. The value ranges from 0 to 100, in the unit of 100 milliseconds. The default value is 0. *interval2:* Indicates the Guard timer interval. The value ranges from 1 to 200, in the unit of 10 milliseconds. The default value is 50. *interval3:* Indicates the WTR timer interval. The value ranges from 1 to 12, in the unit of minutes. The default value is 2.
Command Mode	R-APS VLAN mode
Usage Guide	• Holdoff timer: Is used to minimize frequent ERPS topology switching due to intermittent link failures. After you configure the Holdoff timer, ERPS performs topology switching only if the link failure still persists after the timer times out. • Guard timer: Is used to prevent a device from receiving expired R-APS messages. When the device detects that a link failure is cleared, it sends link recovery packets and starts the Guard timer. During the period before timer expiration, all packets except flush packets indicating a subring topology change will be discarded. • WTR timer: Is effective only for RPL owner devices to avoid ring status misjudgment. When an RPL owner device detects that a failure is cleared, it does perform topology switching immediately but only if the Ethernet ring is recovered after the WTR timer times out. If a ring failure is detected again before timer expiration, the RPL owner device cancels the timer and does not perform topology switching.

Configuration Example

Scenario	
Configuration Steps	• ERPS configuration exists in the ring. The ERPS ports need to be switched because of a physical topology change. • Run the **shutdown** command to shut down a link in the ring and configure the link mode of ports after switching. • Disable ERPS in the ring in R-APS VLAN mode.

	● Reconfigure the ports that will participate in ERPS calculation. ● Enable ERPS in the ring. ● Modify the ERPS timers.
Node 1	# Enter privileged mode. FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. # Shutdown a link in the ring in interface configuration mode to avoid loops. FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)# shutdown FS(config-if-gigabitEthernet 0/1)# exit # Configure the link mode of ports in the Ethernet ring. FS(config)# interface gigabitEthernet 0/3 FS(config-if-gigabitEthernet 0/3)# switchport mode trunk FS(config-if-gigabitEthernet 0/3)# exit # Enter ERPS configuration mode. FS(config)# erps raps-vlan 4093 # Disable ERPS. FS(config-erps 4093)# no state enable # Delete the previous ring configuration. FS(config-erps 4093)# no ring-port # Reconfigure the ports that will participate in ERPS calculation. Change Gig 0/2 to Gig 0/3. FS(config-erps 4093)# ring-port west gigabitEthernet 0/1 east gigabitEthernet 0/3 # Enable ERPS. FS(config-erps 4093)# state enable
Node 4	# Enter privileged mode. FS# configure terminal # Modify timers in ERPS configuration mode. FS(config)# erps raps-vlan 4093 FS(config-erps 4093)# timer wtr-time 1
	Wait for 1 minute. When the ERPS ring is restored to Idle, run the **show erps** command on Node 1 and Node 4 to check the configuration.
Node 1	FS# show erps ERPS Information Global Status : Enabled

	Link monitored by : Not Oam ---------- R-APS VLAN : 4093 Ring Status : Enabled West Port : Gi 0/1 (Forwardin) East Port : Gi 0/3 (Forwardin) RPL Port : None Protected VLANs : ALL RPL Owner : Enabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 2 minutes Current Ring State : Idle Associate R-APS VLAN :
Node 4	FS# show erps ERPS Information Global Status : Enabled Link monitored by : Not Oam ---------- R-APS VLAN : 4093 Ring Status : Enabled West Port : Gi 0/1 (Forwardin) East Port : Gi 0/2 (Blocking) RPL Port : East Port Protected VLANs : ALL RPL Owner : Enabled Holdoff Time : 0 milliseconds Guard Time : 500 milliseconds WTR Time : 1 minutes Current Ring State : Idle Associate R-APS VLAN :

Common Errors

- When the configuration is completed, the R-APS ring is not enabled again or the shutdown ports are not restarted by using the **no shutdown** command.

14.6 Monitoring

Displaying

Description	Command
Displays the ERPS configuration and status of devices.	**show erps** [**global** | **raps_vlan** *vlan-id* [*sub_ring*]]

Chapter3 IP Address & Application Configuration

1. Configuring IP Address and Service
2. Configuring ARP
3. Configuring IPv6
4. Configuring DHCP
5. Configuring DHCPv6
6. Configuring DNS
7. Configuring FTP Server
8. Configuring FTP Client
9. Configuring TFTP
10. Configuration Tunnel
11. Configuring Network Communication Detection Tools
12. Configuring TCP
13. Configuring IPv4/IPv6 REF

1 Configuring IP Addresses and Services

1.1 Overview

Internet Protocol (IP) sends packets to the destination from the source by using logical (or virtual) addresses, namely IP addresses. At the network layer, routers forward packets based on IP addresses.

Protocols and Standards

- RFC 1918: Address Allocation for Private Internets
- RFC 1166: Internet Numbers

1.2 Applications

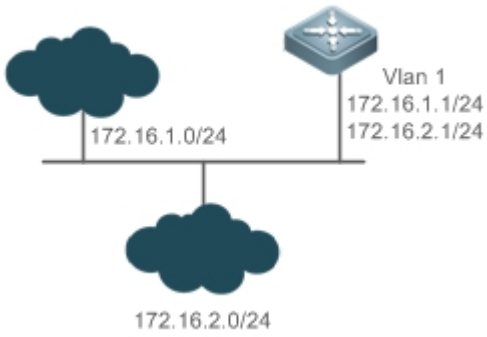
Application	Description
Configuring an IP Address for Communication	Two networks communicate through one switch interface.

1.2.1 Configuring an IP Address for Communication

Scenario

A switch is connected to a Local Area Network (LAN), which is divided into two network segments, namely, 172.16.1.0/24 and 172.16.2.0/24. Computers in the two network segments can communicate with the Internet through switches and computers between the two network segments can communicate with each other.

Figure 1- 1 Configuring IP Addresses



Deployment

- Configure two IP addresses on VLAN1. One is a primary IP address and the other is a secondary IP address.
- On hosts in the network segment 172.16.1.0/24, set the gateway to 172.16.1.1; on hosts in the network segment 172.16.2.0/24, set the gateway to 172.16.2.1.

1.3 Features

Basic Concepts

IP Address

An IP address consists of 32 bits in binary. To facilitate writing and description, an IP address is generally expressed in decimal. When expressed in decimal, an IP address is divided into four groups, with eight bits in each group. The value range of each group is from 0 to 255, and groups are separated by a full stop ".". For example, "192.168.1.1" is an IP address expressed in decimal.

IP addresses are used for interconnection at the IP layer. A 32-bit IP address consists of two parts, namely, the network bits and the host bits. Based on the values of the first several bits in the network part, IP addresses in use can be classified into four classes.

For a class A address, the most significant bit is 0.7 bits indicate a network ID, and 24 bits indicate a local address. There are 128 class A networks in total.

Figure 1- 2

		8	16 24 32
Class A IP address	0	Network ID	Host ID

For a class B address, the first two most significant bits are 10.14 bits indicate a network ID, and 16 bits indicate a local address. There are 16,348 class B networks in total.

Figure 1- 3

			8	16 24 32
Class B IP address	1	0	Network ID	Host ID

For a class C address, the first three most significant bits are 110.21 bits indicate a network ID, and 8 bits indicate a local address. There are 2,097,152 class C networks in total.

Figure 1- 4

				8	16 24 32
Class C IP address	1	1	0	Network ID	Host ID

For a class D address, the first four most significant bits are 1110 and other bits indicate a multicast address.

Figure 1- 5

					8	16 24 32
Class D IP address	1	1	1	0	Multicast address	

The addresses with the first four most significant bits 1111 cannot be assigned. These addresses are called class E addresses and are reserved.

When IP addresses are planned during network construction, IP addresses must be assigned based on the property of the network to be built. If the network needs to be connected to the Internet, users should apply for IP addresses to the corresponding agency. In China, you can apply to China Internet Network Information Center (CNNIC) for IP addresses. Internet Corporation for Assigned Names and Numbers (ICANN) is the final organization responsible for IP address assignment. If the network to be built is an internal private network, users do not need to apply for IP addresses. However, IP addresses cannot be assigned at random. It is recommended to assign dedicated private network addresses.

The following table lists reserved and available addresses.

Class	Address Range	Status
Class A network	0.0.0.0 - 0.255.255.255	Reserved
	1.0.0.0 - 126.255.255.255	Available
	127.0.0.0 - 127.255.255.255	Reserved
Class B network	128.0.0.0 - 191.254.255.255	Available
	191.255.0.0 - 191.255.255.255	Reserved
Class C network	192.0.0.0 - 192.0.0.255	Reserved
	192.0.1.0 - 223.255.254.255	Available
	223.255.255.0 - 223.255.255.255	Reserved
Class D network	224.0.0.0 - 239.255.255.255	Multicast address
Class E network	240.0.0.0 - 255.255.255.254	Reserved
	255.255.255.255	Broadcast address

Three address ranges are dedicated to private networks. These addresses are not used in the Internet. If the networks to which these addresses are assigned need to be connected to the Internet, these IP addresses need to be converted into valid Internet addresses. The following table lists private address ranges. Private network addresses are defined in RFC 1918.

Class	Address Range	Status
Class A network	10.0.0.0 - 10.255.255.255	1 class A network
Class B network	172.16.0.0 - 172.31.255.255	16 class B networks
Class C network	192.168.0.0 - 192.168.255.255	256 class C networks

For assignment of IP addresses, TCP/UDP ports, and other codes, refer to RFC 1166.

Subnet Mask

A subnet mask is also a 32-bit value. The bits that identify the IP address are the network address. In a subnet mask, the IP address bits corresponding to the bits whose values are 1s are the network address, and the IP address bits corresponding to the bits whose values are 0s are the host address. For example, for class A networks, the subnet mask is 255.0.0.0. By using network masks, you can divide a network into several subnets. Subnetting means to use some bits of the host address as the network address, thus decreasing the host capacity, and increasing the number of networks. In this case, network masks are called subnet masks.

Broadcast Packet

Broadcast packets refer to the packets destined for all hosts on a physical network. FS products support two types of broadcast packets: (1) directed broadcast, which indicates that all hosts on the specified network are packet receivers and the host bits of a destination address are all 1s; (2) limited broadcast, which indicates that all hosts on all networks are packet receivers and the 32 bits of a destination address are all 1s.

ICMP Packet

Internet Control Message Protocol (ICMP) is a sub-protocol in the TCP/IP suite for transmitting control messages between IP hosts and network devices. It is mainly used to notify corresponding devices when the network performance becomes abnormal.

TTL

Time To Live (TTL) refers to the number of network segments where packets are allowed to pass before the packets are discarded. The TTL is a value in an IP packet. It informs the network whether packets should be discarded as the packets stay on the network for a long

time.

Features

Feature	Description
IP Address	The IP protocol can run on an interface only after the interface is configured with an IP address.
Broadcast Packet Processing	Broadcast addresses are configured and broadcast packets are forwarded and processed.
Sending ICMP Packets	ICMP packets are sent and received.
Limiting Transmission Rate of ICMP Error Packets	This function prevents Denial of Service (DoS) attacks.
IP MTU	Maximum Transmission Unit (MTU) of IP packets on an interface is configured.
IP TTL	The TTL of unicast packets and broadcast packets is configured.
IP Source Route	Source routes are checked.
IP Address Pool	An IP address is assigned to the peer end during PPP negotiation.

1.3.1 IP Address

IP addresses are obtained on an interface in the following ways:

8. Manually configuring IP addresses
9. Obtaining IP addresses through DHCP
10. Obtaining IP addresses through PPP negotiation

These approaches are mutually exclusive. If you configure a new approach to obtain an IP address , the old IP address will be overwritten.

For details on how to obtain IP addresses through DHCP, see the "DHCP" chapter. The following describes the other three approaches for obtaining IP addresses.

Configuring the IP Address for an Interface

A device can receive and send IP packets only after the device is configured with an IP address. Only the interface configured with an IP address can run the IP protocol.

Configuring Multiple IP Addresses for an Interface

FS products support multiple IP address configuration on one interface, of which one is a primary IP address and the others are secondary IP addresses. Theoretically, the number of secondary IP addresses is not limited. However, secondary IP addresses must belong to different networks and secondary IP addresses must be in different networks from primary IP addresses. In network construction, secondary IP addresses are often used in the following circumstances:

- A network does not have enough host addresses. For example, a LAN now needs one class C network to allocate 254 addresses. However, when the number of hosts exceeds 254, one class C network is not enough and another class C network is needed. In this case, two networks need to be connected. Therefore, more IP addresses are needed.
- Many old networks are based on L2 bridged networks without subnetting. You can use secondary IP addresses to upgrade the network to a routing network based on IP layer. For each subnet, one device is configured with one IP address.
- When two subnets of one network are isolated by another network, you can connect the isolated subnets by creating a subnet of

the isolated network and configuring a secondary address. One subnet cannot be configured on two or more interfaces of a device.

Obtaining an IP Addresses through PPP Negotiation

This command is supported on point-to-point interfaces only.

Through this configuration, a point-to-point interface accepts the IP address assigned by the peer end through PPP negotiation.

Related Configuration

Configuring an Interface with One or More IP Addresses

- By default, an interface is not configured with an IP address.
- The **ip address** command is used to configure an IP address for an interface.
- After an IP address is configured, the IP address can be used for communication when it passes conflict detection.
- The **ip address** *ip-address mask* **secondary** command can be used to configure multiple secondary IP addresses.

1.3.2 Broadcast Packet Processing

Working Principle

Broadcast is divided into two types. One is limited broadcast, and the IP address is 255.255.255.255. Because the broadcast is prohibited by routers, the broadcast is called local network broadcast. The other is directed broadcast. All host bits are 1s, for example, 192.168.1.255/24. The broadcast packets with these IP addresses can be forwarded.

If IP network devices forward limited broadcast packets (destination IP address is 255.255.255.255), the network may be overloaded, which severely affects network performance. This circumstance is called broadcast storm. Devices provide some approaches to confine broadcast storms within the local network and prevent continuous spread of broadcast storms. L2 network devices such as bridges and switches forward and spread broadcast storms.

The best way to avoid broadcast storm is to assign a broadcast address to each network, which is directed broadcast. This requires the IP protocol to use directed broadcast rather than limited broadcast to spread data.

For details about broadcast storms, see RFC 919 and RFC 922.

Directed broadcast packets refer to the broadcast packets destined for a subnet. For example, packets whose destination address is 172.16.16.255 are called directed broadcast packets. However, the node that generates the packets is not a member of the destination subnet.

After receiving directed broadcast packets, the devices not directly connected to the destination subnet forward the packets. After directed broadcast packets reach the devices directly connected to the subnet, the devices convert directed broadcast packets to limited broadcast packets (destination IP address is 255.255.255.255) and broadcast the packets to all hosts on the destination subnet at the link layer.

Related Configuration

Configuring an IP Broadcast Address

- By default, the IP broadcast address of an interface is 255.255.255.255.
- To define broadcast packets of other addresses, run the **ip broadcast-address** command on the interface.

Forwarding Directed Broadcast Packets

- By default, directed broadcast packets cannot be forwarded.
- On the specified interface, you can run the **ip directed-broadcast** command to enable directed broadcast packets forwarding. In this way, the interface can forward directed broadcast packets to networks that are directly connected. Broadcast packets can be transmitted within the destination subnet without affecting forwarding of other directed broadcast packets.
- On an interface, you can define an Access Control List (ACL) to transmit certain directed broadcast packets. After an ACL is defined, only directed broadcast packets that match the ACL are forwarded.

1.3.3 Sending ICMP Packets

Working Principle

ICMP Protocol Unreachable Message

A device receives non-broadcast packets destined for itself, and he packets contain the IP protocol that cannot be processed by the device. The device sends an ICMP protocol unreachable message to the source host. Besides, if the device does not know a route to forward packets, it also sends an ICMP host unreachable message.

ICMP Redirection Message

Sometimes, a route may be less than optimal, which makes a device send packets from the interface that receives packets. If a device sends packets from an interface on which it receives the packets, the device sends an ICMP redirection message to the source, informing the source that the gateway is another device on the same subnet. In this way, the source sends subsequent packets according to the optimal path.

ICMP Mask Response Message

Sometimes, a network device sends an ICMP mask request message to obtain the mask of a subnet.. The network device that receives the ICMP mask request message sends a mask response message.

Enabling Notifications of Expired TTL

- By default, notifications of expired TTL are enabled.
- You can run the **[no] ip ttl-expires enable** command to enable or disable the function.

Enabling the Device to Return a Timestamp Reply

- By default, the device returns a Timestamp Reply.
- You can run the **[no] ip icmp timestamp** command to enable or disable the function.

Related Configuration

Enabling ICMP Protocol Unreachable Message

- By default, the ICMP Protocol unreachable message function is enabled on an interface.
- You can run the [**no**] **ip unreachables** command to disable or enable the function.

Enabling ICMP Redirection Message

- By default, the ICMP redirection message function is enabled on an interface.

- You can run the [**no**] **ip redirects** command to disable or enable the function.

Enabling ICMP Mask Response Message

- By default, the ICMP mask response message function is enabled on an interface.
- You can run the [**no**] **ip mask-reply** command to disable or enable the function.

1.3.4 Limiting Transmission Rate of ICMP Error Packets

Working Principle

This function limits the transmission rate of ICMP error packets to prevent DoS attacks by using the token bucket algorithm.

If an IP packet needs to be fragmented but the Don't Fragment (DF) bit in the header is set to 1, the device sends an ICMP destination unreachable packet (code 4) to the source host. This ICMP error packet is used to discover the path MTU. When there are too many other ICMP error packets, the ICMP destination unreachable packet (code 4) may not be sent. As a result, the path MTU discovery function fails. To avoid this problem, you should limit the transmission rate of ICMP destination unreachable packets and other ICMP error packets respectively.

Related Configuration

Configuring the Transmission Rate of ICMP Destination Unreachable Packets Triggered by DF Bit in the IP Header

- The default transmission rate is 10 packets every 100 milliseconds.
- The **ip icmp error-interval DF** command can be used to configure the transmission rate.

Configuring the Transmission Rate of Other ICMP Error Packets

- The default transmission rate is 10 packets every 100 milliseconds.
- The **ip icmp error-interval** command can be used to configure the transmission rate.

1.3.5 IP MTU

Working Principle

If an IP packet exceeds the IP MTU size, the FSOS software splits the packet. For all devices in the same physical network segment, the IP MTU of interconnected interfaces must be the same. You can adjust the link MTU of interfaces on FS products. After the link MTU of interfaces is changed, the IP MTU of interfaces will be changed. The IP MTU of interfaces automatically keeps consistent with the link MTU of interfaces. However, if the IP MTU of interfaces is adjusted, the link MTU of interfaces will not be changed.

Related Configuration

Setting the IP MTU

- By default, the IP MTU of an interface is 1500.
- The **ip mtu** command can be used to set the IP packet MTU.

1.3.6 IP TTL

Working Principle

An IP packet is transmitted from the source address to the destination address through routers. After a TTL value is set, the TTL value

decreases by 1 every time when the IP packet passes a router. When the TTL value drops to zero, the router discards the packet. This prevents infinite transmission of useless packets and waste of bandwidth.

Related Configuration

Setting the IP TTL

- By default, the IP TTL of an interface is 64.
- The **ip ttl** command can be used to set the IP TTL of an interface.

1.3.7 IP Source Route

Working Principle

FS products support IP source routes. When a device receives an IP packet, it checks the options such as source route, loose source route, and record route in the IP packet header. These options are detailed in RFC 791. If the device detects that the packet enables one option, it responds; if the device detects an invalid option, it sends an ICMP parameter error message to the source and then discards the packet.

After the IP source route is enabled, the source route option is added to an IP packet to test the throughput of a specific network or help the packet bypasses the failed network. However, this may cause network attacks such as source address spoofing and IP spoofing.

Related Configuration

Configuring an IP Source Route

- By default, the IP source route function is enabled.
- The **ip source-route** command can be used to enable or disable the function.

1.4 Configuration

Configuration	Description and Command	
Configuring the IP Addresses of an Interface	(Mandatory) It is used to configure an IP address and allow the IP protocol to run on an interface.	
	ip address	Manually configures the IP address of an interface.
Configuring Broadcast Forwarding	(Optional) It is used to set an IP broadcast address and enable directed broadcast forwarding.	
	ip broadcast-address	Configures an IP broadcast address.
	ip directed-broadcast	Enables directed broadcast forwarding.
Configuring ICMP Forwarding	(Optional) It is used to enable ICMP packet forwarding.	
	ip unreachables	Enables ICMP unreachable messages and host unreachable messages.
	ip redirects	Enables ICMP redirection messages.
	ip mask-reply	Enables ICMP mask response messages.
	ip icmp timestamp	Enables the device to return a Timestamp Reply.
Configuring the Transmission	Optional.	

Configuration	Description and Command	
Rate of ICMP Error Packets	**ip icmp error-interval DF**	Configures the transmission rate of ICMP destination unreachable packets triggered by the DF bit in the IP header.
	ip icmp error-interval	Configures the transmission rate of ICMP error packets and ICMP redirection packets.
Setting the IP MTU	(Optional) It is used to configure the IP MTU on an interface.	
	ip mtu	Sets the MTU value.
Setting the IP TTL	(Optional) It is used to configure the TTL of unicast packets and broadcast packets.	
	ip ttl	Sets the TTL value.
Configuring an IP Source Route	(Optional) It is used to check the source routes.	
	ip source-route	Enables the IP source route function.

1.4.1 Configuring the IP Addresses of an Interface

Configuration Effect

Configure the IP address of an interface for communication.

Notes

- N/A

Configuration Steps

Configuring the IP Address of an Interface

- Mandatory
- Perform the configuration in L3 interface configuration mode.

Obtaining the IP Address of an Interface through PPP Negotiation

- Optional
- If a point-to-point interface is not configured with an IP address, obtain an IP address through PPP negotiation.
- Perform the configuration in L3 interface configuration mode.

Verification

Run the **show ip interface** command to check whether the configuration takes effect.

Related Commands

Manually Configuring the IP Address of an Interface

Command	**ip address** *ip-address network-mask* [**secondary**]
Parameter Description	*ip-address*: 32-bit IP address, with 8 bits for each group. The IP address is expressed in decimal and groups are separated by a full stop (.).

	network-mask: 32-bit network mask. Value 1 indicates the mask bit and 0 indicates the host bit. Every 8 bits form one group. The network mask is expressed in decimal and groups are separated by a full stop (.). **secondary**: Secondary IP address. .
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring an IP Address for an Interface

Configuration Steps	Configure IP address 192.168.23.110 255.255.255.0 on interface GigabitEthernet 0/0.
	FS#configure terminal FS(config)#interface gigabitEthernet 0/0 FS(config-if-GigabitEthernet 0/0)# no switchport FS(config-if-GigabitEthernet 0/0)#ip address 192.168.23.110 255.255.255.0
Verification	Run the **show ip interface** command to check whether the configuration takes effect.
	FS# show ip interface gigabitEthernet 0/0 GigabitEthernet 0/0 IP interface state is: UP IP interface type is: BROADCAST IP interface MTU is: 1500 IP address is: 192.168.23.110/24 (primary)

1.4.2 Configuring Broadcast Forwarding

Configuration Effect

Set the broadcast address of an interface to 0.0.0.0 and enable directed broadcast forwarding.

Notes

N/A

Configuration Steps

Configuring an IP Broadcast Address

- (Optional) Some old hosts may identify broadcast address 0.0.0.0 only. In this case, set the broadcast address of the target interface to 0.0.0.0.
- Perform the configuration in L3 interface configuration mode.

Enabling Directed Broadcast Forwarding

- (Optional) If you want to enable a host to send broadcast packets to all hosts in a domain that it is not in, enable directed broadcast forwarding.
- Perform the configuration in L3 interface configuration mode.

Verification

Run the **show running-config interface** command to check whether the configuration takes effect.

Related Commands

Configuring an IP Broadcast Address

Command	**ip broadcast-address** *ip-address*
Parameter Description	*ip-address*: Broadcast address of an IP network.
Command Mode	Interface configuration mode
Usage Guide	Generally, the destination address of IP broadcast packets is all 1s, which is expressed as 255.255.255.255. The FSOS software can generate broadcast packets of other IP addresses through definition and receive self-defined broadcast packets and the broadcast packets with address 255.255.255.255.

Allowing Forwarding of Directed Broadcast Packets

Command	**ip directed-broadcast** [*access-list-number*]
Parameter Description	*access-list-number*: Access list number, ranging from 1 to 199 and from1300 to 2699. After an ACL is defined, only directed broadcast packets that match the ACL are forwarded.
Command Mode	Interface configuration mode
Usage Guide	If the **no ip directed-broadcast** command is run on an interface, the FSOS software will discard directed broadcast packets received from the network that is directly connected.

Configuration Example

Configuration Steps	On interface gigabitEthernet 0/1, set the destination address of IP broadcast packets to 0.0.0.0 and enable directed broadcast forwarding. FS#configure terminal FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# no switchport FS(config-if-GigabitEthernet 0/1)#ip broadcast-address 0.0.0.0 FS(config-if-GigabitEthernet 0/1)#ip directed-broadcast
Verification	Run the **show ip interface** command to check whether the configuration takes effect. FS#show running-config interface gigabitEthernet 0/1 ip directed-broadcast ip broadcast-address 0.0.0.0

1.4.3 Configuring ICMP Forwarding

Configuration Effect

Enable ICMP unreachable messages, ICMP redirection messages, and mask response messages on an interface.

Notes

N/A

Configuration Steps

Enabling ICMP Unreachable Messages

- By default, ICMP unreachable messages are enabled.
- (Optional)The **no ip unreachables** command can be used to disable ICMP unreachable messages.
- Perform the configuration in L3 interface configuration mode.

Enabling ICMP Redirection Messages

- By default, ICMP redirection messages are enabled.
- (Optional)The **no ip redirects** command can be used to disable ICMP redirection messages.
- Perform the configuration in L3 interface configuration mode.

Enabling ICMP Mask Response Messages

- By default, ICMP mask response messages are enabled.
- (Optional)The **no ip mask-reply** command can be used to disable ICMP mask response messages.
- Perform the configuration in L3 interface configuration mode.

Enabling Notifications of Expired TTL

- By default, notifications of expired TTL are enabled.
- (Optional)The **[no] ip ttl-expires enable** command can be used to enable or disable the function.
- Perform the configuration in global configuration mode.

Enabling the Device to Return a Timestamp Reply

- By default, the device returns a Timestamp Reply.
- (Optional)The **no ip icmp timest** command can be used to disable the function.
- Perform the configuration in global configuration mode.

Verification

Run the **show ip interface** command to check whether the configuration takes effect.

Run the **show running-config** command to check whether notifications of expired TTL are enabled.

Run the **show running-config** command to check whether the device returns a Timestamp Reply.

Related Commands

Enabling ICMP Unreachable Messages

Command	**ip unreachables**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Enabling ICMP Redirection Messages

Command	**ip redirects**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Enabling ICMP Mask Response Messages

Command	**ip mask-reply**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Disabling Notifications of Expired TTL

Command	**no ip ttl-expires enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Disabling the Sending of a Timestamp Reply

Command	**no ip icmp timestamp**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuration Steps	Enable ICMP unreachable messages, ICMP redirection messages, and mask response messages on interface gigabitEthernet 0/1.
	FS#configure terminal FS(config)# no ip ttl-expires enable FS(config)# no ip icmp timestamp FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# no switchport FS(config-if-GigabitEthernet 0/1)# ip unreachables FS(config-if-GigabitEthernet 0/1)# ip redirects FS(config-if-GigabitEthernet 0/1)# ip mask-reply
Verification	Run the **show ip interface** command to check whether the configuration takes effect.
	FS#show running-config | include ip ttl-expires enable no ip ttl-expires enable FS#show running-config | include ip icmp timestamp no ip icmp timestamp FS#show ip interface gigabitEthernet 0/1 GigabitEthernet 0/1 ICMP mask reply is: ON Send ICMP redirect is: ON Send ICMP unreachabled is: ON

1.4.4 Configuring the Transmission Rate of ICMP Error Packets

Configuration Effect

Configure the transmission rate of ICMP error packets.

Notes

N/A

Configuration Steps

Configuring the Transmission Rate of ICMP Destination Unreachable Packets Triggered by the DF Bit in the IP Header

- Optional
- Perform the configuration in global configuration mode.

Configuring the Transmission Rate of Other ICMP Error Packets

- Optional

- Perform the configuration in global configuration mode.

Verification

Run the **show running-config** command to check whether the configuration takes effect.

Related Commands

Configuring the Transmission Rate of ICMP Destination Unreachable Packets Triggered by the DF Bit in the IP Header

Command	**ip icmp error-interval DF** *milliseconds* [*bucket-size*]
Parameter Description	*milliseconds*: Refresh cycle of a token bucket. The value range is from 0 to 2,147,483,647 and the default value is 100 milliseconds. When the value is 0, the transmission rate of ICMP error packets is not limited. *bucket-size*: Number of tokens contained in a token bucket. The value range is from 1 to 200 and the default value is 10.
Command Mode	Global configuration mode.
Usage Guide	This function limits the transmission rate of ICMP error packets to prevent DoS attacks by using the token bucket algorithm. If an IP packet needs to be fragmented but the DF bit in the header is set to 1, the device sends an ICMP destination unreachable packet (code 4) to the source host. This ICMP error packet is used to discover the path MTU. When there are too many other ICMP error packets, the ICMP destination unreachable packet (code 4) may not be sent. As a result, the path MTU discovery function fails. To avoid this problem, you should limit the transmission rate of ICMP destination unreachable packets and other ICMP error packets respectively. It is recommended to set the refresh cycle to integral multiples of 10 milliseconds. If the refresh cycle is set to a value greater than 0 and smaller than 10 milliseconds, the refresh cycle that actually takes effect is 10 milliseconds. For example, if the refresh rate is set to 1 per 5 milliseconds, the refresh rate that actually takes effect is 2 per 10 milliseconds. If the refresh cycle is not integral multiples of 10 milliseconds, the refresh cycle that actually takes effect is automatically converted to integral multiples of 10 milliseconds. For example, if the refresh rate is set to 3 per 15 milliseconds, the refresh rate that actually takes effect is 2 per 10milliseconds.

Configuring the Transmission Rate of Other ICMP Error Packets

Command	**ip icmp error-interval** *milliseconds* [*bucket-size*]
Parameter Description	*milliseconds*: Refresh cycle of a token bucket. The value range is 0to 2,147,483,647, and the default value is **100** (ms). When the value is **0**, the transmission rate of ICMP error packets is not limited. *bucket-size*: Number of tokens contained in a token bucket. The value range is 1to 200 and the default value is **10**.
Command Mode	Global configuration mode.
Usage Guide	This function limits the transmission rate of ICMP error packets to prevent DoS attacks by using the token bucket algorithm. It is recommended to set the refresh cycle to integral multiples of 10 milliseconds. If the refresh cycle is set to a value greater than 0 and smaller than 10 milliseconds, the refresh cycle that actually takes effect is 10 milliseconds. For example, if the refresh rate is set to 1 per 5 milliseconds, the refresh rate that actually takes effect is 2 per 10 milliseconds. If the refresh cycle is not integral multiples of 10 milliseconds, the refresh cycle that actually takes effect is automatically converted to integral multiples of 10 milliseconds. For example, if the refresh rate is set to 3 per 15 milliseconds, the refresh rate that actually takes effect is 2 per 10 milliseconds.

Configuration Example

Configuration Steps	Set the transmission rate of ICMP destination unreachable packets triggered the DF bit in IP header to 100 packets per second and the transmission rate of other ICMP error packets to 10 packets per second.
	FS(config)# ip icmp error-interval DF 1000 100 FS(config)# ip icmp error-interval 1000 10
Verification	Run the **show running-config** command to check whether the configuration takes effect.
	FS#show running-config | include ip icmp error-interval ip icmp error-interval 1000 10 ip icmp error-interval DF 1000 100

1.4.5 Setting the IP MTU

Configuration Effect

Adjust the IP packet MTU.

Notes

N/A

Configuration Steps

- (Optional) When the IP MTU of interconnected interfaces is different on devices in the same physical network segment, set the IP MTU to the same value.
- Perform the configuration in L3 interface configuration mode.

Verification

Run the **show ip interface** command to check whether the configuration takes effect.

Related Commands

Setting the IP MTU

Command	**ip mtu***bytes*
Parameter Description	*bytes*: IP packet MTU. The value range is from 68 to 1,500 bytes.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuration Steps	Set the IP MTU of interface gigabitEthernet 0/1 to 512 bytes.

	FS#configure terminal FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# no switchport FS(config-if-GigabitEthernet 0/1)#ip mtu 512
Verification	Run the **show ip interface** command to check whether the configuration takes effect.
	FS# show ip interface gigabitEthernet 0/1 IP interface MTU is: 512

1.4.6 Setting the IP TTL

Configuration Effect

Modify the IP TTL value of an interface.

Notes

N/A

Configuration Steps

- Optional
- Perform the configuration in L3 interface configuration mode.

Verification

Run the **show run-config** command to check whether the configuration takes effect.

Related Commands

Setting the IP TTL

Command	**ip ttl** *value*
Parameter Description	*value*: TTL value. The value range is from 0 to 255.
Command Mode	Global configuration mode.
Usage Guide	N/A

Configuration Example

Configuration Steps	• Set the TTL of unicast packets to 100.
	FS#configure terminal FS(config)#ip ttl 100
Verification	Run the **show run-config** command to check whether the configuration takes effect.

	FS#show running-config ip ttl 100

1.4.7 Configuring an IP Source Route

Configuration Effect

Enable or disable the IP source route function.

Notes

N/A

Configuration Steps

- By default, the IP source route function is enabled.
- Optional) The **no ip source-route** command can be used to disable the IP source route function.

Verification

Run the **show run-config** command to check whether the configuration takes effect.

Related Commands

Configuring an IP Source Route

Command	**ip source-route**
Parameter Description	N/A
Command Mode	Global configuration mode.
Usage Guide	N/A

Configuration Example

Configuration Steps	• Disable the IP source route function.
	FS#configure terminal FS(config)#no ip source-route
Verification	Run the **show run-config** command to check whether the configuration takes effect.
	FS#show running-config no ip source-route

1.5 Monitoring

Displaying

Description	Command
Displays the IP address of an interface.	**show ip interface** [*interface-type interface-number* | **brief**]
Displays IP packet statistics.	**show ip packet statistics** [**total** | *interface-name*]
Displays statistics on sent and received IP packets in the protocol stack.	**show ip packet queue**

2 Configuring ARP

2.1 Overview

In a local area network (LAN), each IP network device has two addresses: 1) local address. Since the local address is contained in the header of the data link layer (DLL) frame, it is a DLL address. However, it is processed by the MAC sublayer at the DLL and thereby is usually called the MAC address. MAC addresses represent IP network devices on LANs. 2) network address. Network addresses on the Internet represent IP network devices and also indicate the networks where the devices reside.

In a LAN, two IP devices can communicate with each other only after they learn the 48-bit MAC address of each other. The process of obtaining the MAC address based on the IP address is called address resolution. There are two types of address resolution protocols: 1) Address Resolution Protocol (ARP); 2) Proxy ARP. ARP and Proxy ARP are described respectively in RFC 826 and RFC 1027.

ARP is used to bind the MAC address with the IP address. When you enter an IP address, you can learn the corresponding MAC address through ARP. Once the MAC address is obtained, the IP-MAC mapping will be saved to the ARP cache of the network device. With the MAC address, the IP device can encapsulate DLL frames and send them to the LAN. By default, IP and ARP packets on the Ethernet are encapsulated in Ethernet II frames.

Protocols and Standards

- RFC 826: An Ethernet Address Resolution Protocol
- RFC 1027: Using ARP to implement transparent subnet gateways

2.2 Applications

Application	Description
LAN-based ARP	A user learns the MAC addresses of other users in the same network segment through ARP.
Proxy ARP-based Transparent Transmission	With Proxy ARP, a user can directly communicate with users in another network without knowing that it exists.

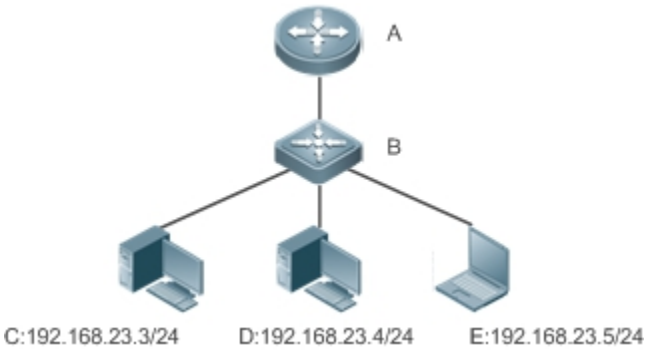
2.2.1 LAN-based ARP

Scenario

ARP is required in all IPv4 LANs.

- A user needs to learn the MAC addresses of other users through ARP to communicate with them.

Figure 2- 1



Remarks	A is a router. B is a switch. It acts as the gateway. C, D, and E are hosts.

Deployment

- Enable ARP in a LAN to implement IP-MAC mapping.

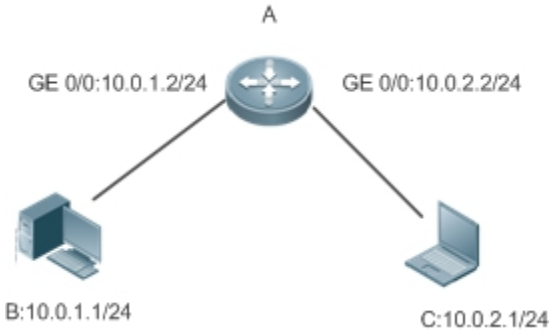
2.2.2 Proxy ARP-based Transparent Transmission

Scenario

Transparent transmission across IPv4 LANs is performed.

- Enable Proxy ARP on the router to achieve direct communication between users in different network segments.

Figure 2- 2



Remarks	A is a router connecting two LANs. B and C are hosts in different subnets. No default gateway is configured for them.

Deployment

- Enable Proxy ARP on the subnet gateway. After configuration, the gateway can act as a proxy to enable a host without any route information to obtain MAC addresses of IP users in other subnets.

2.3 Features

Overview

Feature	Description
Static ARP	Users can manually specify IP-MAC mapping to prevent the device from learning incorrect ARP entries.
ARP Attributes	Users can specify the ARP entry timeout, ARP request retransmission times and interval, and maximum number of unresolved ARP entries.
Trusted ARP	Trusted ARP is used to prevent ARP spoofing.
Gratuitous ARP	Gratuitous ARP is used to detect IP address conflicts and enable peripheral devices to update ARP entries.
Proxy ARP	A proxy replies to the ARP requests from other devices in different subnets.
Local Proxy ARP	A proxy replies to the ARP requests from other devices in the same subnet.
ARP Trustworthiness Detection	Neighbor Unreachable Detection (NUD) is used to ensure that correct ARP entries are learned.

ARP-based IP Guard	You can set the number of IP packets for triggering ARP drop to prevent a large number of unknown unicast packets from being sent to the CPU.

2.3.1 Static ARP

Static ARP entries can be configured manually or assigned by the authentication server. The manually configured ones prevail. Static ARP can prevent the device from learning incorrect ARP entries.

Working Principle

If static ARP entries are configured, the device does not actively update ARP entries and these ARP entries permanently exist.

When the device forwards Layer-3 packets, the static MAC address is encapsulated in the Ethernet header as the destination MAC address.

Related Configuration

Enabling Static ARP

Run the **arp** *ip-address mac-address type* command in global configuration mode to configure static ARP entries. By default, no static ARP entry is configured. ARP encapsulation supports only the Ethernet II type, which is represented by ARPA.

2.3.2 ARP Attributes

Users can specify the ARP timeout, ARP request retransmission interval and times, maximum number of unresolved ARP entries, maximum number of ARP entries on an interface, and maximum number of ARP entries on a board.

Working Principle

ARP Timeout

The ARP timeout only applies to the dynamically learned IP-MAC mapping. When the ARP entry timeout expires, the device sends a unicast ARP request packet to detect whether the peer end is online. If it receives an ARP reply from the peer end, it does not delete this ARP entry. Otherwise, the device deletes this ARP entry.

When the ARP timeout is set to a smaller value, the mapping table stored in the ARP cache is more accurate but ARP consumes more network bandwidth.

ARP Request Retransmission Interval and Times

The device consecutively sends ARP requests to resolve an IP address to a MAC address. The shorter the retransmission interval is, the faster the resolution is. The more times the ARP request is retransmitted, the more likely the resolution will succeed and the more bandwidth ARP will consume.

Maximum Number of Unresolved ARP Entries

In a LAN, ARP attacks and scanning may cause a large number of unresolved ARP entries generated on the gateway. As a result, the gateway fails to learn the MAC addresses of the users. To prevent such attacks, users can configure the maximum number of unresolved ARP entries.

Maximum Number of ARP Entries on an Interface

Configure the maximum number of ARP entries on a specified interface to prevent ARP entry resource waste.

Related Configuration

Configuring the ARP Timeout

Run the **arp timeout** *seconds* command in interface configuration mode to configure the ARP timeout. The default timeout is 3,600 seconds. You can change it based on actual situations.

Configuring the ARP Request Retransmission Interval and Times

- Run the **arp retry interval** *seconds* command in global configuration mode to configure the ARP request retransmission interval. The default interval is 1 second. You can change it based on actual situations.
- Run the **arp retry times** *number* command in global configuration mode to configure the ARP request retransmission times. The default number of retransmission times is 5. You can change it based on actual situations.

Configuring the Maximum Number of Unresolved ARP Entries

Run the **arp unresolve** *number* command in global configuration mode to configure the maximum number of unresolved ARP entries. The default value is the maximum number of ARP entries supported by the device. You can change it based on actual situations.

Configuring the Maximum Number of ARP Entries on an Interface

Run the **arp cache interface-limit** *limit* command in interface configuration mode to configure the maximum number of ARP entries learned on an interface. The default number is 0. You can change it based on actual situations. This command also applies to static ARP entries.

2.3.3 Trusted ARP

Working Principle

As a type of special ARP entries, trusted ARP entries are added to the ARP table to prevent ARP spoofing. Trusted ARP entries have characteristics of both static and dynamic ARP entries, with a priority higher than that of dynamic ARP entries and lower than that of static ARP entries. Trusted ARP has an aging mechanism similar to that of dynamic ARP. When an ARP entry ages, the device actively sends an ARP request packet to detect whether the corresponding user exists. If the user sends a reply, the device regards the user active and updates the ARP timeout. Otherwise, the device deletes the ARP entry. Trusted ARP has characteristics of static ARP, that is, the device does not learn ARP packets to update the MAC address and interface ID in the ARP entry.

When a user goes online on a GSN client, the authentication server obtains the user's reliable IP-MAC mapping through the access switch, and adds trusted ARP entries to the user's gateway. This process is transparent to the network administrator and does not affect the administrator’s work on network management.

Since trusted ARP entries come from authentic sources and will not be updated, they can efficiently prevent ARP spoofing targeted at the gateway.

Related Configuration

Enabling Trusted ARP

- Run the **service trustedarp** command in global configuration mode to enable trusted ARP. This function is disabled by default.
- Run the **arp trusted user-vlan** *vid1* **translated-vlan** *vid2* command in global configuration mode to implement VLAN redirection. This function is disabled by default. If the VLAN pushed by the server differs from the VLAN in the trusted ARP entry, users need to enable VLAN redirection.

- Run the **arp trusted aging** command in global configuration mode to enable ARP aging. Trusted ARP entries are not aged by default.
- Run the **arp trusted number** command in global configuration mode to configure the capacity of trusted ARP entries. The default value is half of the total capacity of ARP entries. You can change it based on actual situations.

2.3.4 Gratuitous ARP

Working Principle

Gratuitous ARP packets are a special type of ARP packets. In a gratuitous ARP packet, the source and destination IP addresses are the IP address of the local device. Gratuitous ARP packets have two purposes:

1. IP address conflict detection. If the device receives a gratuitous packet and finds the IP address in the packet the same as its own IP address, it sends an ARP reply to notify the peer end of the IP address conflict.
2. ARP update. When the MAC address of an interface changes, the device sends a gratuitous ARP packet to notify other devices to update ARP entries.

The device can learn gratuitous ARP packets. After receiving a gratuitous ARP packet, the device checks whether the corresponding dynamic ARP entry exists. If yes, the device updates the ARP entry based on the information carried in the gratuitous ARP packet.

Related Configuration

Enabling Gratuitous ARP

Run the **arp gratuitous-send interval** *seconds* [*number*] command in interface configuration mode to enable gratuitous ARP. This function is disabled on interfaces by default. Generally you need to enable this function on the gateway interface to periodically update the MAC address of the gateway on the downlink devices, which prevents others from faking the gateway.

2.3.5 Proxy ARP

Working Principle

The device enabled with Proxy ARP can help a host without any route information to obtain MAC addresses of IP users in other subnets. For example, if the device receiving an ARP request finds the source IP address in a different network segment from the destination IP address and knows the route to the destination address, the device sends an ARP reply containing its own Ethernet MAC address. This is how Proxy ARP works.

Related Configuration

Enabling Proxy ARP

- Run the **ip proxy-arp** command in interface configuration mode to enable Proxy ARP.
- This function is enabled on routers while disabled on switches by default.

2.3.6 Local Proxy ARP

Working Principle

Local Proxy ARP means that a device acts as a proxy in the local VLAN (common VLAN or sub VLAN).

After local Proxy ARP is enabled, the device can help users to obtain the MAC addresses of other users in the same subnet. For example,

when port protection is enabled on the device, users connected to different ports are isolated at Layer 2. After local Proxy ARP is enabled, the device receiving an ARP request acts as a proxy to send an ARP reply containing its own Ethernet MAC address. In this case, different users communicate with each other through Layer-3 routes. This is how local Proxy ARP works.

Related Configuration

Enabling Local Proxy ARP

- Run the **local-proxy-arp** command in interface configuration mode to enable local Proxy ARP.
- This function is disabled by default.
- This command is supported only on switch virtual interfaces (SVIs).

2.3.7 ARP Trustworthiness Detection

Working Principle

The **arp trust-monitor enable** command is used to enable anti-ARP spoofing to prevent excessive useless ARP entries from occupying device resources. After ARP trustworthiness detection is enabled on a Layer-3 interface, the device receives ARP request packets from this interface:

1. If the corresponding entry does not exist, the device creates a dynamic ARP entry and performs NUD after 1 to 5 seconds. That is, the device begins to age the newly learned ARP entry and sends a unicast ARP request. If the device receives an ARP update packet from the peer end within the aging time, it stores the entry. If not, it deletes the entry.
2. If the corresponding ARP entry exists, NUD is not performed.
3. If the MAC address in the existing dynamic ARP entry is updated, the device also performs NUD.

Since this function adds a strict confirmation procedure in the ARP learning process, it affects the efficiency of ARP learning.

After this function is disabled, NUD is not required for learning and updating ARP entries.

Related Configuration

Enabling ARP Trustworthiness Detection

Run the **arp trust-monitor enable** command in interface configuration mode to enable ARP trustworthiness detection. This function is disabled by default.

2.3.8 ARP-based IP Guard

Working Principle

When receiving unresolved IP packets, the switch cannot forward them through the hardware and thereby need to send them to the CPU for address resolution. If a large number of such packets are sent to the CPU, the CPU will be congested, affecting other services on the switch.

After ARP-based IP guard is enabled, the switch receiving ARP request packets counts the number of packets in which the destination IP address hits this ARP entry. If this number is equal to the configured number, the switch sets a drop entry in the hardware so that the hardware will not send the packets with this destination IP address to the CPU. After the address resolution is complete, the switch continues to forward the packets with this destination IP address.

Related Configuration

Enabling ARP-based IP Guard

- Run the **arp anti-ip-attack** command in global configuration mode to configure the number of IP packets for triggering ARP drop.
- By default, the switch discards the corresponding ARP entry after it receives three unknown unicast packets containing the same destination IP address.

2.4 Configuration

Configuration	Description and Command	
Enabling Static ARP	(Optional) It is used to enable static IP-MAC binding.	
	arp	Enables static ARP.
Configuring ARP Attributes	(Optional) It is used to specify the ARP timeout, ARP request retransmission interval and times, maximum number of unresolved ARP entries, and maximum number of ARP entries on an interface.	
	arp timeout	Configures the ARP timeout.
	arp retry interval	Configures the ARP request retransmission interval.
	arp unresolve	Configures the maximum number of unresolved ARP entries.
	arp cache interface-limit	Configures the maximum number of ARP entries on an interface.
Enabling Trusted ARP	(Optional) It is used to enable anti-ARP spoofing.	
	service trustedarp	Enables trusted ARP.
	arp trusted user-vlan	Enables VLAN redirection when a trusted ARP entry is added.
	arp trusted aging	Enables trusted ARP aging.
	arp trusted	Configures the capacity of trusted ARP entries.
Enabling Gratuitous ARP	(Optional) It is used to detect IP address conflicts and enables peripheral devices to update ARP entries.	
	arp gratuitous-send interval	Enables gratuitous ARP.
Enabling Proxy ARP	(Optional) It is used to act as a proxy to reply to ARP requests from the devices in different subnets.	
	ip proxy-arp	Enables Proxy ARP.
Enabling Local Proxy ARP	(Optional) It is used to act as a proxy to reply to ARP requests from other devices in the same subnet.	
	local-proxy-arp	Enables local Proxy ARP.
Enabling ARP Trustworthiness Detection	(Optional) It is used to unicast ARP request packets to ensure that correct ARP entries are learned.	
	arp trusted-monitor enable	Enables ARP trustworthiness detection.
Enabling ARP-based IP Guard	(Optional) It is used to prevent a large number of IP packets from being sent to the CPU.	

Configuration	Description and Command	
	arp anti-ip-attack	Configures the number of IP packets for triggering ARP drop.

2.4.1 Enabling Static ARP

Configuration Effect

Users can manually specify IP-MAC mapping to prevent the device from learning incorrect ARP entries.

Notes

After a static ARP entry is configured, the Layer-3 switch learns the physical port corresponding to the MAC address in the static ARP entry before it performs Layer-3 routing.

Configuration Steps

Configuring Static ARP Entries

- Optional.
- You can configure a static ARP entry to bind the IP address of the uplink device with its MAC address to prevent MAC change caused by ARP attacks.
- Configure static ARP entries in global configuration mode.

Verification

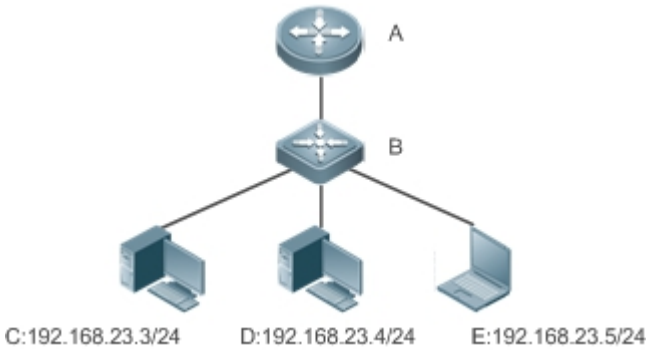
Run the **show running-config** command to check whether the configuration takes effect. Or run the **show arp static** command to check whether a static ARP cache table is created.

Related Commands

Configuring Static ARP Entries

Command	**arp** *ip-address mac-address type*
Parameter Description	*ip-address*: Indicates the IP address mapped to a MAC address, which is in four-part dotted-decimal format. *mac-address*: Indicates the DLL address, consisting of 48 bits. *type*: Indicates the ARP encapsulation type. For an Ethernet interface, the keyword is **arpa**.
Command Mode	Global configuration mode
Usage Guide	The FSOS queries a 48-bit MAC address based on a 32-bit IP address in the ARP cache table. Since most hosts support dynamic ARP resolution, usually the static ARP mapping are not configured. Use the **clear arp-cache** command to delete the dynamic ARP entries.

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	Configure a static ARP entry on B to statically bind the IP address of A with the MAC address. FS(config)#arp 192.168.23.1 00D0.F822.334B arpa
Verification	Run the **show arp static** command to display the static ARP entry. FS(config)#show arp static Protocol Address Age(min) Hardware Type Interface Internet 192.168.23.1 <static> 00D0.F822.334B arpa 1 static arp entries exist.

Common Errors

- The MAC address in static ARP is incorrect.

2.4.2 Configuring ARP Attributes

Configuration Effect

Users can specify the ARP timeout, ARP request retransmission interval and times, maximum number of unresolved ARP entries, maximum number of ARP entries on an interface, and maximum number of ARP entries on a board.

Configuration Steps

Configuring the ARP Timeout

- Optional.
- In a LAN, if a user goes online/offline frequently, it is recommended to set the ARP timeout small to delete invalid ARP entries as soon as possible.
- Configure the ARP timeout in interface configuration mode.

Configuring the ARP Request Retransmission Interval and Times

- Optional.
- If the network resources are insufficient, it is recommended to set the ARP request retransmission interval great and the retransmission times small to reduce the consumption of network bandwidths.

- Configure the ARP request retransmission interval and times in global configuration mode.

Configuring the Maximum Number of Unresolved ARP Entries

- Optional.
- If the network resources are insufficient, it is recommended to set the maximum number of unresolved ARP entries small to reduce the consumption of network bandwidths.
- Configure the maximum number of unresolved ARP entries in global configuration mode.

Configuring the Maximum Number of ARP Entries on an Interface

- Optional.
- Configure the maximum number of ARP entries on an interface in interface configuration mode.

Verification

Run the **show arp timeout** command to display the timeouts of all interfaces.

Run the **show running-config** command to display the ARP request retransmission interval and times, maximum number of unresolved ARP entries, maximum number of ARP entries on an interface, and maximum number of ARP entries on a board.

Related Commands

Configuring the ARP Timeout

Command	**arp timeout** *seconds*
Parameter Description	*seconds*: Indicates the timeout in seconds, ranging from 0 to 2,147,483. The default value is 3,600.
Command Mode	Interface configuration mode
Usage Guide	The ARP timeout only applies to the dynamically learned IP-MAC mapping. When the ARP timeout is set to a smaller value, the mapping table stored in the ARP cache is more accurate but ARP consumes more network bandwidth. Unless otherwise specified, do not configure the ARP timeout.

Configuring the ARP Request Retransmission Interval and Times

Command	**arp retry interval** *seconds*
Parameter Description	*seconds*: Indicates the ARP request retransmission interval in seconds, ranging from 1 to 3,600. The default value is 1.
Command Mode	Global configuration mode
Usage Guide	If a device frequently sends ARP requests, affecting network performance, you can set the ARP request retransmission interval longer. Ensure that this interval does not exceed the ARP timeout.

Configuring the Maximum Number of Unresolved ARP Entries

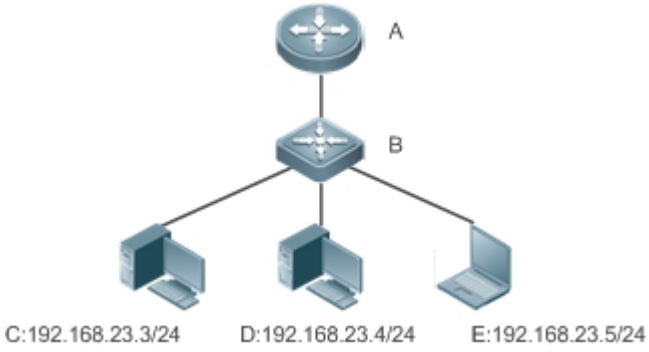
Command	**arp unresolve** *number*
Parameter Description	*number*: Indicates the maximum number of unresolved ARP entries, ranging from 1 to 8,192. The default value is 8,192.
Command	Global configuration mode

Mode	
Usage Guide	If a large number of unresolved entries exist in the ARP cache table and remain in the table after a while, it is recommended to use this command to limit the number of unresolved ARP entries.

Configuring the Maximum Number of ARP Entries on an Interface

Command	**arp cache interface-limit** *limit*
Parameter Description	*limit*: Indicates the maximum number of ARP entries that can be learned on an interface, including configured ARP entries and dynamically learned ARP entries. The value ranges from 0 to the ARP entry capacity supported by the device. 0 indicates no limit on this number.
Command Mode	Interface configuration mode
Usage Guide	Limiting the number of ARP entries on an interface can prevent malicious ARP attacks from generating excessive ARP entries on the device and occupying entry resources. The configured value must be equal to or greater than the number of the ARP entries learned by the interface. Otherwise, the configuration does not take effect. The configuration is subject to the ARP entry capacity supported by the device.

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	● Set the ARP timeout to 60 seconds on port GigabitEthernet 0/1. ● Set the maximum number of learned ARP entries to 300 on port GigabitEthernet 0/1. ● Set the ARP request retransmission interval to 3 seconds. ● Set the ARP request retransmission times to 4. ● Set the maximum number of unresolved ARP entries to 4,096. ● Set the maximum number of learned ARP entries to 1,000 on Sub Slot 2 of Slot 1.
	FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#arp timeout 60 FS(config-if-GigabitEthernet 0/1)#arp cache interface-limit 300 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#arp retry interval 3 FS(config)#arp retry times 4

	FS(config)#arp unresolve 4096
Verification	● Run the **show arp timeout** command to display the timeout of the interface. ● Run the **show running-config** command to display the ARP request retransmission interval and times, maximum number of unresolved ARP entries, and maximum number of ARP entries on the interface.
	FS#show arp timeout Interface arp timeout(sec) ---------- ---------- GigabitEthernet 0/1 60 GigabitEthernet 0/2 3600 GigabitEthernet 0/4 3600 GigabitEthernet 0/5 3600 GigabitEthernet 0/7 3600 VLAN 100 3600 VLAN 111 3600 FS(config)# show running-config arp unresolve 4096 arp retry times 4 arp retry interval 3 ! interface GigabitEthernet 0/1 arp cache interface-limit 300

2.4.3 Enabling Trusted ARP

Configuration Effect

The gateway is protected from ARP spoofing.

Notes

Trusted ARP is supported only on switches.

Configuration Steps

- To deploy a GSN solution, enable trusted ARP.
- To deploy a GSN solution, enable trusted ARP.
- Enable trusted ARP in global configuration mode.

Verification

Run the **show arp trusted** command to display trusted ARP entries.

Run the **show running** command to check whether the configuration takes effect.

Related Commands

Enabling Trusted ARP

Command	**service trustedarp**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Trusted ARP is an anti-ARP spoofing function. As a part of the GSN solution, trusted ARP needs to be used with the GSN solution.

Enabling VLAN Redirection When a Trusted ARP Entry Is Added

Command	**arp trusted user-vlan** *vid1* **translated-vlan** *vid2*
Parameter Description	*vid1*: Indicates the VLAN ID configured on the server. *vid2*: Indicates the ID of the VLAN redirected.
Command Mode	Global configuration mode
Usage Guide	This command takes effect only after trusted ARP is enabled. Configure this command only when the VLAN pushed by the server differs from the VLAN in the trusted ARP entry.

Displaying Trusted ARP Entries

Command	**show arp trusted** [*ip* [*mask*]]
Parameter Description	*ip*: Indicates the IP address. The ARP entry of the specified IP address is displayed. If keyword **trusted** is specified, only the trusted ARP entries are displayed. Otherwise, the non-trusted ARP entries are displayed. *mask*: ARP entries within the IP subnet are displayed. If keyword **trusted** is specified, only the trusted ARP entries are displayed. Otherwise, the non-trusted ARP entries are displayed.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Deleting Trusted ARP Entries

Command	**clear arp trusted** [*ip* [*mask*]]
Parameter Description	*ip*: Indicates the IP address. The ARP entry of the specified IP address is displayed. If keyword **trusted** is specified, only the trusted ARP entries are displayed. Otherwise, the non-trusted ARP entries are displayed. *mask*: ARP entries within the IP subnet are displayed. If keyword **trusted** is specified, only the trusted ARP entries are displayed. Otherwise, the non-trusted ARP entries are displayed.
Command Mode	Privileged EXEC mode
Usage Guide	After you run the **clear arp trusted** command to delete all trusted ARP entries on the switch, users may fail to access the network. It is recommended to use the **clear arp trusted** *ip* command to delete a specified trusted ARP entry.

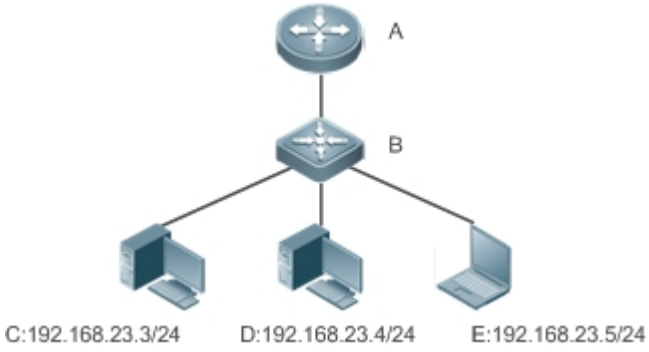
Enabling Trusted ARP Aging

Command	**arp trusted aging**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After you configure this command, trusted ARP entries begin to age, with the aging time the same as the dynamic ARP aging time. You can run the **arp timeout** command in interface configuration mode to configure the aging time.

Adjusting the Capacity of Trusted ARP Entries

Command	**arp trusted** *number*
Parameter Description	*number*: The minimum value is 10. The maximum number is the capacity supported by the device minus 1,024. By default, the maximum number of trusted ARP entries is half of the total capacity of ARP entries.
Command Mode	Global configuration mode
Usage Guide	To make this command take effect, enable trusted ARP first. Trusted ARP entries and other entries share the memory. If trusted ARP entries occupy much space, dynamic ARP entries may not have sufficient space. Set the number of ARP entries based on the actual requirement. Do not set it to an excessively large value.

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	● Enable trusted ARP. ● Enable VLAN redirection. ● Enable trusted ARP aging. ● Set the maximum number of trusted ARP entries to 1,024.
	FS(config)#service trustedarp FS(config)#arp trusted user-vlan 2-9 translated-vlan 10 FS(config)#arp trusted aging FS(config)#arp trusted 1024
Verification	● Run the **show running-config** command to check whether the configurations take effect.

	FS(config)# show running-config service trustedarp arp trusted user-vlan 2-9 translated-vlan 10 arp trusted aging arp trusted 1024

Common Errors

- Trusted ARP is disabled, causing failure to assign ARP entries.

2.4.4 Enabling Gratuitous ARP

Configuration Effect

The interface periodically sends gratuitous ARP packets.

Configuration Steps

- Optional.
- When a switch acts as the gateway, enable gratuitous ARP on an interface to prevent other users from learning incorrect gateway MAC address in case of ARP spoofing.
- Enable gratuitous ARP in interface configuration mode.

Verification

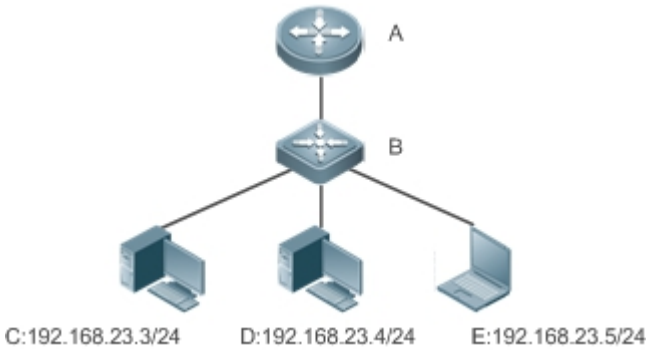
Run the **show running-config interface** [*name*] command to check whether the configuration is successful.

Related Commands

Enabling Gratuitous ARP

Command	**arp gratuitous-send interval** *seconds* [*number*]
Parameter Description	*seconds*: Indicates the interval for sending a gratuitous ARP request. The unit is second. The value ranges from 1 to 3,600. *number*: Indicates the number of gratuitous ARP requests that are sent. The default value is 1. The value ranges from 1 to 100.
Command Mode	Interface configuration mode
Usage Guide	If a network interface of a device acts as the gateway for downstream devices but a downstream device pretends to be the gateway, enable gratuitous ARP on the interface to advertise itself as the real gateway.

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	Configure the GigabitEthernet 0/0 interface to send a gratuitous ARP packet every 5 seconds.
	FS(config-if-GigabitEthernet 0/0)#arp gratuitous-send interval 5
Verification	Run the **show running-config interface** command to check whether the configuration takes effect.
	FS#sh running-config interface gigabitEthernet 0/0 Building configuration... Current configuration : 127 bytes ! interface GigabitEthernet 0/0 duplex auto speed auto ip address 30.1.1.1 255.255.255.0 arp gratuitous-send interval 5

2.4.5 Enabling Proxy ARP

Configuration Effect

The device acts as a proxy to reply to ARP request packets from other users.

Notes

By default, Proxy ARP is disabled on Layer-3 switches while enabled on routers.

Configuration Steps

- Optional.
- If a user without any route information needs to obtain the MAC addresses of the IP users in other subnets, enable Proxy ARP on the device so that the device can act as a proxy to send ARP replies.

- Enable Proxy ARP in interface configuration mode.

Verification

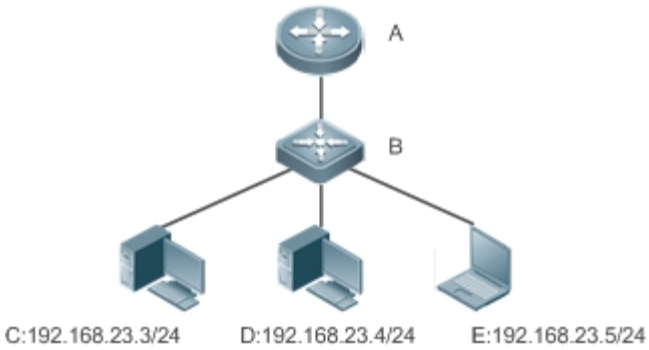
Run the **show run interface** [*name*] command to check whether the configuration takes effect.

Related Commands

Enabling Proxy ARP

Command	**ip proxy-arp**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	Enable Proxy ARP on port GigabitEthernet 0/0 .
	FS(config-if-GigabitEthernet 0/0)#ip proxy-arp
Verification	Run the **show ip interface** command to check whether the configuration takes effect.
	FS#show ip interface gigabitEthernet 0/0 GigabitEthernet 0/0 IP interface state is: DOWN IP interface type is: BROADCAST IP interface MTU is: 1500 IP address is: No address configured IP address negotiate is: OFF

	Forward direct-broadcast is: OFF ICMP mask reply is: ON Send ICMP redirect is: ON Send ICMP unreachable is: ON DHCP relay is: OFF Fast switch is: ON Help address is: 0.0.0.0 Proxy ARP is: ON ARP packet input number: 0 Request packet : 0 Reply packet : 0 Unknown packet : 0 TTL invalid packet number: 0 ICMP packet input number: 0 Echo request : 0 Echo reply : 0 Unreachable : 0 Source quench : 0 Routing redirect : 0

2.4.6 Enabling Local Proxy ARP

Configuration Effect

The device acts as a proxy to reply to ARP request packets from other users in the same subnet.

Notes

Local Proxy ARP is supported only on SVIs.

Configuration Steps

- Optional.
- If a user enabled with port protection needs to communicate with users in the VLAN, enable local Proxy ARP on the device.
- Enable local Proxy ARP in interface configuration mode.

Verification

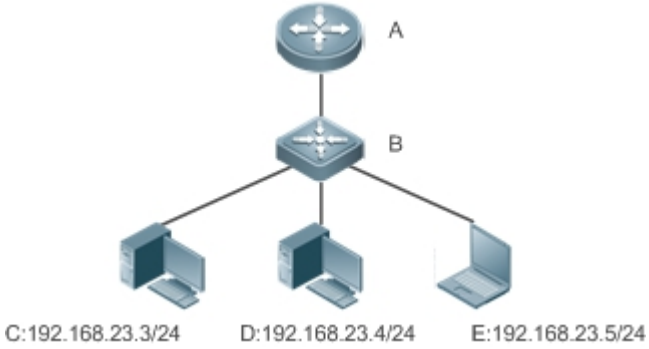
Run the **show run interface** [*name*] command to check whether the configuration takes effect.

Related Commands

Enabling Local Proxy ARP

Command	**local-proxy-arp**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	Enable local Proxy ARP on the VLAN 1 interface.
	FS(config-if-VLAN 1)#local-proxy-arp
Verification	Run the **show ip interface** command to check whether the configuration takes effect.

	FS#show running-config interface vlan 1
	
	Building configuration...
	Current configuration : 53 bytes
	
	interface VLAN 1
	  ip address 192.168.1.2 255.255.255.0
	  local-proxy-arp

2.4.7 Enabling ARP Trustworthiness Detection

Configuration Effect

Enable ARP trustworthiness detection. If the device receiving an ARP request packet fails to find the corresponding entry, it performs NUD. If the MAC address in the existing dynamic ARP entry is updated, the device immediately performs NUD to prevent ARP attacks.

Notes

Since this function adds a strict confirmation procedure in the ARP learning process, it affects the efficiency of ARP learning.

Configuration Steps

- Optional.
- If there is a need for learning ARP entries, enable ARP trustworthiness detection on the device. If the device receiving an ARP request packet fails to find the corresponding entry, it needs to send a unicast ARP request packet to check whether the peer end exists. If yes, the device learns the ARP entry. If not, the device does not learn the ARP entry. If the MAC address in the ARP entry changes, the device will immediately perform NUD to prevent ARP spoofing.
- Enable ARP trustworthiness detection in interface configuration mode.

Verification

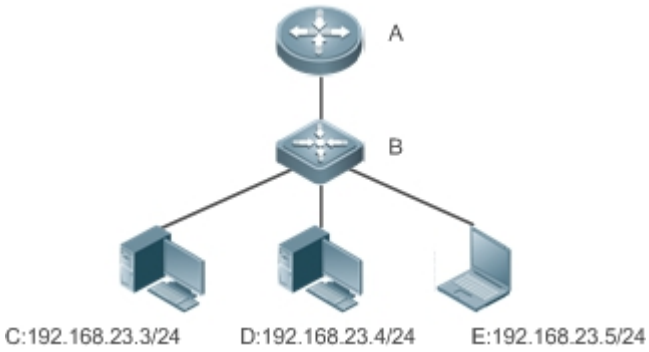
Run the **show running-config interface** [*name*] command to check whether the configuration take effect

Related Commands

Enabling ARP Trustworthiness Detection

Command	**arp trust-monitor enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Enable this function. If the corresponding ARP entry exists and the MAC address is not updated, the device does not perform NUD. Enable this function. If the MAC address of the existing dynamic ARP entry is updated, the device immediately performs NUD. After this function is disabled, the device does not perform NUD for learning or updating ARP entries.

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users

Configuration Steps	Enable ARP trustworthiness detection on port GigabitEthernet 0/0.
	FS(config-if-GigabitEthernet 0/0)#arp trust-monitor enable
Verification	Run the **show running-config interface** command to check whether the configuration takes effect.
	FS#show running-config interface gigabitEthernet 0/0 Building configuration... Current configuration : 184 bytes ! interface GigabitEthernet 0/0 duplex auto speed auto ip address 30.1.1.1 255.255.255.0 arp trust-monitor enable

2.4.8 Enabling ARP-based IP Guard

Configuration Effect

When the CPU receives the specified number of packets in which the destination IP address hits the ARP entry, all packets with this destination IP address will not be sent to the CPU afterwards.

Notes

ARP-based IP guard is supported on switches.

Configuration Steps

- Optional.
- By default, when three unknown unicast packets are sent to the switch CPU, the drop entry is set. Users can run this command to adjust the number of packets for triggering ARP drop based on the network environment. Users can also disable this function.
- Configure ARP-based IP guard in global configuration mode.

Verification

Run the **show run** command to check whether the configuration takes effect.

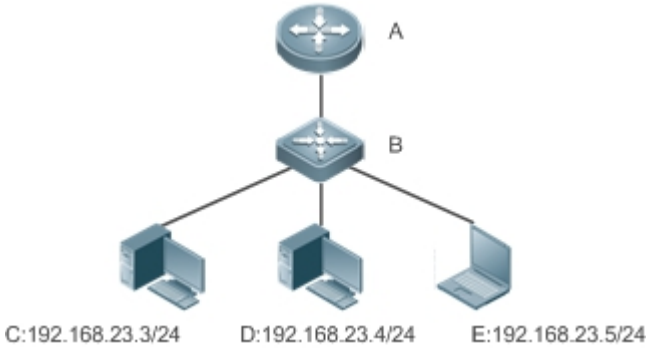
Related Commands

Enabling ARP-based IP Guard

Command	**arp anti-ip-attack** *num*
Parameter Description	*num*: Indicates the number of IP packets for triggering ARP drop. The value ranges from 0 to 100. 0 indicates that ARP-based IP guard is disabled. The default value is 3.
Command	Global configuration mode

Mode	
Usage Guide	If hardware resources are sufficient, run the **arp anti-ip-attack** *num* command to set the number of IP packets for triggering ARP drop to a small value. If hardware resources are insufficient, run the **arp anti-ip-attack** *num* command to set the number of IP packets for triggering ARP drop to a large value, or disable this function.

Configuration Example

Scenario	
Remakrs	A: Router B: Switch serving as a gateway C, D and E: Users
Configuration Steps	Enable ARP-based IP guard on B. FS(config)#arp anti-ip-attack 10
Verification	Run the **show running-config** command to check whether the configuration takes effect.
	FS#show running-config Building configuration... Current configuration : 53 bytes arp anti-ip-attack 10

2.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears dynamic ARP entries. In gateway authentication mode, dynamic ARP entries in authentication VLANs are not cleared.	**clear arp-cache**

Displaying

Description	Command
Displays the ARP table in detail.	**show arp** [**detail**] [*interface-type interface-number* [*ip* [*mask*] | *mac-address* | **static** | **complete** | **incomplete**]]
Displays the ARP table.	**show ip arp**
Displays the trusted ARP table.	**show arp** [**detail**] **trusted** [*ip* [*mask*]]
Displays the ARP entry counter.	**show arp counter**
Displays the timeout of dynamic ARP entries.	**show arp timeout**

Debugging

System resources are occupied when debugging information is output. Therefore, disable the debugging switch immediately after use.

Description	Command
Debugs ARP packet sending and receiving.	**debug arp**
Debugs the creation and deletion of ARP entries.	**debug arp event**

3 Configuring IPv6

3.1 Overview

As the Internet develops rapidly and IPv4 address space is becoming exhausted, IPv4 limitations become more and more obvious. At present, many researches and practices on Internet Protocol Next Generation (IPng) have been conducted. The IPng working group of the Internet Engineering Task Force (IETF) has formulated an IPng protocol named IP Version 6 (IPv6), which is described in RFC 2460.

Main Features

Larger Address Space

Compared with 32 bits in an IPv4 address, the length of an IPv6 address is extended to 128 bits. Therefore, the address space has approximately 2^{128} addresses. IPv6 adopts a hierarchical address allocation mode to support address allocation of multiple subnets from the Internet core network to intranet subnet.

Simpler Packet Header Format

Since the design principle of the IPv6 packet header is to minimize the overhead of the packet header, some non-key fields and optional fields are removed from the packet header to the extended packet header. Therefore, although the length of an IPv6 address is four times of that of an IPv4 address, the IPv6 packet header is only two times of the IPv4 packet header. The IPv6 packet header makes device forwarding more efficient. For example, with no checksum in the IPv6 packet header, the IPv6 device does not need to process fragments (fragmentation is completed by the initiator).

Efficient Hierarchical Addressing and Routing Structure

IPv6 uses a convergence mechanism and defines a flexible hierarchical addressing and routing structure. Multiple networks at the same layer are represented as a uniform network prefix on the upstream device, greatly reducing routing entries maintained by the device and routing and storage overheads of the device.

Easy Management: Plug and Play (PnP)

IPv6 provides automatic discovery and auto-configuration functions to simplify management and maintenance of network nodes. For example, Neighbor Discovery (ND), MTU Discovery, Router Advertisement (RA), Router Solicitation (RS), and auto-configuration technologies provide related services for PnP. Particularly, IPv6 offers two types of auto-configuration: stateful auto-configuration and stateless auto-configuration. In IPv4, Dynamic Host Configuration Protocol (DHCP) realizes auto-configuration of the host IP address and related parameters. IPv6 inherits this auto-configuration service from IPv4 and called it stateful auto-configuration (see DHCPv6). Besides, IPv6 also offers the stateless auto-configuration service.During stateless auto-configuration, a host automatically obtains the local address of the link, address prefix of the local device, and other related configurations.

Security

As an optional extension protocol of IPv4, Internet Protocol Security (IPSec) is a part of IPv6 to provide security for IPv6 packets. At present, IPv6 provides two mechanisms: Authentication Header (AH) and Encapsulated Security Payload (ESP). AH provides data integrity and authenticates IP packet sources to ensure that the packets originate from the nodes identified by the source addresses. ESP provides data encryption to realize end-to-end encryption.

Better QoS Support

A new field in the IPv6 packet header defines how to identify and process data streams. The Flow Label field in the IPv6 packet header is used to authenticate a data flow. Using this field, IPv6 allows users to propose requirements on the communication quality. , A device can

identify all packets belonging to a specific data stream based on this field and process these packets according to user requirements.

New Protocol for Neighboring Node Interaction

IPv6 Neighbor Discovery Protocol (NDP) uses a series of Internet Control Message Protocol Version 6 (ICMPv6) packets to implement interactive management of neighboring nodes (nodes on the same link). IPv6 uses NDP packets and efficient multicast/unicast ND packets instead of broadcast-based Address Resolution Protocol (ARP) and Control Message Protocol Version 4 (ICMPv4) router discovery packets.

Extensibility

With strong extensibility, IPv6 features can be added to the extended packet header following the IPv6 packet header. Unlike IPv4, the IPv6 packet header can support at most 40 bytes of options. For an IPv6 packet, the length of the extended packet header is restricted only by the maximum number of bytes in the packet.

Protocols and Standards

- RFC 4291 - IP Version 6 Addressing Architecture
- RFC 2460 - Internet Protocol, Version 6 (IPv6) Specification
- RFC 4443 - Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification
- RFC 4861 - Neighbor Discovery for IP version 6 (IPv6)
- RFC 4862 - IPv6 Stateless Address Auto-configuration
- RFC 5059 - Deprecation of Type 0 Routing Headers in IPv6

3.2 Applications

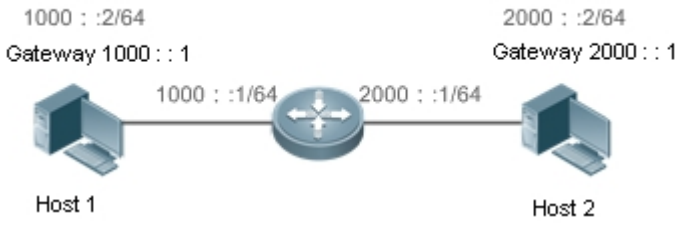
Application	Description
Communication Based on IPv6 Addresses	Two PCs communicate with each other using IPv6 addresses.

3.2.1 Communication Based on IPv6 Addresses

Scenario

As shown in Figure 3- 1, Host 1 and Host 2 communicate with each other using IPv6 addresses.

Figure 3- 1



Deployment

Hosts can use the stateless address auto-configuration or DHCPv6 address assignment mode. After addresses are configured, hosts can

communicate with each other using IPv6 addresses.

3.3 Features

Overview

Feature	Description
IPv6 Address Format	The IPv6 address format makes IPv6 have a larger address space and flexible representation approach.
IPv6 Address Type	IPv6 identifies network applications based on addresses.
IPv6 Packet Header Format	IPv6 simplifies the fixed and extended packet headers to improve the data packet processing and forwarding efficiency of the device.
IPv6 PMTUD	A host dynamically discovers and adjusts the MTU size on the data Tx path, saving router resources and improving IPv6 network efficiency.
IPv6 Neighbor Discovery	ND functions include router discovery, prefix discovery, parameter discovery, address auto-configuration, address resolution (like ARP), next-hop determination, Neighbor Unreachability Detection (NUD), Duplicate Address Detection (DAD), and redirection.
IPv6 Source Routing	This feature is used to specify the intermediate nodes that a packet passes through along the path to the destination address. It is similar to the IPv4 loose source routing option and loose record routing option.
Restricting the Sending Rate of ICMPv6 Error Messages	This feature prevents DoS attacks.
IPv6 HOP-LIMIT	This feature prevents useless unicast packets from being unlimitedly transmitted on the network and wasting network bandwidth.

3.3.1 IPv6 Address Format

An IPv6 address is represented in the X:X:X:X:X:X:X format, where X is a 4-digit hexadecimal integer (16 bits). Each address consists of 8 integers, with a total of 128 bits (each integer contains 4 hexadecimal digits and each digit contains four bits). The following are three valid IPv6 addresses:

2001:ABCD:1234:5678:AAAA:BBBB:1200:2100

800:0:0:0:0:0:0:1

1080:0:0:0:8:800:200C:417A

These integers are hexadecimal, where A to F represent 10 to 15. Each integer in the address must be represented, except the leading zeros in each integer. If an IPv6 address contains a string of zeros (as shown in the second and third examples above), a double colon (::) can be used to represent these zeros. That is, 800:0:0:0:0:0:0:1 can be represented as 800::1.

A double colon indicates that this address can be extended to a complete 128-bit address. In this approach, only when the 16-bit integers are all 0s, can they can be replaced with a double colon. A double colon can exist once in an IPv6 address.

In IPv4/IPv6 mixed environment, an address has a mixed representation. In an IPv6 address, the least significant 32 bits can be used to represent an IPv4 address. This IPv6 address can be represented in a mixed manner, that is, X:X:X:X:X:X:d.d.d.d, where X is a hexadecimal integer and d is a 8-bit decimal integer. For example, 0:0:0:0:0:0:192.168.20.1 is a valid IPv6 address. It can be abbreviated to ::192.168.20.1. Typical applications are IPv4-compatible IPv6 addresses and IPv4-mapped IPv6 addresses. If the first 96 bits are 0 in an IPv4-compatible IPv6 address, this address can be represented as ::A.B.C.D, e.g., ::1.1.1.1. IPv4-compatible addresses have been abolished at present. IPv4-mapped IPv6 addresses are represented as ::FFFF:A.B.C.D to represent IPv4 addresses as IPv6 addresses. For example, IPv4 address 1.1.1.1 mapped to an IPv6 address is represented as ::FFFF:1.1.1.1.

Since an IPv6 address is divided into two parts: subnet prefix and interface ID, it can be represented as an address with an additional value according to an address allocation method like Classless Inter-Domain Routing (CIDR). The additional value indicates how many bits (subnet prefix) in the address represent the network part. That is, the IPv6 node address contains the prefix length. The prefix length is separated from the IPv6 address by a slash. For example, in 12AB::CD30:0:0:0:0/60, the prefix length used for routing is 60 bits.

Related Configuration

Configuring an IPv6 Address

- No IPv6 address is configured on interfaces by default.
- Run the **ipv6 address** command to configure an IPv6 address on an interface.
- After configuration, a host can communicate with others using the configured IPv6 address based on DAD.

3.3.2 IPv6 Address Type

RFC 4291 defines three types of IPv6 addresses:

- Unicast address: ID of a single interface. Packets destined to a unicast address are sent to the interface identified by this address.
- Multicast address: ID of an interface group (the interfaces generally belong to different nodes). Packets destined to a multicast address are sent to all interfaces included in this address.
- Anycast address: ID of an interface group. Packets destined to an anycast address are sent to one interface included in this address (the nearest interface according to the routing protocol).

IPv6 does not define broadcast addresses.

These three types of addresses are described as follows:

Unicast Addresses

Unicast addresses fall into five types: unspecified address, loopback address, link-local address, site-local address, and global unicast address. At present, site-local addresses have been abolished. Except unspecified, loopback, and link-local addresses, all other addresses are global unicast addresses.

- Unspecified address

The unspecified address is 0:0:0:0:0:0:0:0, which is usually abbreviated to ::. It has two general purposes:

1. If a host has no unicast address when started, it uses the unspecified address as the source address to send an RS packet to obtain prefix information from the gateway and thereby generate a unicast address.
2. When an IPv6 address is configured for a host, the device detects whether the address conflicts with addresses of other hosts in the same network segment and uses the unspecified address as the source address to send a Neighbor Solicitation (NS) packet (similar to a free ARP packet).

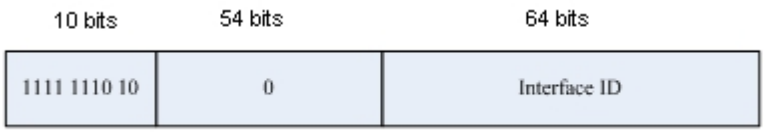
- Loopback address

The loopback address is 0:0:0:0:0:0:0:1, which is usually abbreviated to ::1. Similar to IPv4 address 127.0.0.1, the loopback address is generally used by a node to send itself packets.

- Link-local address

The format of a link-local address is as follows:

Figure 3- 2

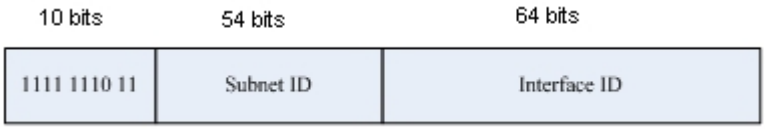


The link-local address is used on a single network link to assign IDs to hosts. The address identified by the first 10 bits in the prefix is the link-local address. A device never forwards packets in which the source or destination address contains the link-local address. The intermediate 54 bits in the address are all 0s. The last 64 bits represent the interface ID, which allows a single network to connect 2^{64}-1 hosts.

- Site-local address

The format of a site-local address is as follows:

Figure 3- 3

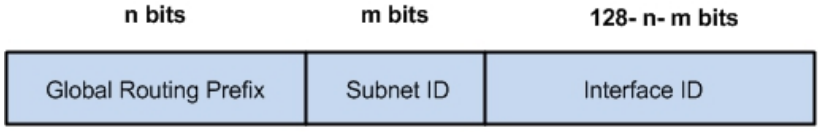


A site-local address is used to transmit data within a site. A device never forwards packets in which the source or destination address contains the site-local address to the Internet. That is, these packets can be forwarded only within the site. A site can be assumed as an enterprise's local area network (LAN). Such addresses are similar to IPv4 private addresses such as 192.168.0.0/16. RFC 3879 has abolished site-local addresses. New addresses do not support the first 10 bits as the prefix and are all regarded as global unicast addresses. Existing addresses can continue to use this prefix.

- Global unicast address

The format of a global unicast address is as follows:

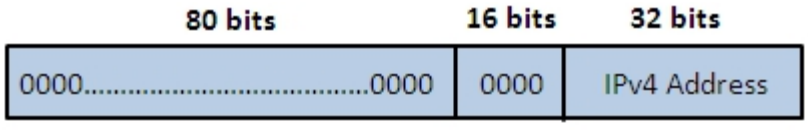
Figure 3- 4



Among global unicast addresses, there is a type of IPv4-embedded IPv6 addresses, including IPv4-compatible IPv6 addresses and IPv4-mapped IPv6 addresses. They are used for interconnection between IPv4 nodes and IPv6 nodes.

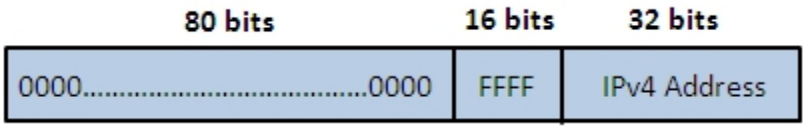
The format of an IPv4-compatible IPv6 address is as follows:

Figure 3- 5



The format of an IPv4-mapped IPv6 address is as follows:

Figure 3- 6



IPv4-compatible IPv6 addresses are mainly used on automatic tunnels. Nodes on automatic tunnels support both IPv4 and IPv6. Using these addresses, IPv4 devices transmit IPv6 packets over tunnels. At present, IPv4-compatible IPv6 addresses have been abolished. IPv4-mapped IPv6 addresses are used by IPv6 nodes to access IPv4-only nodes. For example, if the IPv6 application on an IPv4/IPv6 host requests to resolve the name of an IPv4-only host, the name server dynamically generates an IPv4-mapped IPv6 address and returns it to the IPv6 application.

↘ Multicast Addresses

The format of an IPv6 multicast address is as follows:

```
|    8      |   4 |   4 |                    112 bits                    |
+--------+----+----+-------------------------------------------+
|11111111|flgs|scop|                    group ID                    |
+--------+----+----+-------------------------------------------+
```

The first byte in the address is all 1s, representing a multicast address.

- Flag field

The flag field consists of four bits. Currently only the fourth bit is specified to indicate whether this address is a known multicast address assigned by the Internet Assigned Numbers Authority (IANA) or a temporary multicast address in a certain scenario. If the flag bit is 0, this address is a known multicast address. If the flag bit is 1, this address is a temporary multicast address. The remaining three flag bits are reserved for future use.

- Scope field

The scope field consists of four bits to indicate the multicast range. That is, a multicast group includes the local node, local link, local site, and any node in the IPv6 global address space.

- Group ID field

The group ID consists of 112 bits to identify a multicast group. A multicast ID can represent different groups based on the flag and scope fields.

IPv6 multicast addresses are prefixed with FF00::/8. One IPv6 multicast address usually identifies interfaces on a series of different nodes. After a packet is sent to a multicast address, the packet is then forwarded to the interfaces on each node identified by this multicast address. For a node (host or device), you must add the following multicast addresses:

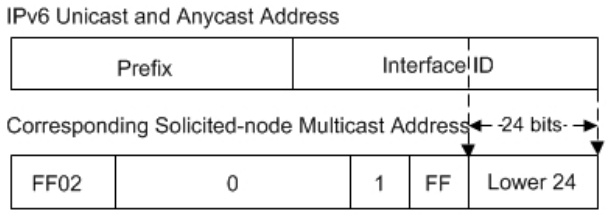
3. Multicast address for all nodes on the local link, that is, FF02::1
4. Solicited-node multicast address, prefixed with FF02:0:0:0:0:1:FF00:0000/104

If the node is a device, it also has to be added to the multicast address of all devices on the local link, that is, FF02::2.

The solicited-node multicast address corresponds to the IPv6 unicast and anycast address. You must add a corresponding solicited-node multicast address for each configured unicast and anycast address of an IPv6 node. The solicited-node multicast address is prefixed with FF02:0:0:0:0:1:FF00:0000/104. The remaining 24 bits are composed of the least significant 24 bits of the unicast or anycast address. For example, if the unicast address is FE80::2AA:FF:FE21:1234, the solicited-node multicast address is FF02::1:FF21:1234.

The solicited-node multicast address is usually used in NS packets. Its address format is as follows:

Figure 3- 7



Anycast Addresses

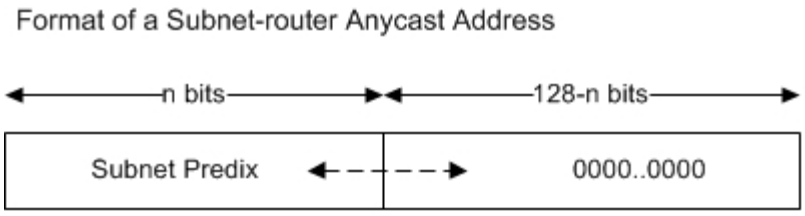
Similar to a multicast address, an anycast address can also be shared by multiple nodes. The difference is that only one node in the anycast address receives data packets while all nodes included in the multicast address receive data packets. Since anycast addresses are allocated to the normal IPv6 unicast address space, they have the same formats with unicast addresses. Every member in an anycast address must be configured explicitly for easier recognition.

Anycast addresses can be allocated only to devices and cannot be used as source addresses of packets.

RFC 2373 redefines an anycast address called subnet-router anycast address. Figure 3- 8 shows the format of a subnet-router anycast address. Such an address consists of the subnet prefix and a series of 0s (interface ID).

The subnet prefix identifies a specified link (subnet). Packets destined to the subnet-router anycast address will be forwarded to a device on this subnet. A subnet-router anycast address is usually used by the application on a node to communicate with a device on a remote subnet.

Figure 3- 8



Related Configuration

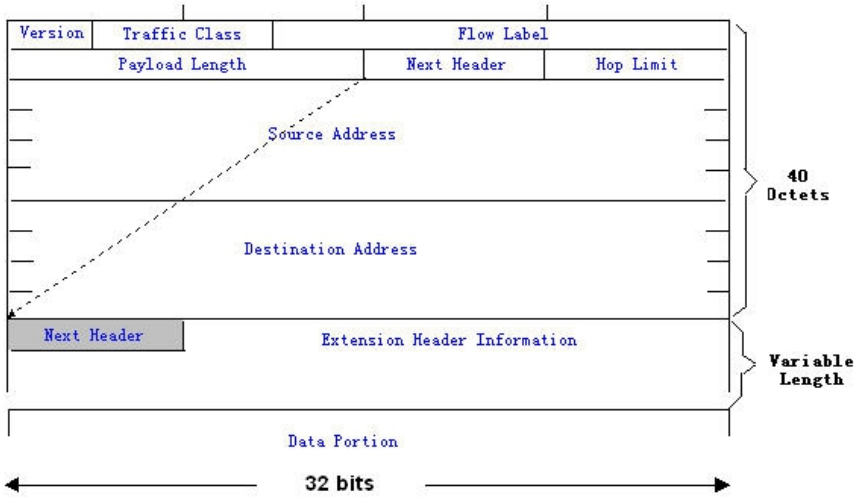
Configuring an IPv6 Address

- No IPv6 address is configured on interfaces by default.
- Run the **ipv6 address** command to configure the IPv6 unicast address and anycast address of an interface.
- After an interface goes up, it will automatically join the corresponding multicast group.

3.3.3 IPv6 Packet Header Format

Figure 3- 9 shows the format of the IPv6 packet header.

Figure 3- 9



The IPv4 packet header is in unit of four bytes. The IPv6 packet header consists of 40 bytes, in unit of eight bytes. The IPv6 packet header has the following fields:

- Version

This field consists of 4 bits. In an IPv6 address, this field must be 6.

- Traffic Class

This field consists of 8 bits. This field indicates the service provided by this packet, similar to the TOS field in an IPv4 address.

- Flow Label

This field consists of 20 bits to identify packets belonging to the same service flow. One node can act as the Tx source of multiple service flows. The flow label and source address uniquely identify one service flow.

- Payload Length

This field consists of 16 bits, including the packet payload length and the length of IPv6 extended options (if available). That is, it includes the IPv6 packet length except the IPv6 packet header.

- Next Header

This field indicates the protocol type in the header field following the IPv6 packet header. Similar to the Protocol field in the IPv4 address header, the Next Header field is used to indicate whether the upper layer uses TCP or UDP. It can also be used to indicate existence of the IPv6 extension header.

- Hop Limit

This field consists of 8 bits. Every time a device forwards a packet, the field value reduced by 1. If the field value reaches 0, this packet will be discarded. It is similar to the Lifetime field in the IPv4 packet header.

- Source Address

This field consists of 128 bits and indicates the sender address in an IPv6 packet.

- Destination Address

This field consists of 128 bits and indicates the receiver address in an IPv6 packet.

At present, IPv6 defines the following extension headers:

- Hop-By-Hop Options

This extension header must follow the IPv6 packet header. It consists of option data to be checked on each node along the path.

- Routing Options (Type 0 routing header)

This extension header indicates the nodes that a packet passes through from the source address to the destination address. It consists of the address list of the passerby nodes. The initial destination address in the IPv6 packet header is the first address among the addresses in the routing header, but not the final destination address of the packet. After the node corresponding to the destination address in the IPv6 packet header receives a packet, it processes the IPv6 packet header and routing header, and sends the packet to the second address, the third address, and so on in the routing header list till the packet reaches the final destination address.

- Fragment

The source node uses this extension header to fragment the packets of which the length exceeds the path MTU (PMTU).

- Destination Options

This extension header replaces the option fields of IPv4. At present, the Destination Options field can only be filled with integral multiples of 64 bits (eight bytes) if required. This extension header can be used to carry information to be checked by the destination node.

- Upper-layer header

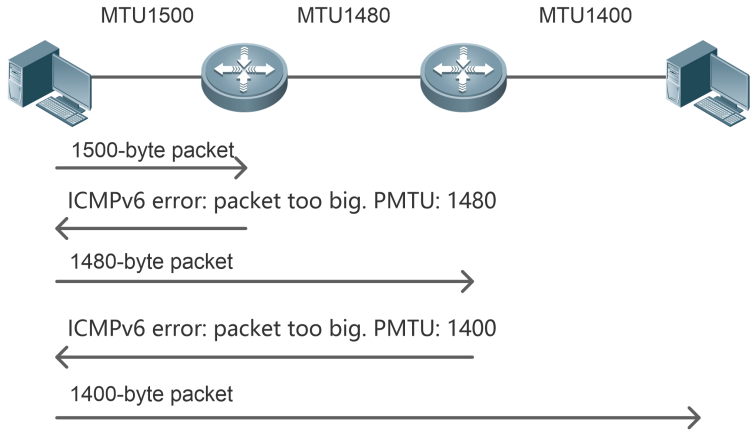
This extension header indicates the protocol used at the upper layer, such as TCP (6) and UDP (17).

Another two extension headers AH and ESP will be described in the *Configuring IPSec*.

3.3.4 IPv6 PMTUD

Similar to IPv4 Path MTU Discovery (PMTUD), IPv6 PMTUD allows a host to dynamically discover and adjust the MTU size on the data Tx path. If the length of a data packet to be sent by a host is greater than the PMTU, the host performs packet fragmentation on its own. In this manner, the IPv6 device does not need to perform fragmentation, saving device resources and improving the IPv6 network efficiency.

Figure 3- 10



As shown in Figure 3- 10, if the length of a packet to be sent by the host is greater than the PMTU, the router discards this packet and sends an ICMPv6 Packet Too Big message containing its PMTU to the host. The host then fragments the packet based on the new PMTU.

In this manner, the router does not need to perform fragmentation, saving router resources and improving the IPv6 network efficiency.

Related Configuration

Configuring the IPv6 MTU of an Interface

- The default IPv6 MTU is 1500 on an Ethernet interface.
- Run the **ipv6 mtu** command to modify the IPv6 MTU of an interface.

Configuring the Static PMTU

- Run the **ipv6 path-mtu** command to configure the static PMTU.

Configuring the Aging Time of the Dynamic PMTU

- The default aging time is 10 minutes.
- Run the **ipv6 path-mtu age** command to configure the aging time of the dynamic PMTU.

3.3.5 IPv6 Neighbor Discovery

NDP is a basic part of IPv6. Its main functions include router discovery, prefix discovery, parameter discovery, address auto-configuration, address resolution (like ARP), next-hop determination, NUD, DAD, and redirection. NDP defines five ICMP packets: RS (ICMP type: 133), RA (ICMP type: 134), NS (similar to ARP request, ICMP type: 135), NA (similar to ARP reply, ICMP type: 136), ICMP Redirect (ICMP type: 137).

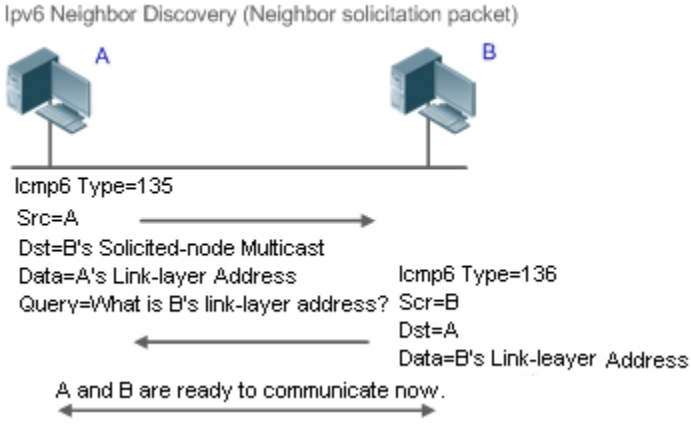
All the above ICMP packets carry one or multiple options. These options are optional in some cases but are significant in other cases. NDP mainly defines five options: Source Link-Layer Address Option, Type=1; Target Link-Layer Address Option, Type=2; Prefix Information Option, Type=3; Redirection Header Option, Type=4; MTU Option, Type=5.

Address Resolution

When a node attempts to communicate with another, the node has to obtain the link-layer address of the peer end by sending it an NS packet. In this packet, the destination address is the solicited-node multicast address corresponding to the IPv6 address of the destination node. This packet also contains the link-layer address of the source node. After receiving this NS packet, the peer end replies with an NA packet in which the destination address is the source address of the NS packet, that is, the link-layer address of the solicited node. After receiving this NA packet, the source node can communicate with the destination node.

Figure 3- 11 shows the address resolution process.

Figure 3- 11



NUD

If the reachable time of a neighbor has elapsed but an IPv6 unicast packet needs to be sent to it, the device performs NUD.

While performing NUD, the device can continue to forward IPv6 packets to the neighbor.

DAD

To know whether the IPv6 address configured for a host is unique, the device needs to perform DAD by sending an NS packet in which the source IPv6 address is the unspecified address.

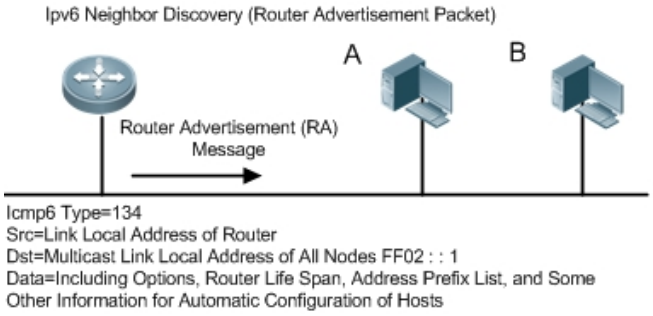
If a device detects an address conflict, this address is set to the duplicate status so that the device cannot receive IPv6 packets with this address being the destination address. Meanwhile, the device also starts a timer for this duplicate address to periodically perform DAD. If no address conflict is detected in re-detection, this address can be properly used.

Router, Prefix, and Parameter Discovery

A device periodically sends RA packets to all local nodes on the link.

Figure 3- 12 shows the RA packet sending process.

Figure 3- 12



An RA packet usually contains the following content:

- One or multiple IPv6 address prefixes (used for on-link determination or stateless address auto-configuration)
- Validity of the IPv6 address prefix
- Host auto-configuration method (stateful or stateless)
- Default device information (whether the device acts as the default device; if yes, the interval for acting as the default device is also included.)
- Other information provided for host configuration, such as hop limit, MTU, and NS retransmission interval

RA packets can also be used as replies to the RS packets sent by a host. Using RS packets, a host can obtain the auto-configured information immediately after started rather than wait for the RA packets sent by the device. If no unicast address is configured for a newly started host, the host includes the unspecified address (0:0:0:0:0:0:0:0) as the source address in the RS packet. Otherwise, the host uses the configured unicast address as the source address and the multicast address of all local routing devices (FF02::2) as the destination address in the RS packet. As an reply to the RS packet, the RA packet uses the source address of the RS packet as the destination address (if the source address is the unspecified address, it uses the multicast address of all local nodes (FF02::1).

In an RA packet, the following parameters can be configured:

- Ra-interval: Interval for sending the RA packet.

- Ra-lifetime: Lifetime of a router, that is, whether the device acts as the default router on the local link and the interval for acting as the default router.
- Prefix: Prefix of an IPv6 address on the local link. It is used for on-link determination or stateless address auto-configuration, including other parameter configurations related to the prefix.
- Ns-interval: NS packet retransmission interval.
- Reachabletime: Period when the device regards a neighbor reachable after detecting a Confirm Neighbor Reachability event.
- Ra-hoplimit: Hops of the RA packet, used to set the hop limit for a host to send a unicast packet.
- Ra-mtu: MTU of the RA packet.
- Managed-config-flag: Whether a host receiving this RA packet obtains the address through stateful auto-configuration.
- Other-config-flag: Whether a host receiving this RA packet uses DHCPv6 to obtain other information except the IPv6 address for auto-configuration.

Configure the above parameters when configuring IPv6 interface attributes.

Redirection

If a router receiving an IPv6 packet finds a better next hop, it sends the ICMP Redirect packet to inform the host of the better next hop. The host will directly send the IPv6 packet to the better next hop next time.

Maximum Number of Unresolved ND Entries

- You can configure the maximum number of unresolved ND entries to prevent malicious scanning network segments from generating a large number of unresolved ND entries and occupying excessive memory space.

Maximum Number of ND Options

- You can configure the maximum number of ND options to prevent forged ND packets from carrying unlimited ND options and occupying excessive CPU space on the device.

Maximum Number of Neighbor Learning Entries on an Interface

- You can configure the maximum number of neighbor learning entries on an interface to prevent neighbor learning attacks from occupying ND entries and memory space of the device and affecting forwarding efficiency of the device.

Related Configuration

Enabling IPv6 Redirection

- By default, ICMPv6 Redirect packets can be sent on IPv6 interfaces.
- Run the **no ipv6 redirects** command in interface configuration mode to prohibit an interface from sending Redirect packets.

Configuring IPv6 DAD

- By default, an interface sends one NS packet to perform IPv6 DAD.
- Run the **ipv6 nd dad attempts** *value* command in interface configuration mode to configure the number of NS packets consecutively sent by DAD. Value 0 indicates disabling DAD for IPv6 addresses on this interface.
- Run the **no ipv6 nd dad attempts** command to restore the default configuration.
- By default, the device performs DAD on duplicate IPv6 addresses every 60 seconds.

- Run the **ipv6 nd dad retry** value command in global configuration mode to configure the DAD interval. Value 0 indicates disabling DAD for the device.
- Run the **no ipv6 nd dad retry** command to restore the default configuration.

Configuring the Reachable Time of a Neighbor

- The default reachable time of an IPv6 neighbor is 30s.
- Run the **ipv6 nd reachable-time** *milliseconds* command in interface configuration mode to modify the reachable time of a neighbor.

Configuring the Stale Time of a Neighbor

- The default stale time of an IPv6 neighbor is 1 hour. After the time elapses, the device performs NUD.
- Run the **ipv6 nd stale-time** *seconds* command in interface configuration mode to modify the stale time of a neighbor.

Configuring Prefix Information

- By default, the prefix in an RA packet on an interface is the prefix configured in the **ipv6 address** command on the interface.
- Run the **ipv6 nd prefix** command in interface configuration mode to add or delete prefixes and prefix parameters that can be advertised.

Enabling/disabling RA Suppression

- By default, an IPv6 interface does not send RA packets.
- Run the **no ipv6 nd suppress-ra** command in interface configuration mode to disable RA suppression.

Configuring the Maximum Number of Unresolved ND Entries

- The default value is 0, indicating no restriction. It is only restricted to the ND entry capacity supported by the device.
- Run the **ipv6 nd unresolved** *number* command in global configuration mode to restrict the number of unresolved neighbors. After the entries exceed this restriction, the device does not actively resolve subsequent packets.

Configuring the Maximum Number of ND Options

- Run the **ipv6 nd max-opt** *value* command in global configuration mode to restrict the number of ND options to be processed. The default value is 10.

Configuring the Maximum Number of ND Entries Learned on an Interface

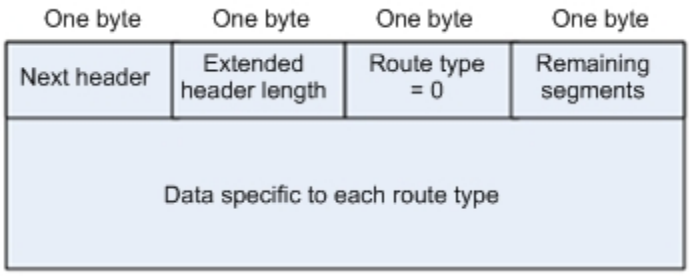
- Run the **ipv6 nd cache interface-limit** *value* command in interface configuration mode to restrict the number of neighbors learned on an interface. The default value is 0, indicating no restriction.

3.3.6 IPv6 Source Routing

Working Principle

Similar to the IPv4 loose source routing and loose record routing options, the IPv6 routing header is used to specify the intermediate nodes that the packet passes through along the path to the destination address. It uses the following format:

Figure 3- 13



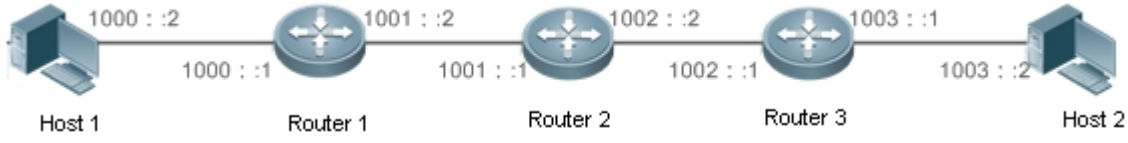
The Segments Left field is used to indicate how many intermediate nodes are specified in the routing header for the packet to pass through from the current node to the final destination address.

Currently, two routing types are defined: 0 and 2. The Type 2 routing header is used for mobile communication. RFC 2460 defines the Type 0 routing header (similar to the loose source routing option of IPv4). The format of the Type 0 routing header is as follows:

Figure 3- 14

The following example describes the application of the Type 0 routing header, as shown in Figure 3- 15.

Figure 3- 15



Host 1 sends Host 2 a packet specifying the intermediate nodes Router 2 and Router 3. The following table lists the changes of fields related to the IPv6 header and routing header during the forwarding process.

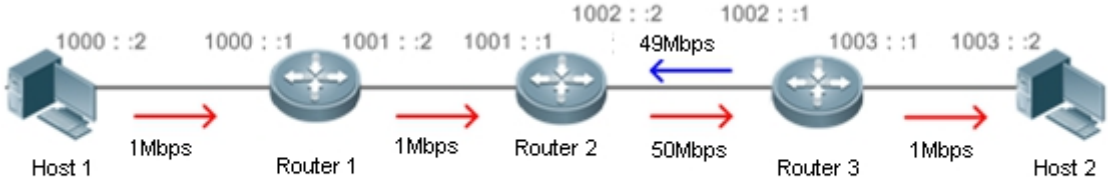
Transmission Node	Fields in the IPv6 Header	Fields Related to the Type 0 Routing Header
Host 1	Source address=1000::2 Destination address=1001::1 (Address of Router 2)	Segments Left=2 Address 1=1002::1 (Address of Router 3) Address 2=1003::2 (Address of Host 2)
Router 1	No change	
Router 2	Source address=1000::2 Destination address=1002::1 (Address of Router 3)	Segments Left=1 Address 1=1001::1 (Address of Router 2) Address 2=1003::2 (Address of Host 2)
Router 3	Source address=1000::2 Destination address=1003::2 (Address of Host 2)	Segments Left=0 Address 1=1001::1 (Address of Router 2) Address 1=1002::2 (Address of Router 3)
Host 2	No change	

The forwarding process is as follows:

1. Host 1 sends a packet in which the destination address is Router 2's address 1001::1, the Type 0 routing header is filled with Router 3's address 1002::1 and Host 2's address 1003::2, and the value of the Segments Left field is 2.
2. Router 1 forwards this packet to Router 2.
3. Router 2 changes the destination address in the IPv6 header to Address 1 in the routing header. That is, the destination address becomes Router 3's address 1002::1, Address 1 in the routing header becomes Router 2's address 1001::1, and the value of the Segments Left field becomes 1. After modification, Router 2 forwards the packet to Router 3.
4. Router 3 changes the destination address in the IPv6 header to Address 2 in the routing header. That is, the destination address becomes Host 2's address 1003::2, Address 2 in the routing header becomes Router 3's address 1002::1, and the value of the Segments Left field becomes 0. After modification, Router 3 forwards the packet to Host 2.

The Type 0 routing header may be used to initiate DoS attacks. As shown in Figure 3- 16, Host 1 sends packets to Host 2 at 1 Mbps and forges a routing header to cause multiple round-trips between Router 2 and Router 3 (50 times from Router 2 to Router 3 and 49 times from Router 3 to Router 2). At the time, the routing header generates the traffic amplification effect:" 50 Mbps from Router 2 to Router 3 and 49 Mbps from Router 3 to Router 2." Due to this security problem, RFC 5095 abolished the Type 0 routing header.

Figure 3- 16



IPv6 Packet
Source Address 1000::2
Destination Address 1001:1

Segments Left in the Type 0
Routing Header: 100
Address 1: 1002::1
Address 2: 1001::1
Address 3: 1002::1
Address 4: 1002::1
...
Address 99: 1002::1
Address 100: 100:3::2

Host 1 sends packets to Host 2, passing through Router 2, Router 3, Router 2, and Router 3, ...
Each packet is sent 50 times from Router 2 to Router 3 and 49 times from Router 3 to Router 2.

Related Configuration

Enabling IPv6 Source Routing

- The Type 0 routing header is not supported by default.
- Run the **ipv6 source-route** command in global configuration mode to enable IPv6 source routing.

3.3.7 Restricting the Sending Rate of ICMPv6 Error Messages

Working Principle

The destination node or intermediate router sends ICMPv6 error messages to report the errors incurred during IPv6 data packet forwarding and transmission. There are mainly four types of error messages: Destination Unreachable, Packet Too Big, Time Exceeded, and Parameter Problem.

When receiving an invalid IPv6 packet, a device discards the packet and sends back an ICMPv6 error message to the source IPv6 address. In the case of invalid IPv6 packet attacks, the device may continuously reply to ICMPv6 error messages till device resources are exhausted and thereby fail to properly provide services. To solve this problem, you can restrict the sending rate of ICMPv6 error messages.

If the length of an IPv6 packet to be forwarded exceeds the IPv6 MTU of the outbound interface, the router discards this IPv6 packet and sends back an ICMPv6 Packet Too Big message to the source IPv6 address. This error message is mainly used as part of the IPv6 PMTUD process. If the sending rate of ICMPv6 error messages is restricted due to excessive other ICMPv6 error messages, ICMPv6 Packet Too Big messages may be filtered, causing failure of IPv6 PMTUD. Therefore, it is recommended to restrict the sending rate of ICMPv6 Packet Too Big messages independently of other ICMPv6 error messages.

Although ICMPv6 Redirect packets are not ICMPv6 error messages, FS recommends restricting their rates together with ICMPv6 error messages except Packet Too Big messages.

Related Configuration

Configuring the Sending Rate of ICMPv6 Packet Too Big Messages

- The default rate is 10 per 100 ms.
- Run the **ipv6 icmp error-interval too-big** command to configure the sending rate of ICMPv6 Packet Too Big messages.

Configuring the Sending Rate of Other ICMPv6 Error Messages

- The default rate is 10 per 100 ms.
- Run the **ipv6 icmp error-interval** command to configure the sending rate of other ICMPv6 error messages.

3.3.8 IPv6 Hop Limit

Working Principle

An IPv6 data packet passes through routers from the source address and destination address. If a hop limit is configured, it decreases by one every time the packet passes through a router. When the hop limit decreases to 0, the router discards the packet to prevent this useless packet from being unlimitedly transmitted on the network and wasting network bandwidth. The hop limit is similar to the TTL of IPv4.

Related Configuration

Configuring the IPv6 Hop Limit

- The default IPv6 hop limit of a device is 64.
- Run the **ipv6 hop-limit** command to configure the IPv6 hop limit of a device.
-

3.4 Configuration

Configuration	Description and Command	
Configuring an IPv6 Address	(Mandatory) It is used to configure IPv6 addresses and enable IPv6.	
	ipv6 enable	Enables IPv6 on an interface.
	ipv6 address	Configures the IPv6 unicast address of an interface.
Configuring IPv6 NDP	(Optional) It is used to enable IPv6 redirection on an interface.	
	ipv6 redirects	Enables IPv6 redirection on an interface.
	(Optional) It is used to enable DAD.	
	ipv6 nd dad attempts	Configures the number of consecutive NS packets sent during DAD.
	(Optional) It is used to configure ND parameters.	
	ipv6 nd reachable-time	Configures the reachable time of a neighbor.
	ipv6 nd prefix	Configures the address prefix to be advertised in an RA packet.
	ipv6 nd suppress-ra	Enables RA suppression on an interface.
	(Optional) It is used to configure the maximum number of unresolved ND entries.	
	ipv6 nd unresolved	Configures the maximum number of unresolved ND entries.
	(Optional) It is used to configure the maximum number of ND options.	
	ipv6 nd max-opt	Configures the maximum number of ND options.

Configuration	Description and Command	
	(Optional) It is used to configure the maximum number of neighbors learned on an interface.	
	ipv6 nd cache interface-limit	Configures the maximum number of neighbors learned on an interface.
Enabling PMTUD	(Optional) It is used to restrict the MTU of IPv6 packets sent on an interface.	
	ipv6 mtu	Configures the IPv6 MTU.
Enabling IPv6 Source Routing	(Optional) It is used to enable IPv6 source routing.	
	ipv6 source-route	Configures the device to forward IPv6 packets carrying the routing header.
Configuring the Sending Rate of ICMPv6 Error Messages	Optional.	
	ipv6 icmp error-interval too-big	Configures the sending rate of ICMPv6 Packet Too Big messages.
	ipv6 icmp error-interval	Configures the sending rates of other ICMPv6 error messages and ICMPv6 Redirect packets.
Configuring the IPv6 Hop Limit	(Optional) It is used to restrict the hop limit of IPv6 unicast packets sent on an interface.	
	ipv6 hop-limit	Configures the IPv6 hop limit.

3.4.1 Configuring an IPv6 Address

Configuration Effect

Configure the IPv6 address of an interface to implement IPv6 network communication.

Configuration Steps

Enabling IPv6 on an Interface

- (Optional) If you do not want to enable IPv6 by configuring an IPv6 address, run the **ipv6 enable** command.

Configuring the IPv6 Unicast Address of an Interface

- Mandatory.

Verification

Run the **show ipv6 interface** command to check whether the configured address takes effect.

Related Commands

Enabling IPv6 on an Interface

Command	ipv6 enable
Parameter Description	N/A
Command Mode	Interface configuration mode

Usage Guide	IPv6 can be enabled on an interface by two methods: 1) running the **ipv6 enable** command in interface configuration mode; 2) configuring an IPv6 address on the interface. If an IPv6 address is configured on an interface, IPv6 is automatically enabled on this interface. In this case, IPv6 cannot be disabled even when you run the **no ipv6 enable** command.

Configuring the IPv6 Unicast Address of an Interface

Command	**ipv6 address** *ipv6-address / prefix-length* **ipv6 address** *ipv6-prefix / prefix-length* **eui-64** **ipv6 address** *prefix-name sub-bits / prefix-length* [**eui-64**]
Parameter Description	*ipv6-address*: Indicates the IPv6 address, which must comply with the address format defined in RFC 4291. Separated by a colon (:), each address field consists of 16 bits and is represented by hexadecimal digits. *ipv6-prefix*: Indicates the IPv6 address prefix, which must comply with the address format defined in RFC 4291. *prefix-length*: Indicates the length of the IPv6 address prefix, that is, the part representing the network in the IPv6 address. *prefix-name*: Indicates the name of the universal prefix. This specified universal prefix is used to create the interface address. *sub-bits*: Indicates the subprefix bits and host bits of the address to be concatenated with the prefixes provided by the general prefix specified with the *prefix-name* parameter. This value is combined with the universal prefix to create the interface address. This value must be in the form documented in RFC 4291. *eui-64*: Indicates the created IPv6 address, consisting of the configured address prefix and 64-bit interface ID.
Command Mode	Interface configuration mode
Usage Guide	If an IPv6 interface is created and is Up state, the system automatically generates a link-local address for this interface. The IPv6 address of an interface can also be created by the universal prefix mechanism. That is, IPv6 address = Universal prefix + Sub prefix + Host bits. The universal prefix can be configured by running the **ipv6 general-prefix** command or learned by the prefix discovery function of the DHCPv6 client (see the *Configuring DHCPv6*). Sub prefix + Host bits are specified by the *sub-bits* and *prefix-length* parameters in the **ipv6 address** command. If you run the **no ipv6 address** command without specifying an address, all manually configured addresses will be deleted. Run the **no ipv6 address** *ipv6-prefix/prefix-length* **eui-64** command to delete the configured address.

Configuration Example

Configuring an IPv6 Address on an Interface

Configuration Steps	Enable IPv6 on the GigabitEthernet 0/0 interface and add IPv6 address 2000::1 to the interface.
	FS(config)#interface gigabitEthernet 0/0 FS(config-if-GigabitEthernet 0/0)#ipv6 enable FS(config-if-GigabitEthernet 0/0)#ipv6 address 2000::1/64
Verification	Run the **show ipv6 interface** command to verify that an address is successfully added to the GigabitEthernet 0/0 interface.
	FS(config-if-GigabitEthernet 0/0)#show ipv6 interface gigabitEthernet 0/0

```
interface GigabitEthernet 0/0 is Down, ifindex: 1, vrf_id 0
  address(es):
    Mac Address: 00:00:00:00:00:00
    INET6: FE80::200:FF:FE00:1 [ TENTATIVE ], subnet is FE80::/64
    INET6: 2000::1 [ TENTATIVE ], subnet is 2000::/64
  Joined group address(es):
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND advertised reachable time is 0 milliseconds
  ND retransmit interval is 1000 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 200 seconds<160--240>
  ND router advertisements live for 1800 seconds
```

The current version does not support VRF. The configuration example is only for reference.

3.4.2 Enabling PMTUD

Configuration Effect

When sending an IPv6 packet, a host fragments the packet based on the PMTU.

Notes

The IPv6 MTU of an interface must be less than or equal to the interface MTU.

Configuration Steps

Configuring the IPv6 MTU of an Interface

- Optional.

Verification

- Run the **show run** command to check whether the configuration is correct.
- Run the **show ipv6 interface** command to check whether the IPv6 MTU of an interface is correct.

Related Commands

Configuring the IPv6 MTU of an Interface

Command	**ipv6 mtu** *bytes*
Parameter Description	*bytes*: Indicates the MTU of an IPv6 packet, ranging from 1280 to 1500. The unit is byte.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring the IPv6 MTU of an Interface

Configuration Steps	Change the IPv6 MTU of interface GigabitEthernet 0/0 to 1,300.
	FS(config-if-GigabitEthernet 0/0)#ipv6 mtu 1300
Verification	Run the **show ipv6 interface** command to check whether the configuration takes effect.
	FS(config-if-GigabitEthernet 0/0)#show ipv6 interface interface GigabitEthernet 0/ is Down, ifindex: 1, vrf_id 0 address(es): Mac Address: 00:d0:f8:22:33:47 INET6: FE80::2D0:F8FF:FE22:3347 [TENTATIVE], subnet is FE80::/64 INET6: 1020::1 [TENTATIVE], subnet is 1020::/64 INET6: 1023::1 [TENTATIVE], subnet is 1023::/64 Joined group address(es): MTU is 1300 bytes ICMP error messages limited to one every 100 milliseconds ICMP redirects are enabled ND DAD is enabled, number of DAD attempts: 1 ND reachable time is 30000 milliseconds ND advertised reachable time is 0 milliseconds ND retransmit interval is 1000 milliseconds ND advertised retransmit interval is 0 milliseconds ND router advertisements are sent every 200 seconds<160--240> ND router advertisements live for 1800 seconds

3.4.3 Configuring IPv6 NDP

Configuration Effect

Configure NDP-related attributes, for example, enable IPv6 redirection and DAD.

Notes

RA suppression is enabled on interfaces by default. To configure a device to send RA packets, run the **no ipv6 nd suppress-ra** command in interface configuration mode.

Configuration Steps

Enabling IPv6 Redirection on an Interface

- (Optional) IPv6 redirection is enabled by default.
- To disable IPv6 redirection on an interface, run the **no ipv6 redirects** command.

Configuring the Number of Consecutive NS Packets Sent During DAD

- Optional.
- To prevent enabling DAD for IPv6 addresses on an interface or modify the number of consecutive NS packets sent during DAD, run the **ipv6 nd dad attempts** command.

Configuring the Reachable Time of a Neighbor

- Optional.
- To modify the reachable time of a neighbor, run the **ipv6 nd reachable-time** command.

Configuring the Address Prefix to Be Advertised in an RA Packet

- By default, the prefix in an RA packet on an interface is the prefix configured in the **ipv6 address** command on the interface.

Enabling/Disabling RA Suppression on an Interface

- Optional.
- If a device needs to send RA packets, run the **no ipv6 nd suppress-ra** command.

Configuring the Maximum Number of Unresolved ND Entries

- Optional.
- If a large number of unresolved ND entries are generated due to scanning attacks, run the **ipv6 nd unresolved** command to restrict the number of unresolved neighbors.

Configuring the Maximum Number of ND Options

- Optional.
- If a device needs to process more options, run the **ipv6 nd max-opt** command.

Configuring the Maximum Number of ND Entries Learned on an Interface

- Optional.
- If the number of IPv6 hosts is controllable, run the **ipv6 nd cache interface-limit** command to restrict the number of neighbors learned on an interface. This prevents ND learning attacks from occupying the memory space and affecting device performance.

Verification

Run the following commands to check whether the configuration is correct:

- **show ipv6 interface** *interface-type interface-num*: Check whether the configurations such as the redirection function, reachable time of a neighbor, and NS sending interval take effect.
- **show ipv6 interface** *interface-type interface-num* **ra-inifo**: Check whether the prefix and other information configured for RA packets are correct.
- **show run**

Related Commands

Enabling IPv6 Redirection on an Interface

Command	**ipv6 redirects**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	All ICMPv6 error messages are transmitted at a limited transmission rate. By default, a maximum number of 10 ICMPv6 error messages are transmitted per second (10 pps).

Configuring the Number of Consecutive NS Packets Sent During DAD

Command	**ipv6 nd dad attempts** *value*
Parameter Description	*value*: Indicates the number of NS packets.
Command Mode	Interface configuration mode
Usage Guide	You need to enable DAD before configuring an IPv6 address on an interface. Then the address is in tentative state. If no address conflict is detected by DAD, this address can be correctly used. If an address conflict is detected and the interface ID of this address uses EUI-64, duplicate link-layer addresses exist on this link. In this case, the system automatically disables this interface to prevent IPv6-related operations on this interface). At the time, you must configure a new address and restart the interface to re-enable DAD. When an interface changes from the down state to the up state, DAD is re-enabled for the addresses on this interface.

Configuring the Reachable Time of a Neighbor

Command	**ipv6 nd reachable-time** *milliseconds*
Parameter Description	*milliseconds*: Indicates the reachable time of a neighbor, ranging from 0 to 3,600,000. The unit is millisecond. The default value is 30s.
Command Mode	Interface configuration mode
Usage Guide	A device detects unreachable neighbors based on the configured reachable time. The shorter the configured reachable time, the faster the device detects unreachable neighbors but the more it consumes network bandwidth and device resources. Therefore, it is not recommended to set this time too small. The configured value is advertised in an RA packet and is also used on the device. If the value is 0, the reachable time is not specified on the device and it is recommended to use the default value.

Configuring the Address Prefix to Be Advertised in an RA Packet

Command	**ipv6 nd prefix** {*ipv6-prefix/prefix-length* | **default**} [[*valid-lifetime* { **infinite** | *preferred-lifetime* }] | [**at** *valid-date preferred-date*] | [**infinite** {**infinite** | *preferred-lifetime*}]] [**no-advertise**] | [[**off-link**] [**no-autoconfig**]]
Parameter Description	*ipv6-prefix*: Indicates the network ID of IPv6, which must comply with the address representation format in RFC 4291. *prefix-length*: Indicates the length of the IPv6 address prefix. A slash (/) must be added before the prefix. *valid-lifetime*: Indicates the period when a host receiving the prefix of an RA packet regards the prefix valid. The value ranges from 0 to 4,294,967,295. The default value is 30 days. *preferred-lifetime*: Indicates the preferred period when a host receiving the prefix of an RA packet regards the prefix valid. The value ranges from 0 to 4,294,967,295. The default value is 7 days. **at** *valid-date preferred-date*: Indicates the valid date and preferred deadline configured for the RA prefix. It uses the format of *dd+mm+yyyy+hh+mm*. **infinite**: Indicates that the prefix is permanently valid. **default**: Indicates that the default parameter configuration is used. **no-advertise**: Indicates that the prefix is not advertised by a router. **off-link**: If the prefix of the destination address in the IPv6 packet sent by a host matches the configured prefix, the device regards the destination address on the same link and directly reachable. This parameter indicates that this prefix does not require on-link determination. **no-autoconfig**: Indicates that the prefix in the RA packet received by a host cannot be used for address auto-configuration.
Command Mode	Interface configuration mode
Usage Guide	This command can be used to configure parameters related to each prefix, including whether to advertise this prefix. By default, an RA packet uses the prefix configured by running the **ipv6 address** command. Run the **ipv6 nd prefix** command to add other prefixes. Run the **ipv6 nd prefix default** command to configure the default parameters for an interface. That is, if no parameter is specified when a prefix is added, use the parameters configured in the **ipv6 nd prefix default** command as the parameters of the new prefix. The default parameter configurations are abandoned once a parameter is specified for the prefix. That is, when you use the **ipv6 nd prefix default** command to modify the default parameter configurations, only the prefix configured for the default parameters changes and configurations of the prefix remain the same. **at** *valid-date preferred-date*: You can specify the valid date of the prefix in two methods: 1) specifying a fixed time for each prefix in an RA packet; 2) specifying the deadline. In the second method, the valid date of the prefix in each RA packet decreases till it becomes 0.

Enabling/Disabling RA Suppression on an Interface

Command	**ipv6 nd suppress-ra**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	To enable RA suppression on an interface, run the **ipv6 suppress-ra** command.

Configuring the Maximum Number of Unresolved ND Entries

Command	**ipv6 nd unresolved** *number*
Parameter	*number*: Indicates the maximum number of unresolved ND entries.

Description	
Command Mode	Global configuration mode
Usage Guide	To prevent malicious scanning attacks from creating a large number of unresolved ND entries and occupying entry resources, you can restrict the number of unresolved ND entries.

Configuring the Maximum Number of ND Options

Command	**ipv6 nd max-opt** *value*
Parameter Description	*value*: Indicates the number of supported ND options.
Command Mode	Global configuration mode
Usage Guide	Configure the maximum number of ND options processed by a device, such as link-layer address option, MTU option, redirection option, and prefix option.

Configuring the Maximum Number of ND Entries Learned on an Interface

Command	**ipv6 nd cache interface-limit** *value*
Parameter Description	*value*: Indicates the maximum number of neighbors learned by an interface.
Command Mode	Interface configuration mode
Usage Guide	Restricting the number of ND entries learned on an interface can prevent malicious neighbor attacks. If this number is not restricted, a large number of ND entries will be generated on the device, occupying excessive memory space. The configured value must be equal to or greater than the number of the ND entries learned by the interface. Otherwise, the configuration does not take effect. The configuration is subject to the ND entry capacity supported by the device.

Configuration Example

Enabling IPv6 Redirection on an Interface

Configuration Steps	Enable IPv6 redirection on interface GigabitEthernet 0/0.
	FS(config-if-GigabitEthernet 0/0)#ipv6 redirects
Verification	Run the **show ipv6 interface** command to check whether the configuration takes effect.

	FS#show ipv6 interface gigabitEthernet 0/0
	
	interface GigabitEthernet 0/0 is Down, ifindex: 1, vrf_id 0
	  address(es):
	  Mac Address: 00:00:00:00:00:00
	  INET6: FE80::200:FF:FE00:1 [TENTATIVE], subnet is FE80::/64
	  Joined group address(es):
	  MTU is 1500 bytes

Configuration Steps	Enable IPv6 redirection on interface GigabitEthernet 0/0.
	FS(config-if-GigabitEthernet 0/0)#ipv6 redirects
Verification	Run the **show ipv6 interface** command to check whether the configuration takes effect.
	ICMP error messages limited to one every 100 milliseconds ICMP redirects are enabled ND DAD is enabled, number of DAD attempts: 1 ND reachable time is 30000 milliseconds ND advertised reachable time is 0 milliseconds ND retransmit interval is 1000 milliseconds ND advertised retransmit interval is 0 milliseconds ND router advertisements are sent every 200 seconds<160--240> ND router advertisements live for 1800 seconds

Configuring IPv6 DAD

Configuration Steps	Configure the interface to send three consecutive NS packets during DAD.
	FS(config-if-GigabitEthernet 0/0)# ipv6 nd dad attempts 3
Verification	Run the **show ipv6 interface** command to check whether the configuration takes effect.
	FS#show ipv6 interface gigabitEthernet 0/0 interface GigabitEthernet 0/0 is Down, ifindex: 1, vrf_id 0 address(es): Mac Address: 00:00:00:00:00:00 INET6: FE80::200:FF:FE00:1 [TENTATIVE], subnet is FE80::/64 Joined group address(es): MTU is 1500 bytes ICMP error messages limited to one every 100 milliseconds ICMP redirects are enabled ND DAD is enabled, number of DAD attempts: 3 ND reachable time is 30000 milliseconds ND advertised reachable time is 0 milliseconds ND retransmit interval is 1000 milliseconds ND advertised retransmit interval is 0 milliseconds ND router advertisements are sent every 200 seconds<160--240>

	ND router advertisements live for 1800 seconds FS(config-if-GigabitEthernet 0/0)# The current version does not support VRF. The configuration example is only for reference.

Configuring Prefix Information in an RA Packet

Configuration Steps	Add a prefix 1234::/64 to interface GigabitEthernet 0/0.
	FS(config-if-GigabitEthernet 0/0)#ipv6 nd prefix 1234::/6
Verification	Run the **show ipv6 interface** command to check whether the configuration takes effect.
	FS#show ipv6 interface gigabitEthernet 0/0 ra-info GigabitEthernet 0/0: DOWN (RA is suppressed) RA timer is stopped waits: 0, initcount: 0 statistics: RA(out/in/inconsistent): 0/0/0, RS(input): 0 Link-layer address: 00:00:00:00:00:00 Physical MTU: 1500 ND router advertisements live for 1800 seconds ND router advertisements are sent every 200 seconds<160--240> Flags: !M!O, Adv MTU: 1500 ND advertised reachable time is 0 milliseconds ND advertised retransmit time is 0 milliseconds ND advertised CurHopLimit is 64 Prefixes: <total: 1> 1234::/64(Def, CFG, vltime: 2592000, pltime: 604800, flags: LA)

Configuring RA Packets to Obtain Prefixes from the Prefix Pool

Configuration Steps	Configure RA packets to obtain prefixes from the prefix pool "ra-pool".
	FS(config-if-GigabitEthernet 0/0)#peel default ipv6 pool ra-pool
Verification	Run the **show run** command to check whether the configuration takes effect.
	FS(config-if-GigabitEthernet 0/0)#show run interface gigabitEthernet 0/0 Building configuration... Current configuration : 125 bytes

Configuration Steps	Configure RA packets to obtain prefixes from the prefix pool "ra-pool".
	FS(config-if-GigabitEthernet 0/0)#peel default ipv6 pool ra-pool
Verification	Run the **show run** command to check whether the configuration takes effect.
	interface GigabitEthernet 0/0 ipv6 enable no ipv6 nd suppress-ra peel default ipv6 pool ra-pool !

Disabling RA Suppression

Configuration Steps	Disable RA suppression on an interface.
	FS(config-if-GigabitEthernet 0/0)# no ipv6 nd suppress-ra
Verification	Run the **show run** command to check whether the configuration takes effect.
	FS(config-if-GigabitEthernet 0/0)#show run interface gigabitEthernet 0/0 Building configuration... Current configuration : 125 bytes interface GigabitEthernet 0/0 ipv6 enable no ipv6 nd suppress-ra !

Configuring the Maximum Number of Unresolved ND Entries

Configuration Steps	Set the maximum number of unresolved ND entries to 200.
	FS(config)# ipv6 nd unresolved 200
Verification	Run the **show run** command to check whether the configuration takes effect.
	FS#show run ipv6 nd unresolved 200 !

Configuring the Maximum Number of ND Options

Configuration Steps	Set the maximum number of ND options to 20.
	FS(config)# ipv6 nd max-opt 20
Verification	Run the **show run** command to check whether the configuration takes effect.
	FS#show run ipv6 nd max-opt 20 !

Configuring the Maximum Number of ND Entries Learned on an Interface

Configuration Steps	Set the maximum number of ND entries learned on an interface to 100.
	FS(config-if-GigabitEthernet 0/1)# ipv6 nd cache interface-limit 100
Verification	Run the **show run** command to check whether the configuration takes effect.
	FS#show run ! interface GigabitEthernet 0/1 ipv6 nd cache interface-limit 100 !

3.4.4 Enabling IPv6 Source Routing

Configuration Effect

RFC 5095 abolished the Type 0 routing header. FS devices do not support the Type 0 routing header by default. The administrator can run the **ipv6 source-route** command to in global configuration mode to enable IPv6 source routing.

Configuration Steps

Enabling IPv6 Source Routing

- Optional.
- To enable IPv6 source routing, run the **ipv6 source-route** command.

Verification

The device can properly forward packets carrying the Type 0 routing header.

Related Commands

Enabling IPv6 Source Routing

Command	**ipv6 source-route**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Since the Type 0 header may cause the device prone to DoS attacks, the device does not forward IPv6 packets carrying the routing header by default, but still processes IPv6 packets with itself being the final destination address and the Type 0 routing header.

Configuration Example

Enabling IPv6 Source Routing

Configuration Steps	Enable IPv6 source routing.
	FS(config)#ipv6 source-route
Verification	Run the **show run** command to check whether the configuration takes effect.
	FS#show run | inc ipv6 source-route ipv6 source-route

3.4.5 Configuring the Sending Rate of ICMPv6 Error Messages

Configuration Effect

Configure the sending rate of ICMPv6 error messages.

Configuration Steps

Configuring the Sending Rate of ICMPv6 Packet Too Big Messages

- Optional.
- If a device receives many IPv6 packets with the packet length exceeding the IPv6 MTU of the outbound interface and thereby sends many ICMPv6 Packet Too Big messages to consume much CPU resources, run the **ipv6 icmp error-interval too-big** command to restrict the sending rate of this error message.

Configuring the Sending Rate of Other ICMPv6 Error Messages

- Optional.
- If a device receives many illegal IPv6 packets and thereby generates many ICMPv6 error messages, run the **ipv6 icmp error-interval** command to restrict the sending rate of ICMPv6 error messages. (This command does not affect the sending rate of ICMPv6 Packet Too Big messages.)

Verification

Run the **show running-config** command to check whether the configuration takes effect.

Related Commands

Configuring the Sending Rate of ICMPv6 Packet Too Big Messages

Command	**ipv6 icmp error-interval too-big** *milliseconds* [*bucket-size*]
Parameter Description	*milliseconds*: Indicates the refresh period of a token bucket, ranging from 0 to 2,147,483,647. The unit is millisecond. The default value is 100. If the value is 0, the sending rate of ICMPv6 error messages is not restricted. *bucket-size*: Indicates the number of tokens in a token bucket, ranging from 1 to 200. The default value is 10.
Command Mode	Global configuration mode
Usage Guide	To prevent DoS attacks, use the token bucket algorithm to restrict the sending rate of ICMPv6 error messages. If the length of an IPv6 packet to be forwarded exceeds the IPv6 MTU of the outbound interface, the router discards this IPv6 packet and sends back an ICMPv6 Packet Too Big message to the source IPv6 address. This error message is mainly used as part of the IPv6 PMTUD process. If other ICMPv6 error messages are excessive, ICMPv6 Packet Too Big messages cannot be sent, causing failure of IPv6 PMTUD. Therefore, it is recommended to restrict the sending rate of ICMPv6 Packet Too Big messages independently of other ICMPv6 error messages. Since the precision of the timer is 10 milliseconds, it is recommended to set the refresh period of a token bucket to an integer multiple of 10 milliseconds. If the refresh period of the token bucket is between 0 and 10, the actual refresh period is 10 milliseconds. For example, if the sending rate is set to 1 every 5 milliseconds, two error messages are sent every 10 milliseconds in actual situations. If the refresh period of the token bucket is not an integer multiple of 10 milliseconds, it is automatically converted to an integer multiple of 10 milliseconds. For example, if the sending rate is set to 3 every 15 milliseconds, two tokens are refreshed every 10 milliseconds in actual situations.

Configuring the Sending Rate of Other ICMPv6 Error Messages

Command	**ipv6 icmp error-interval** *milliseconds* [*bucket-size*]
Parameter Description	*milliseconds*: Indicates the refresh period of a token bucket, ranging from 0 to 2,147,483,647. The unit is millisecond. The default value is 100. If the value is 0, the sending rate of ICMPv6 error messages is not restricted. *bucket-size*: Indicates the number of tokens in a token bucket, ranging from 1 to 200. The default value is 10.
Command Mode	Global configuration mode
Usage Guide	To prevent DoS attacks, use the token bucket algorithm to restrict the sending rate of ICMPv6 error messages. Since the precision of the timer is 10 milliseconds, it is recommended to set the refresh period of a token bucket to an integer multiple of 10 milliseconds. If the refresh period of the token bucket is between 0 and 10, the actual refresh period is 10 milliseconds. For example, if the sending rate is set to 1 every 5 milliseconds, two error messages are sent every 10 milliseconds in actual situations. If the refresh period of the token bucket is not an integer multiple of 10 milliseconds, it is automatically converted to an integer multiple of 10 milliseconds. For example, if the sending rate is set to 3 every 15 milliseconds, two tokens are refreshed every 10 milliseconds in actual situations.

Configuration Example

Configuring the Sending Rate of ICMPv6 Error Messages

Configuration Steps	Set the sending rate of the ICMPv6 Packet Too Big message to 100 pps and that of other ICMPv6 error messages to 10 pps.
	FS(config)#ipv6 icmp error-interval too-big 1000 100 FS(config)#ipv6 icmp error-interval 1000 10
Verification	Run the **show running-config** command to check whether the configuration takes effect.

	FS#show running-config | include ipv6 icmp error-interval ipv6 icmp error-interval 1000 10 ipv6 icmp error-interval too-big 1000 100

3.4.6 Configuring the IPv6 Hop Limit

Configuration Effect

Configure the number of hops of a unicast packet to prevent the packet from being unlimitedly transmitted.

Configuration Steps

Configuring the IPv6 Hop Limit

- Optional.
- To modify the number of hops of a unicast packet, run the **ipv6 hop-limit value** command.

Verification

- Run the **show running-config** command to check whether the configuration is correct.
- Capture the IPv6 unicast packets sent by a host. The packet capture result shows that the hop-limit field value in the IPv6 header is the same as the configured hop limit.

Related Commands

Configuring the IPv6 Hop Limit

Command	**ipv6 hop-limit value**
Parameter Description	*value*: Indicates the number of hops of a unicast packet sent by the device. The value ranges from 1 to 255.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the IPv6 Hop Limit

Configuration Steps	Change the IPv6 hop limit of a device to 250.
	FS(config)#ipv6 hop-limit 250
Verification	Run the **show running-config** command to check whether the configuration takes effect.
	FS#show running-config ipv6 hop-limit 254

3.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the dynamically learned neighbors.	**clrear ipv6 neighbors** [*interface-id*]

Displaying

Description	Command
Displays IPv6 information of an interface.	**show ipv6 interface** [[*interface-id*] [ra-info]] [brief [*interface-id*]]
Displays neighbor information.	**show ipv6 neighbors** [**verbose**] [*interface-id*] [*ipv6-address*] [static]

Debugging

System resources are occupied when debugging information is output. Therefore, disable the debugging switch immediately after use.

Description	Command
Debugs ND entry learning.	**debug ipv6 nd**

4 Configuring DHCP

4.1 Overview

The Dynamic Host Configuration Protocol (DHCP) is a LAN protocol based on the User Datagram Protocol (UDP) for dynamically assigning reusable network resources, for example, IP addresses.

The DHCP works in Client/Server mode. A DHCP client sends a request message to a DHCP server to obtain an IP address and other configurations. When a DHCP client and a DHCP server are not in a same subnet, they need a DHCP relay to forward DHCP request and reply packets.

Protocols and Standards

- RFC2131: Dynamic Host Configuration Protocol
- RFC2132: DHCP Options and BOOTP Vendor Extensions
- RFC3046: DHCP Relay Agent Information Option

4.2 Applications

Application	Description
Providing DHCP Service in a LAN	Assigns IP addresses to clients in a LAN.
Enabling DHCP Client	Enable DHCP Client.
Deploying DHCP Relay in Wired Network	In a wired network, users from different network segments requests IP addresses.

4.2.1 Providing DHCP Service in a LAN

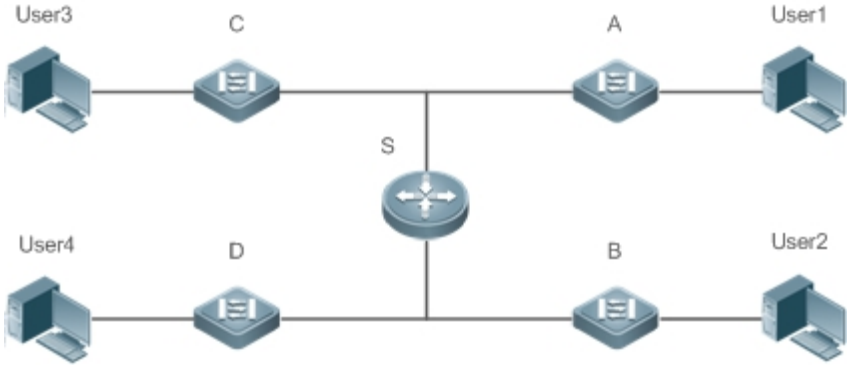
Scenario

Assign IP addresses to four users in a LAN.

For example, assign IP addresses to User 1, User 2, User 3 and User 4, as shown in the following figure.

- The four users are connected to Server S through A, B, C and D.

Figure 4- 1



Remarks	S is an egress gateway working as a DHCP server. A, B, C and D are access switches achieving layer-2 transparent transmission. User 1, User 2, User 3 and User 4 are LAN users.

Deployment

- Enable DHCP Server on S.
- Deploy layer-2 VLAN transparent transmission on A, B, C and D.
- User 1, User 2, User 3 and User 4 initiate DHCP client requests.

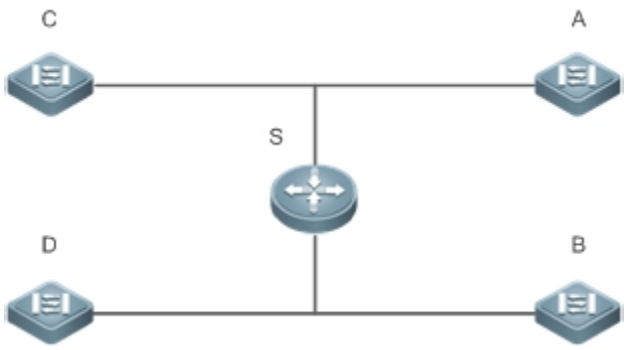
4.2.2 Enabling DHCP Client

Scenario

Access switches A, B, C and D in a LAN request server S to assign IP addresses.

For example, enable DHCP Client on the interfaces of A, B, C and D to request IP addresses, as shown in the following figure.

Figure 4- 2



Remarks	S is an egress gateway working as a DHCP server. A, B, C and D are access switches with DHCP Client enabled on the interfaces.

Deployment

- Enable DHCP Server on S.
- Enable DHCP Client on the interfaces of A, B, C and D.
-

4.2.3 Deploying DHCP Relay in Wired Network

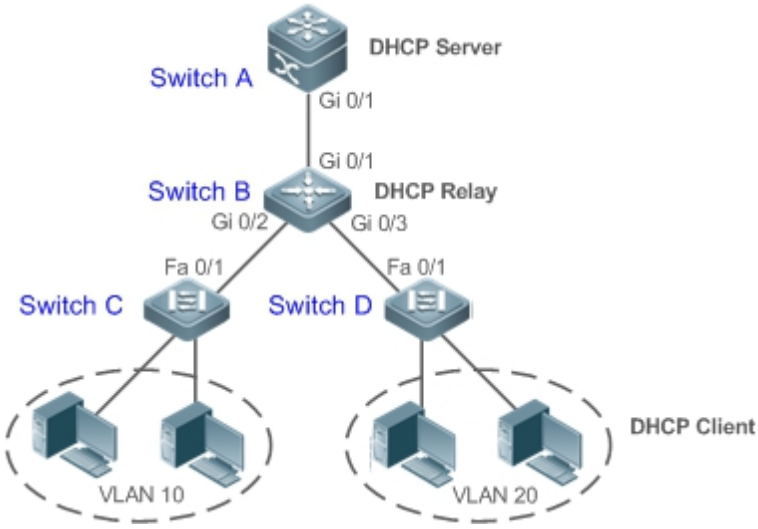
Scenario

As shown in the following figure, Switch C and Switch D are access devices for the users in VLAN 10 and VLAN 20 respectively. Switch B is a gateway, and Switch A a core device. The requirements are listed as follows:

Switch A works as a DHCP server to assign IP addresses of different network segments dynamically to users in different VLANs.

Users in VLAN 10 and VLAN 20 obtain IP addresses dynamically.

Figure 4- 3 DHCP Relay



Remarks	Switch C and Switch D are access devices. Switch B is a gateway. Switch A is a core device.

Deployment

- Configure layer-2 communication between Switch B and Switch C as well as between Switch B and Switch D.
- On Switch B, specify a DHCP server address and enable DHCP Relay.
- On Switch A, create DHCP address pools for VLAN 10 and VLAN 20 respectively, and enable DHCP Server.

4.2.4 Applying Administration Upon Initial Startup

Scenario

If there are no trunk ports, an administrator accesses the device via Web or Telnet. In this case, an administrated address is configured automatically. Then, after startup, the device obtains a dynamic IP address, which the administrator cannot know. Therefore, to make the dynamic address known to the administrator, the administrated address should be configured as the Secondary IP address

Deployment

- An initial startup refers to a startup on a device without the "config.text" file.

4.3 Features

Basic Concepts

DHCP Server

Based on the RFC 2131, FS DHCP server assigns IP addresses to clients and manages these IP addresses.

DHCP Client

DHCP Client enables a device to automatically obtain an IP address and configurations from a DHCP server.

DHCP Relay

When a DHCP client and a DHCP server are not in a same subnet, they need a DHCP relay to forward DHCP request and reply packets.

Lease

Lease is a period of time specified by a DHCP server for a client to use an assigned IP address. An IP address is active when leased to a client. Before a lease expires, a client needs to renew the lease through a server. When a lease expires or is deleted from a server, the lease becomes inactive.

Excluded Address

An excluded address is a specified IP address not assigned to a client by a DHCP server.

Address Pool

An address pool is a collection of IP addresses that a DHCP server may assign to clients.

Option Type

An option type is a parameter specified by a DHCP server when it provides lease service to a DHCP client. For example, a public option include the IP addresses of a default gateway (router), WINS server and a DNS server. DHCP server allows configuration of other options. Though most options are defined in the RFC 2132, you can add user-defined options.

Overview

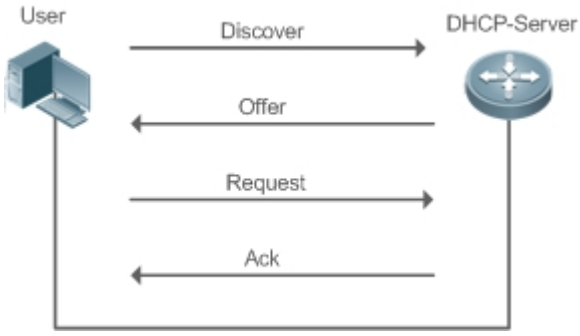
Feature	Description
DHCP Server	Enable DHCP Server on a device, and it may assign IP addresses dynamically and pushes configurations to DHCP clients.
DHCP Relay Agent	Enable DHCP Relay on a device, and it may forward DHCP request and reply packets across different network segments.
DHCP Client	Enable DHCP Client on a device, and it may obtain IP addresses and configurations automatically from a DHCP server.

4.3.1 DHCP Server

Working Principle

DHCP Working Principle

Figure 4- 7



A host requests an IP address through DHCP as follows:

1. A host broadcasts a DHCP discover packet to find DHCP servers in a network.

2. DHCP servers unicast/broadcast (based on the property of the host packet) DHCP offer packets to the host, containing an IP address, a MAC address, a domain name and a lease.
3. The host broadcasts a DHCP request packet to formally request an IP address.
4. A DHCP server sends a DHCP ACK unitcast packet to the host to acknowledge the request.

A DHCP client may receive DHCPOFFER packets from multiple DHCP servers, but usually it accepts only the first DHCPOFFER packet. Besides, the address specified in a DHCPOFFER packet is not necessarily assigned. Instead, it is retained by the DHCP server until a client sends a formal request.

To formally request an IP address, a client broadcasts a DHCPREQUEST packet so that all DHCP servers sending DHCPOFFER packets may receive the packet and release OFFER IP addresses.

If a DHCPOFFER packet contains invalid configuration parameters, a client will send a DHCPDECLINE packet to the server to decline the configuration.

During the negotiation, if a client does not respond to the DHCPOFFER packets in time, servers will send DHCPNAK packets to the client and the client will reinitiate the process.

During network construction, FS DHCP servers have the following features:

- Low cost. Usually the static IP address configuration costs more than DHCP configuration.
- Simplified configuration. Dynamic IP address assignment dramatically simplifies device configuration
- Centralized management. You can modify the configuration for multiple subnets by simply modifying the DHCP server configuration.

Address Pool

After a server receives a client's request packet, it chooses a valid address pool, determines an available IP address from the pool through PING, and pushes the pool and address configuration to the client. The lease information is saved locally for validity check upon lease renewal.

An address pool may carry various configuration parameters as follows:

- An IP address range, which is the range of IP addresses that are available.
- A gateway address. A maximum of 8 gateway addresses are supported.
- A DNS address. A maximum of 8 DNS addresses are supported.
- A lease period notifying clients of when to age an address and request a lease renewal.

VRRP Monitoring

In a Virtual Router Redundancy Protocol (VRRP) scenario, FS devices enabled with DHCP provide a command to monitor the VRRP status of the interface. To an interface configured with VRRP address and VRRP monitoring, a DHCP server only processes the DHCP clients' request packets from the interface in Master state, and other packets are discarded. If no VRRP address is configured, the DHCP server does not monitor the VRRP status, and all DHCP packets are processed. VRRP monitoring is configured on only layer-3 interfaces. It is disabled by default, namely, only the Master device processes the DHCP service.

ARP-Based Offline Detection

FS devices enabled with DHCP provide a command to enable ARP-based offline detection. After this function is enabled, a DHCP server will receive an ARP aging notification when a client gets offline, and start retrieving the client's address. If the client does not get online within a period of time (5 minutes by default), the DHCP server will retrieve the address and assign it to another client. If the client gets

online again, the address is still valid.

Adding Pseudo Server Detection

If a DHCP server is deployed illegally, a client interacts with this server while requesting an IP address and a wrong address will be assigned to the client. This server is a pseudo server. FS devices enabled with DHCP provides a command to enable pseudo server detection. After it is enabled, DHCP packets are checked for Option 54 (Server Identifier Option). If the content of Option 54 is different from the actual DHCP server identifier, the IP address of the pseudo server and port receiving the packets will be recorded. The pseudo server detection is only an after-event security function and cannot prevent an illegal DHCP server from assigning IP addresses to clients.

Related Configuration

Enabling DHCP Server Globally

- By default, DHCP Server is disabled.
- Run the **service dhcp** command to enable the DHCP Server.
- Run the **service dhcp** command globally to enable DHCP service.

Configuring Address Pool

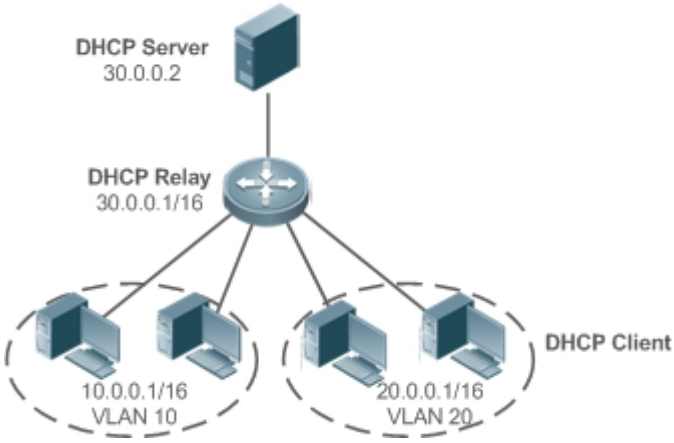
- By default, no address pool is configured.
- Run the **ip dhcp pool** command to configure an IP address range, a gateway and a DNS.
- If no address pool is configured, no addresses will be assigned.

4.3.2 DHCP Relay Agent

Working Principle

The destination IP address of DHCP request packets is 255.255.255.255, and these packets are forwarded within a subnet. To achieve IP address assignment across network segments, a DHCP relay agent is needed. The DHCP relay agent unicasts DHCP request packets to a DHCP server and forwards DHCP reply packets to a DCHP client. The DHCP relay agent serves as a repeater connecting a DHCP client and a DHCP server of different network segments by forwarding DHCP request packets and DHCP reply packets. The Client-Relay-Server mode achieves management of IP addresses across multiple network segments by only one DHCP server. See the following figure.

Figure 4- 8 DHCP Relay Scenario



VLAN 10 and VLAN 20 correspond to the segments 10.0.0.1/16 and 20.0.0.1/16 respectively. A DHCP server with IP address 30.0.0.2 is in segment 30.0.0.1/16. To achieve management of dynamic IP addresses in VLAN 10 and VLAN 20 by the DHCP server, you only need to

enable DHCP Relay on a gateway and configure IP address 30.0.0.2 for the DHCP server.

DHCP Relay Agent Information (Option 82)

As defined in RFC3046, an option can be added to indicate a DHCP client's network information when DHCP Relay is performed, so that a DHCP server may assign IP addresses of various privileges based on more accurate information.

Relay agent information option82: This scheme serves without correlation with other protocol modules. A DHCP relay agent forms an Option 82 based on the physical port receiving DHCP request packets and the MAC address of the device. The option format is shown in the following figure.

Figure 4- 10 Agent Circuit ID

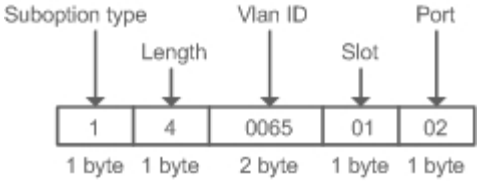
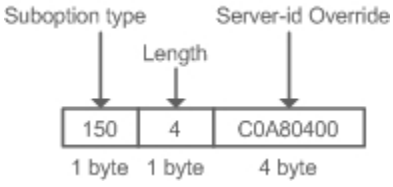


Figure 4- 15 Server-Identifier-Override



- Relay agent information option82: This scheme serves without correlation with other protocol modules. Compared with previous Option 82, this option supports user-defined content, which may change. By default, a DHCP relay agent forms Option 82 according to the information of the physical port receiving DHCP packets, device MAC address and device name. The option format is shown in the following figure.

Figure 4- 16 Option 82.1-circuit-id

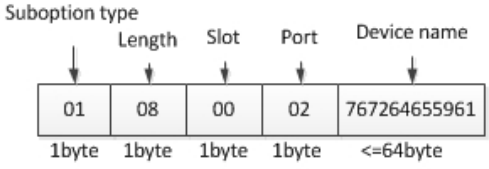
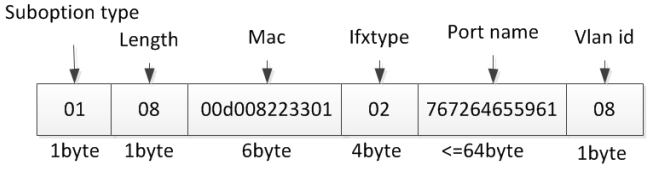


Figure 4- 17 Option82-remote-id



DHCP Relay Check Server-ID

In DHCP environment, multiple DHCP servers are deployed for a network, achieving server backup to ensure uninterrupted network operation. After this function is enabled, the DHCP request packet sent by a client contains a **server-id** option specifying a DHCP server. In alleviating the burden on servers in specific environments, you need to enable this function on a relay agent to send a packet to a

specified DHCP server rather than all DHCP servers.

DHCP Relay suppression

After you configure the **ip DHCP Relay suppression** command on an interface, DHCP request packets received on the interface will be filtered, and the other DHCP request packets will be forwarded.

Related Configuration

Enabling DHCP Relay

- By default, DHCP Relay is disabled.
- You may run the **service dhcp** command to enable DHCP Relay.
- You need to enable DHCP Relay before it works.

Configuring IP Address for DHCP Server

- By default, no IP address is configured for a DHCP server.
- You may run the **ip helper-address** command to configure an IP address for a DHCP server. The IP address can be configured globally or on a layer-3 interface. A maximum of 20 IP addresses can be configured for a DHCP server.
- When an interface receives a DHCP request packet, the DHCP server configuration on the interface prevails over that configured globally. If the interface is not configured with DHCP server addresses, the global configuration takes effect.

Enabling DHCP Option 82

- By default, DHCP Option 82 is disabled.
- You may run the **ip dhcp relay information option82** command to enable DHCP Option 82.

Enabling DHCP Relay Check Server-ID

- By default, DHCP Relay check server-id is disabled.
- You may run the **ip dhcp relay check server-id** command to enable DHCP Relay check server-id.

Enabling DHCP Relay Suppression

- By default, DHCP Relay suppression is disabled on all interfaces.
- You may run the **ip dhcp relay suppression** command to enable it on an interface.

4.3.3 DHCP Client

Working Principle

A DHCP client broadcasts a DHCP discover packet after entering the Init state. Then it may receive multiple DHCP offer packets. It chooses one of them and responds to the corresponding DHCP server. After that, it sends lease renewal request packets in the Renew and Rebind processes of an aging period to request lease renewal.

Related Configuration

Enabling DHCP Client on Interface

- By default, DHCP Client is disabled.

- In interface configuration mode, you may run the **ip address dhcp** command to enable DHCP Client.
- You need to enable DHCP Client to enable DHCP service.
- The configuration takes effect on a layer-3 interface, for example, an SVI or a routed port.
-

4.4 Configuration

Configuring DHCP Server

Configuration	Description and Command	
Configuring Dynamic IP Address	(Mandatory) It is used to enable DHCP Server to achieve dynamic IP address assignment.	
	service dhcp	Enables DHCP Server.
	ip dhcp pool	Configures an address pool.
	network	Configures the network number and subnet mask of a DHCP address pool.
	(Optional) It is used to configure the properties of an address pool.	
	default-router	Configures a default gateway of a client.
	lease	Configures an address lease.
	next-server	Configures a TFTP server address
	bootfile	Configures a boot file of a client.
	domain-name	Configures a domain name of a client.
	dns-server	Configures a domain name server.
	netbios-name-server	Configures a NetBIOS WINS server.
	netbios-node-type	Configures a NetBIOS node type on a client.
	lease-threshold	Configures an alarm threshold of an address pool.
	option	Configures a user-defined option.
	pool-status	Enables or disables an address pool.
Configuring Static IP Address	(Optional) It is used to statically assign an IP address to a client.	
	ip dhcp pool	Configures an address pool name and enters address pool configuration mode.
	host	Configures the IP address and subnet mask of a client host.
	hardware-address	Configures a client hardware address.
	client-identifier	Configures a unique client identifier.
	client-name	Configures a client name.
Configuring Global Properties of DHCP Server	(Optional) It is used to configure the properties of a DHCP server.	
	ip dhcp excluded-address	Configures an excluded IP address.
	ip dhcp force-send-nak	Configures Compulsory NAK reply by a DHCP server.

Configuration	Description and Command	
	ip dhcp monitor-vrrp-state	Configures VRRP status monitoring.
	ip dhcp ping packets	Configures ping times.
	ip dhcp ping timeout	Configures a ping timeout.
	ip dhcp server arp-detect	Configures a DHCP server to detect user offline.

Configuring DHCP Relay

Configuration	Description and Command	
Configuring Basic DHCP Relay Functions	(Mandatory) It is used to enable DHCP Relay.	
	service dhcp	Enables DHCP Relay.
	ip helper-address	Configures an IP Address of a DHCP Server.
Configuring DHCP Relay Option 82	(Optional) It is used to assign IP addresses of different privileges to clients in combination with the information of a physical port. This function cannot be used together with the **dhcp option dot1x** command.	
	ip dhcp relay information option82	Enables DHCP option82.
Configuring DHCP Relay Check Server-ID	(Optional) It is used to enable a DHCP Relay agent to send DHCP request packets only to a specified server.	
	ip dhcp relay check server-id	Enables a DHCP Relay agent to send DHCP request packets only to a specified server
Configuring DHCP Relay Suppression	(Optional) It is used to shield DHCP request packets on an interface.	
	ip dhcp relay suppression	Enables DHCP Relay Suppression.

Configuring DHCP Client

Configuration	Description and Command	
Configuring DHCP Client	(Mandatory) It is used to enable DHCP Client.	
	ip address dhcp	Enables an Ethernet interface, a PPP/HDLC-encapsulated or FR-encapsulated interface to obtain IP addresses through DHCP.

4.4.1 Configuring Dynamic IP Address

Configuration Effect

Provide all DHCP clients with DHCP service including assigning IP addresses and gateways.

Notes

A DHCP server and a DHCP relay share the **service dhcp** command, but a device cannot function as a DHCP server and relay at the same time. When a device is configured with a valid address pool, it acts as a server and forwards packets. Otherwise, it serves as a relay agent.

Configuration Steps

Enabling DHCP Server

- Mandatory. It achieves dynamic IP address assignment.
- Run the **service dhcp** command in global configuration mode.

Configuring Address Pool

- Mandatory. It is used to create an IP address pool.
- Run the **ip dhcp pool** command in global configuration mode.

Configuring Network Number and Subnet Mask of DHCP Address Pool

- Mandatory. It defines a range of dynamically assigned addresses.
- Run the **network** command in DHCP address pool configuration mode.

Configuring Default Gateway of Client

- Optional. It is used to configure a gateway address.
- Run the **default-router** command in DHCP address pool configuration mode.

Configuring Address Lease

- Optional. It is used to configure an IP address lease, which is 24h by default.
- Run the **lease** command in DHCP address pool configuration mode.

Configuring TFTP Server Address

- Optional. It is used to configure a TFTP server address.
- Run the **next-server** command in DHCP address pool configuration mode.

Configuring Domain Name of Client

- Optional. It is used to configure the domain name of a client.
- Run the **domain-name** command in DHCP address pool configuration mode.

Configuring DNS

- Optional. It is used to configure a DNS address.
- Run the **dns** command in DHCP address pool configuration mode.

Configuring NetBIOS WINS Server

- Optional. It is used to configure a NetBIOS WINS server address.
- Run the **netbios-name-server** command in DHCP address pool configuration mode.

Configuring NetBIOS Node Type on Client

- Optional. It is used to configure a NetBIOS node type.
- Run the **netbios-name-type** command in DHCP address pool configuration mode.

Configuring Alarm Threshold of Address Pool

- Optional. It is used to manage the number of leases. When a threshold (90% by default) is reached, an alarm will be printed.

- Run the **lease-threshold** command in DHCP address pool configuration mode.

Configuring User-Defined Option

- Optional. It is used to configure user-defined options.
- Run the **option** command in DHCP address pool configuration mode.

Enabling or Disabling Address Pool

- Optional. It is used to enable or disable an address pool. It is enabled by default.
- Run the **pool-status** command in DHCP address pool configuration mode.

Verification

Connect a DHCP client and a DHCP server.

- Check whether the client obtains configurations on the server.

Related Commands

Enabling DHCP Server

Command	**service dhcp**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Enable DHCP Server and DHCP Relay. A DHCP server and a DHCP relay share the **service dhcp** command. When a device is configured with a valid address pool, it acts as a server and forwards packets. Otherwise, it serves as a relay agent.

Configuring Address Pool

Command	**ip dhcp pool** *dhcp-pool*
Parameter Description	*pool-name*: Indicates the name of an address pool.
Command Mode	Global configuration mode
Usage Guide	Before assigning an IP address to a client, you need to configure an address pool name and enter DHCP address pool configuration mode.

Configuring Network Number and Subnet Mask of DHCP Address Pool

Command	**network** *network-number mask* [*low-ip-address high-ip-address*]
Parameter Description	*network-number:* Indicates the network number of an IP address pool. *mask:* Indicates the subnet mask of an IP address pool. If no subnet mask is defined, the natural subnet mask is applied.
Command Mode	DHCP address pool configuration mode

Usage Guide	To configure dynamic address assignment, you need to configure a network number and subnet mask of an address pool to provide a DHCP server with a range of addresses. The IP addresses in a pool are assigned in order. If an address is assigned or exists in the target network segment, the next address will be checked until a valid address is assigned. To avoid such failures, a network administrator needs to prevent path redundancy in network construction, for example, by adjusting physical links or network paths.

Configuring Default Gateway of Client

Command	**default-router** *address* [*address2…address8*]
Parameter Description	*address*: Indicates the IP address of a default gateway. Configure at least one IP address. *ip-address2…ip-address8:* (Optional) A maximum of 8 gateways can be configured.
Command Mode	DHCP address pool configuration mode
Usage Guide	Configure a default gateway of a client, and a server will push the gateway configuration to the client. The IP addresses of the default gateway and the client should be in a same network.

Configuring Address Lease

Command	**lease** {*days* [*hours]* [*minutes*] | **infinite**}
Parameter Description	*days:* Defines a lease in the unit of day. *hours:* (Optional) Defines a lease in the unit of hour. Please define *days* before *hours*. *minutes:* (Optional) Defines a lease in the unit of minute. Please define *days* and *hours* before *minutes*. **infinite:** Defines an unlimited lease.
Command Mode	DHCP address pool configuration mode
Usage Guide	The default lease of an IP address assigned by a DHCP server is 1 day. When a lease is expiring soon, a client needs to request a lease renewal. Otherwise the IP address cannot be used after the lease is expired.

Configures Boot File on Client

Command	**bootfile** *filename*
Parameter Description	*file-name*: Defines a boot file name.
Command Mode	DHCP address pool configuration mode
Usage Guide	A boot file is a bootable image file used when a client starts up. The file is usually an OS downloaded by a DHCP client.

Configuring Domain Name of Client

Command	**domain-name** *domain*
Parameter Description	*domain-name:* Defines a domain name of a DHCP client.
Command Mode	DHCP address pool configuration mode

Usage Guide	You may define a domain name for a client. When the client accesses network through the host name, the domain name will be added automatically to complete the host name.

Configuring DNS

Command	**dns-server** { *ip-address* [*ip-address2…ip-address8*] }
Parameter Description	*ip-address:* Defines an IP address of a DNS server. Configure at least one IP address. *ip-address2…ip-address8:* (Optional) A maximum of 8 DNS servers can be configured.
Command Mode	DHCP address pool configuration mode
Usage Guide	If a client accesses network resources through the domain name, you need to configure a DNS server to resolve the domain name.

Configuring NetBIOS WINS Server

Command	**netbios-name-server** *address* [*address2…address8*]
Parameter Description	*address:* Defines an IP address of a WINS server. Configure at least one IP address. *ip-address2…ip-address8:* (Optional) A maximum of 8 WINS servers can be configured.
Command Mode	DHCP address pool configuration mode
Usage Guide	WINS is a domain name service through which a Microsoft TCP/IP network resolves a NetNBIOS name to an IP address. A WINS server is a Windows NT server. When a WINS server starts, it receives a registration request from a WINS client. When the client shuts down, it sends a name release message, so that the computers in the WINS database and on the network are consistent.

Configuring NetBIOS Node Type on Client

Command	**netbios-node-type** *type*
Parameter Description	type: Defines a NetBIOS node type with one of the following approaches. 1. A hexadecimal number, ranging from 0 to FF. Only followings values are available. ● b-node ● p-node ● m-node ● 8 for h-node 2. A character string. ● b-node for a broadcast node; ● p-node for a peer-to-peer node; ● m-node for a mixed node; ● h-node for a hybrid mode.
Command Mode	DHCP address pool configuration mode
Usage Guide	There are four types of NetBIOS nodes of a Microsoft DHCP client. 1) A broadcast node. For such a node, NetBIOS name resolution is requested through broadcast.2) A peer-to-peer node. The client sends a resolution request to the WINS server. 3) A mixed node. The client broadcasts a resolution request and sends the resolution request to the WINS server.. 4) A hybrid node. The client sends a resolution request to the WINS server. If no reply is received, the client will broadcast the resolution request. By default, a Microsoft operating system is a broadcast or hybrid node. If no WINS server is configured, it is a broadcast node. Otherwise, it is a hybrid node.

Configuring User-Defined Option

Command	**option** *code* { **ascii** *string* | **hex** *string* | **ip** *ip-address* }
Parameter Description	*code:* Defines a DHCP option code. **ascii** *string:* Defines an ASCII character string. **hex** *string:* Defines a hexadecimal character string. **ip** *ip-address:* Defines an IP address.
Command Mode	DHCP address pool configuration mode
Usage Guide	The DHCP allows transmitting configuration information to a host via a TCP/IP network. DHCP packets contain the option field of definable content. A DHCP client should be able to receive a DHCP packet carrying at least 312 bytes option. Besides, the fixed data field in a DHCP packet is also called an option.

Enabling or Disabling Address Pool

Command	**pool-status** {**enable | disable**}
Parameter Description	**enable***:* Enables an address pool. **disable***:* Disable an address pool. It is enabled by default.
Command Mode	DHCP address pool configuration mode
Usage Guide	N/A

Configuration Example

Configuring Address Pool

Configuration Steps	● Define an address pool net172. ● The network segment is 172.16.1.0/24. ● The default gateway is 172.16.1.254. ● The address lease is 1 day. ● xcluded addresses range from 172.16.1.2 to 172.16.1.100.
	FS(config)# ip dhcp excluded-address 172.16.1.2 172.16.1.100 FS(dhcp-config)# ip dhcp pool net172 FS(dhcp-config)# network 172.16.1.0 255.255.255.0 FS(dhcp-config)# default-router 172.16.1.254 FS(dhcp-config)# lease 1
Verification	Run the **show run** command to display the configuration.
	FS(config)#show run | begin ip dhcp ip dhcp excluded-address 172.16.1.2 172.16.1.100 ip dhcp pool net172 network 172.16.1.0 255.255.255.0default-router 172.16.1.254

	lease 1

4.4.2 Configuring Static IP Address

Configuration Effect

Assign specific IP addresses and push configuration to specific DHCP clients.

Notes

N/A

Configuration Steps

Configuring Address Pool Name and Entering Address Pool Configuration Mode

- Mandatory. It is used to create an IP address pool.
- Run the **ip dhcp pool** command in global configuration mode.

Configuring IP Address and Subnet Mask of Client

- Mandatory. It is used to configure a static IP address and a subnet mask.
- Run the **host** command in DHCP address pool configuration mode.

Configuring Hardware Address of Client

- Optional. It is used to configure a MAC address.
- Run the **hardware** command in DHCP address pool configuration mode.

Configures Unique Client Identifier

- Optional. It is used to configure a static user identifier (UID).
- Run the **client-identifier** command in DHCP address pool configuration mode.

Configuring Client Name

- Optional. It is used to configure a static client name.
- Run the **host-name** command in DHCP address pool configuration mode.

Verification

Check whether the client obtains the IP address when it is online.

Related Commands

Configuring Address Pool

Command	**ip dhcp pool** *dhcp-pool*
Parameter Description	*pool-name*: Indicates the name of an address pool.
Command Mode	Global configuration mode

Usage Guide	Before assigning an IP address to a client, you need to configure an address pool name and enter address pool configuration mode.

Manual IP Address Binding

Command	**host** *ip-address* [*netmask*] **client-identifier** *unique-identifier* **client-name** *name*
Parameter Description	*ip-address:* Defines the IP address of a DHCP client. *netmask:* Defines the subnet mask of a DHCP client. *unique-identifier:* Defines the hardware address (for example, aabb.bbbb.bb88) and identifier (for example, 01aa.bbbb.bbbb.88) of a DHCP client. *name:* (Optional) It defines a client name using ASCII characters. The name excludes a domain name. For example, name a host **mary** rather than **mary.fs.com**.
Command Mode	DHCP address pool configuration mode
Usage Guide	Address binding means mapping between an IP address and a client's MAC address. There are two kind of address binding. 1) Manual binding. Manual binding can be deemed as a special DHCP address pool with only one address. 2) Dynamic binding. A DHCP server dynamically assigns an IP address from a pool to a client when it receives a DHCP request, creating mapping between the IP address and the client's MAC address. To configure manual binding, you need to define a host pool and then specify a DHCP client's IP address and hardware address or identifier. A hardware address is a MAC address. A client identifier includes a network medium type and a MAC address. A Microsoft client is usually identified by a client identifier rather than a MAC address. For the codes of medium types, refer to the *Address Resolution Protocol Parameters* section in the RFC 1700. The Ethernet type is **01**.

Configuration Example

Dynamic IP Address Pool

Configuration Steps	● Configure address pool VLAN 1 with IP address 20.1.1.0 and subnet mask 255.255.255.0. ● The default gateway is 20.1.1.1. ● The lease time is 1 day.
	FS(config)# ip dhcp pool vlan1 FS(dhcp-config)# network 20.1.1.0 255.255.255.0 FS(dhcp-config)# default-router 20.1.1.1 FS(dhcp-config)# lease 1 0 0
Verification	● Run the **show run** command to display the configuration.
	FS(config)#show run | begin ip dhcp ip dhcp pool vlan1 network 20.1.1.0 255.255.255.0 default-router 20.1.1.1 lease 1 0 0

Manual Binding

Configuration Steps	● The host address is 172.16.1.101 and the subnet mask is 255.255.255.0. ● The host name is Billy.fs.com. ● The default gateway is 172.16.1.254. ● The MAC address is 00d0.df34.32a3.
	FS(config)# ip dhcp pool Billy FS(dhcp-config)# host 172.16.1.101 255.255.255.0 FS(dhcp-config)# client-name Billy FS(dhcp-config)# hardware-address 00d0.df34.32a3 Ethernet FS(dhcp-config)# default-router 172.16.1.254
Verification	Run the **show run** command to display the configuration.
	FS(config)#show run | begin ip dhcp ip dhcp pool Billy host 172.16.1.101 255.255.255.0 client-name Billy hardware-address 00d0.df34.32a3 Ethernet default-router 172.16.1.254

4.4.3 Configuring Global Properties of DHCP Server

Configuration Effect

Enable a server with specific functions, for example, ping and compulsory NAK.

Notes

Configuring the command may cause exceptions on other servers.

Configuration Steps

Configuring Excluded IP Address

- Optional. Configure some addresses or address ranges as unavailable.
- Run the **ip dhcp excluded-address** command in global configuration mode.

Configuring Compulsory NAK Reply

- Optional. A server replies to a wrong address request with a NAK packet.
- Run the **ip dhcp force-send-nak** command in global configuration mode.

Configuring VRRP Status Monitoring

- Optional. After configuration, DHCP packets are processed by the Master server.

- Run the **ip dhcp monitor-vrrp-state** command in global configuration mode.

Configuring Ping Times

- Optional. Check the address reachability with the **ping** command. The default is 2.
- Run the **ip dhcp ping packet** command in global configuration mode.

Configuring Ping Timeout

- Optional. Check the address reachability with the **ping** command. The default is 500 ms.
- Run the **ip dhcp ping timeout** command in global configuration mode.

Detecting User Offline Detection

- Configure a DHCP server to detect whether the client is offline or not. If a client does not get online after being offline for a period, the address assigned to the client will be retrieved.
- Run the **ip dhcp server arp-detect** command in global configuration mode.

Verification

Run the **dhcp-server** command, and check the configuration during address assignment.

Related Commands

Configuring Excluded IP Address

Command	**ip dhcp excluded-address** *low-ip-address* [*high-ip-address*]
Parameter Description	*low-ip-address:* Indicates a start IP address. *high-ip-address:* Indicates an end IP address.
Command Mode	Global configuration mode
Usage Guide	Unless otherwise specified, a DHCP server assigns all the addresses from an IP address pool to DHCP clients. To reserve some addresses (e.g., addresses already assigned to the server or devices), you need to configure these addresses as excluded addresses. To configure a DHCP server, it is recommended to configure excluded addresses to avoid address conflict and shorten detection time during address assignment.

Configuring Compulsory NAK Reply

Command	**ip dhcp force-send-nak**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	The server sends a NAK packet only when it finds the client's lease record. When a DHCP client crosses the network, a DHCP server cannot find lease record of the client and will not reply with a NAK packet. The client sends request packets continually before obtaining an IP address again after timeout. Consequently, it takes a long to obtain an IP address. This also occurs when a DHCP server loses a lease after restart and a client requests lease renewal. In this case, you may configure a command to force the DHCP server to reply with a NAK packet even though it cannot find the lease record so that the client may obtain an IP address rapidly. Please note that the command is disabled by

	default. To enable it, only one DHCP server can be configured in a broadcast domain.

Configuring Ping Times

Command	**ip dhcp ping packets** [*number*]
Parameter Description	*number:* (Optional) Ranges from 0 to 10. 0 indicates the ping function is disabled. The default is two pings.
Command Mode	Global configuration mode
Usage Guide	By default, when a DHCP server assigns an IP address from a pool, it runs the **Ping** command twice (one packet per time). If there is no reply, the server takes the address as idle and assigns it to a client. If there is a reply, the server takes the address as occupied and assigns another address.

Configuring Ping Timeout

Command	**ip dhcp ping timeout** *milliseconds*
Parameter Description	*milli-seconds:* Indicates the time that it takes for a DHCP server to wait for a ping reply. The value ranges from 100 ms to 10,000 ms.
Command Mode	Global configuration mode
Usage Guide	By default, if a DHCP server receives no Ping reply within 500 ms, the IP address is available. You may adjust the ping timeout to change the time for a server to wait for a reply.

Configuring ARP-Based Offline Detection

Command	**ip dhcp server arp-detect**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	By default, DHCP server does not detect whether a client is offline or not based on ARP. After configuration, a DHCP server may perform the detection. If a client does not get online again after a period (5 minutes by default), a DHCP server retrieves the address assigned to the client.

Configuration Example

Configuring Ping

Configuration Steps	• Set ping times to 5. • Set ping timeout to 800ms.
	FS(config)# ip dhcp ping packet 5 FS(config)# ip dhcp ping timeout 800
Verification	Run the **show run** command to display the configuration.
	FS(config)#show run | begin ip dhcp ip dhcp ping packet 5

	ip dhcp ping timeout 800

Configuring Excluded IP Address

Configuration Steps	● Configure the excluded IP address from 192.168.0.0 to 192.168.255.255.
	FS(config)# ip dhcp excluded-address 192.168.0.0 192.168.255.255
Verification	Run the **show run** command to display the configuration.
	FS(config)#show run | begin ip dhcp ip dhcp excluded-address 192.168.0.0 192.168.255.255

4.4.4 Configuring Basic DHCP Relay Functions

Configuration Effect

- Deploy dynamic IP management in Client–Relay–Server mode to achieve communication between a DHCP client and a DHCP server, which are in different network segments.

Notes

- To enable DHCP Relay, you need to configure IPv4 unicast routing in a network.

Configuration Steps

Enabling DHCP Relay

- Mandatory.
- Unless otherwise specified, you need to enable DHCP Relay on a device.

Configuring IP Address for DHCP Server

- Mandatory.
- You need to configure an IP address for a DHCP server.

Verification

- Check whether a client obtains an IP address through DHCP Relay.

Related Commands

Enabling DHCP Relay

Command	**service dhcp**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring IP Address for DHCP Server

Command	**ip helper-address { cycle-mode | *A.B.C.D* }**
Parameter Description	*cycle-mode*: Indicates that DHCP request packets are forwarded to all DHCP servers. *A.B.C.D*: Indicates the IP address of a server.
Command Mode	Global configuration mode/interface configuration mode
Usage Guide	You may configure the function on a layer-3 interface, such as a routed port, a L3 AP port, SVI and loopback interface. The configured interface must be accessible via IPv4 unicast routing.

Configuration Example

Configuring DHCP Relay in Wired Connection

Scenario **Figure 4- 18**	
Configuration Steps	● Enable a client with DHCP to obtain an IP address. ● Enable the DHCP Relay function on a DHCP relay agent. ● Configure DHCP Server.
A	Enable a client with DHCP to obtain an IP address.
B	Enable DHCP Relay. FS(config)# service dhcp Configure a global IP address of a DHCP server. FS(config)# ip helper-address 172.2.2.1 Configure an IP address for the port connected to the client. FS(config)# interface gigabitEthernet 0/1 FS(config-if)# ip address 192.1.1.1 255.255.255.0 Configure an IP address for the port connected to the server. FS(config)# interface gigabitEthernet 0/2 FS(config-if-gigabitEthernet 0/2)# ip address 172.2.2.2 255.255.255.0
C	Enable DHCP Server. FS(config)# service dhcp Configure an address pool. FS(config)# ip dhcp pool relay FS (dhcp-config)#network 192.1.1.0 255.255.255.0 FS (dhcp-config)#default-router 192.1.1.1 Configure an IP address for the port connected to the relay agent.

	FS(config)# interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/2)# ip address 172.2.2.1 255.255.255.0
Verification	Check whether the client obtains an IP address. ● Check whether the client obtains an IP address. ● Check the DHCP Relay configuration.
A	The user device obtains an IP address.
B	After login to the DHCP relay agent, run the **show running-config** command in privileged EXEC mode to display DHCP Relay configuration. FS# show running-config service dhcp ip helper-address 172.2.2.1 ! interface GigabitEthernet 0/1 ip address 192.1.1.1 255.255.255.0 ! interface GigabitEthernet 0/2 ip address 172.2.2.2 255.255.255.0 !

Common Errors

- IPv4 unicast routing configuration is incorrect.
- DHCP Relay is disabled.
- No routing between DHCP relay agent and DHCP server is configured.
- No IP address is configured for the DHCP server.

4.4.5 Configuring DHCP Relay Option 82

Configuration Effect

- Through a DHCP relay agent, a server may assign IP addresses of different privileges to the clients more accurately based on the option information.

Notes

- You need to enable the DHCP Relay function.

Configuration Steps

↘ **Enabling Basic DHCP Relay Functions**

- Mandatory.

- Unless otherwise specified, you need to enable DHCP Relay on a device.

Enables DHCP Option82

- By default, DHCP Option 82 is disabled.
- You may run the **ip dhcp relay information option82** command to enable or disable DHCP Option 82.

Verification

- Check whether the client obtains an IP address based on Option 82.

Related Commands

Enabling DHCP Option 82

Command	**ip dhcp relay information option82**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Enabling DHCP Option 82

Configuration Steps	• Enable DHCP Option 82. • Configure sub-option commands.
	FS(config)# ip dhcp relay information option82
Verification	After login to the DHCP relay agent, run the **show running-config** command in privileged EXEC mode to display DHCP Relay configuration.
	FS#show ru | incl ip dhcp relay ip dhcp relay information option82

Common Errors

- Basic DHCP Relay functions are not configured.

4.4.6 Configuring DHCP Relay Check Server-ID

Configuration Effect

- After you configure the **ip dhcp relay check server-id**, a DHCP Relay agent will forward DHCP request packets only to the server specified by the **option server-id** command. Otherwise, they are forwarded to all DHCP servers.

Notes

- You need to enable basic DHCP Relay functions.

Configuration Steps

Enabling DHCP Relay Check Server-ID

- By default, DHCP Relay check server-id is disabled.
- You may run the **ip dhcp relay check server-id** command to enable DHCP Relay check server-id.

Verification

Check whether a DHCP Relay agent sends DHCP request packets only to the server specified by the **option server-id** command.

Related Commands

Configuring DHCP Relay Check Server-ID

Command	**ip dhcp relay check server-id**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring DHCP Relay Check Server-ID

Configuration Steps	• Enable DHCP Relay. • Enable **DHCP Relay check server-id** on an interface.
	FS# configure terminal FS(config)# ip dhcp relay check server-id
Verification	After login to the DHCP relay agent, run the **show running-config** command in privileged EXEC mode to display DHCP Relay configuration.
	FS# show running-config | include check server-id ip dhcp relay check server-id FS#

Common Errors

- Basic DHCP Relay functions are not configured.

4.4.7 Configuring DHCP Relay Suppression

Configuration Effect

- After you configure the **ip DHCP Relay suppression** command on an interface, DHCP request packets received on the interface will be filtered, and the other DHCP requests will be forwarded.

Notes

- You need to enable basic DHCP Relay functions.

Configuration Steps

Enabling DHCP Relay Suppression

By default, DHCP Relay suppression is disabled on all interfaces.

You may run the **ip dhcp relay suppression** command to enable DHCP Relay suppression.

Verification

- Check whether the DHCP request packets received on the interface are filtered.

Related Commands

Configuring DHCP Relay Suppression

Command	**ip dhcp relay suppression**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring DHCP Relay Suppression

Configuration Steps	• Configure basic DHCP Relay functions. • Configure DHCP Relay suppression on an interface.
	FS# configure terminal FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ip dhcp relay suppression FS(config-if-GigabitEthernet 0/1)#end FS#
Verification	After login to the DHCP relay agent, run the **show running-config** command in privileged EXEC mode to display DHCP Relay configuration.
	FS# show running-config | include relay suppression ip dhcp relay suppression FS#

Common Errors

Basic DHCP Relay functions are not configured.

4.4.8 Configuring DHCP Client

Configuration Effect

Enable DHCP Client on a device so that it obtains IP addresses and configurations dynamically.

Notes

FS products support DHCP Client configuration on Ethernet, FR, PPP and HDLC interfaces.

Configuration Steps

Run the **ip address dhcp** command on an interface.

Verification

Check whether the interface obtains an IP address.

Related Commands

Configuring DHCP Client

Command	**ip address dhcp**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	● FS products support dynamic IP address obtainment by an Ethernet interface. ● FS products support dynamic IP address obtainment by a PPP-encapsulated interface. ● FS products support dynamic IP address obtainment by an FR-encapsulated interface. ● FS products support dynamic IP address obtainment by an HDLC-encapsulated interface.

Configuration Example

Configuring DHCP Client

Configuration Steps	1: Enable port FastEthernet 0/0 with DHCP to obtain an IP address.
	FS(config)# interface FastEthernet0/0 FS(config-if-FastEthernet 0/0)#ip address dhcp
Verification	1: Run the **show run** command to display the configuration.
	FS(config)#show run | begin ip address dhcp ip address dhcp

4.5 Monitoring

Clearing

Running the clear commands may lose vital information and interrupt services.

Description	Command
Clears DHCP address binding.	**clear ip dhcp binding** { *address* | *}
Clears DHCP address conflict.	**clear ip dhcp conflict** { *address* | *}
Clears statistics of a DHCP server.	**clear ip dhcp server statistics**
Clears statistics of a DHCP relay.	**clear ip dhcp relay statistics**
Clears statistics of DHCP server performance.	**clear ip dhcp server rate**

Displaying

Description	Command
Displays DHCP lease.	**show dhcp lease**
Displays DHCP sockets.	**show ip dhcp socket**
Displays assigned IP addresses.	**show ip dhcp binding**
Displays created address pools.	**show ip dhcp pool**
Displays statistics of DHCP Server.	**show ip dhcp server statistic**
Displays statistics of DHCP Relay.	**show ip dhcp relay statistic**
Displays conflicted addresses.	**show ip dhcp conflict**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs DHCP agent.	**debug ip dhcp server agent**
Debugs DHCP hot backup.	**debug ip dhcp server ha**
Debugs DHCP address pools.	**debug ip dhcp server pool**
Debugs DHCP VRRP.	**debug ip dhcp server vrrp**
Debugs all DHCP servers.	**debug ip dhcp server all**
Debugs DHCP packets.	**debug ip dhcp client**
Debugs DHCP Relay events.	**debug ip dhcp relay**

5 Configuring DHCPv6

5.1 Overview

The Dynamic Host Configuration Protocol for IPv6 (DHCPv6) is a protocol that allows a DHCP server to transfer configurations (such as IPv6 addresses) to IPv6 nodes.

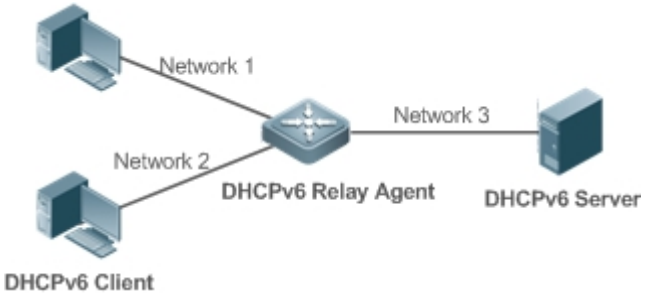
As compared with other IPv6 address allocation methods, such as manual configuration and stateless automatic address configuration, DHCPv6 provides the address allocation, prefix delegation, and configuration parameter allocation.

- DHCPv6 is a stateful protocol for automatically configuring addresses and flexibly adding and reusing network addresses, which can record allocated addresses and enhance network manageability.
- By using the prefix delegation of DHCPv6, uplink network devices can allocate address prefixes to downlink network devices, which implements flexible station-level automatic configuration and flexible control of station address space.
- The DHCPv6 configuration parameter allocation solves the problem that parameters cannot be obtained through a stateless automatic address configuration protocol and allocates DNS server addresses and domain names to hosts.

DHCPv6 is a protocol based on the client/server model. A DHCPv6 client is used to obtain various configurations whereas a DHCPv6 server is used to provide various configurations. If the DHCPv6 client and DHCPv6 server are not on the same network link (the same network segment), they can interact with each other by using a DHCPv6 relay agent.

The DHCPv6 client usually discovers the DHCPv6 server by reserving multicast addresses within a link; therefore, the DHCPv6 client and DHCPv6 server must be able to directly communicate with each other, that is, they must be deployed within the same link. This may cause management inconvenience, economic waste (a DHCPv6 server is deployed for each subnet) and upgrade inconvenience. The DHCPv6 relay agent function can solve these problems by enabling a DHCPv6 client to send packets to a DHCPv6 server on a different link. The DHCP relay agent is often deployed within the link where a DHCPv6 client resides and is used to relay interaction packets between the DHCPv6 client and DHCPv6 server. The DHCP relay agent is transparent to the DHCPv6 client.

Figure 5- 1



Protocols and Standards

- RFC3315: Dynamic Host Configuration Protocol for IPv6
- RFC3633: IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) Version 6
- RFC3646: DNS Configuration Options for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)
- RFC3736: Stateless DHCP Service for IPv6
- RFC5417: Control And Provisioning of Wireless Access Points (CAPWAP) Access Controller DHCP Option

5.2 Applications

Application	Description
Requesting/Allocating Addresses and Configuration Parameters	A DHCPv6 client requests addresses from a DHCPv6 server. The DHCPv6 server allocates addresses and configuration parameters to the DHCPv6 client.
Requesting/Allocating Prefixes	The DHCPv6 client requests a prefix from the DHCPv6 server. The DHCPv6 server allocates a prefix to the DHCPv6 client and then the DHCPv6 client configures IPv6 addresses by using this prefix.
Relay Service	The DHCPv6 relay is used to enable communication between the DHCPv6 client and DHCPv6 server on different links.

5.2.1 Requesting/Allocating Addresses and Configuration Parameters

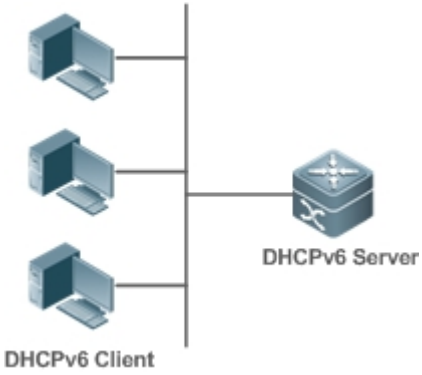
Scenario

In a subnet, a DHCPv6 client requests addresses from a DHCPv6 server. The DHCPv6 server allocates addresses and configuration parameters to the DHCPv6 client.

As shown in Figure 5- 2:

- The DHCPv6 server is configured with IPv6 addresses, DNS servers, domain names and other configuration parameters to be allocated.
- A host works as a DHCPv6 client to request an IPv6 address from the DHCPv6 server. After receiving the request, the DHCPv6 server selects an available address and allocates the address to the host.
- The host can also request a DNS server, domain name and other configuration parameters from the DHCPv6 server.

Figure 5- 2



Deployment

- Run the DHCPv6 client on a host in the subnet to obtain an IPv6 address and other parameters.
- Run the DHCPv6 server on a device and configure the IPv6 address and other parameters to allocate the IPv6 address and parameters.

5.2.2 Requesting/Allocating Prefixes

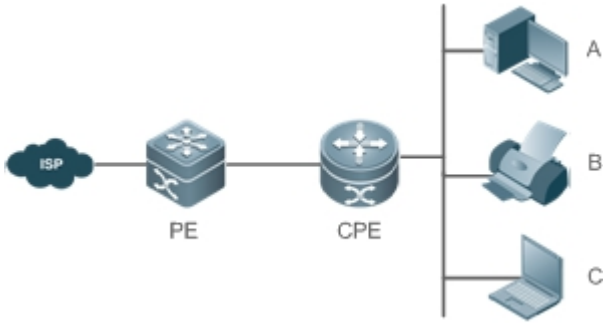
Scenario

As shown in Figure 5- 3, an uplink device (PE) allocates an IPv6 address prefix for a downlink device (CPE). The CPE generates a new address prefix for the internal subnet based on the obtained prefix. Hosts in the internal subnet of the CPE are configured with addresses

through Router Advertisement (RA) by using the new address prefix.

- The PE provides the prefix delegation service as a DHCPv6 server.
- The CPE requests an address prefix from the PE as a DHCPv6 client. After obtaining the address prefix, the CPE generates a new address prefix for the internal subnet and sends an RA message to hosts in the internal subnet.
- The hosts in the internal subnet where CPE resides configure their addresses based on the RA message sent by the CPE.

Figure 5- 3



Remarks	The Provider Edge (PE) works as a DHCPv6 server for providing prefixes and is also called a delegating router. The Customer Premises Equipment (CPE) works as a DHCPv6 client for requesting prefixes and is also called a requesting router. A, B and C are various hosts.

Deployment

- Run the DHCPv6 server on the PE to implement the prefix delegation service.
- Run the DHCPv6 client on the CPE to obtain address prefixes.
- Deploy IPv6 ND between the CPE and the hosts to configure the host addresses in the subnet through RA.

5.2.3 Relay Service

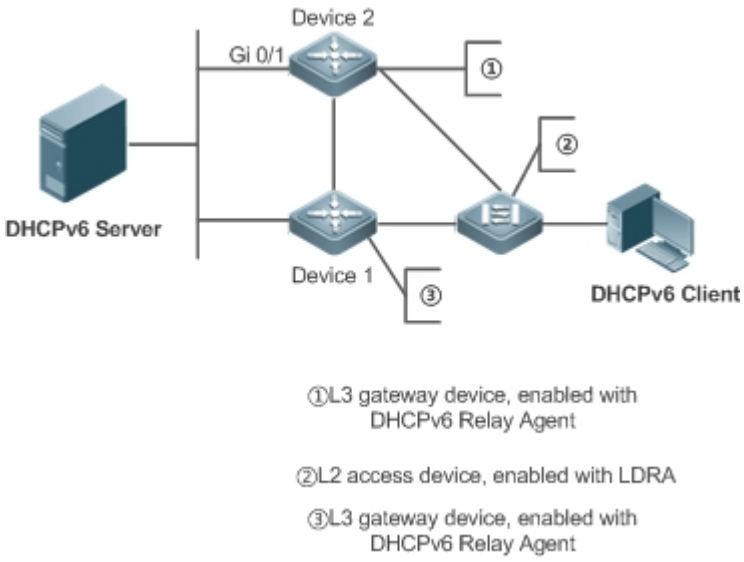
Scenario

The DHCPv6 relay agent provides the relay service for the DHCPv6 client ad DHCPv6 server on different links to enable communication between them.

As shown in Figure 5- 4:

- Device 1 is enabled with the DHCPv6 relay agent and destined to 3001::2.
- Device 2 wants to forward packets to other servers through a next-level relay service. Enable the DHCPv6 relay agent on Device 2, set the destination address to FF02::1:2 (all servers and Relay multicast addresses) and specify the egress interface as the layer-3 interface gi 0/1.

Figure 5- 4



Deployment

- Enable the DHCPv6 relay agent on device 1 and specify the address as 3000::1.
- Enable the DHCPv6 relay agent on device 2 and specify the address as FF02::1:2.

5.3 Features

Basic Concept

↘ DUID

The DHCP Unique Identifier (DUID) identifies a DHCPv6 device. As defined in RFC3315, each DHCPv6 device (DHCPv6 client, relay or server) must have a DUID, which is used for mutual authentication during DHCPv6 message exchange.

RFC3315 defines three types of DUIDs:

- DUID Based on Link-Layer address plus Time (DUID-LLT).
- DUID Assigned by Vendor Based on Enterprise Number (DUID-EN).
- Link-Layer address (DUID-LL).

FS DHCPv6 devices use DUID-LLs. The structure of a DUID-LL is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               3               |    hardware type (16 bits)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.             link-layer address (variable length)              .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The values of *DUID-LL*, *Hardware type*, and *Link-layer address* are 0x0003, 0x0001 (indicating the Ethernet), and MAC address of a device respectively.

↘ Identity Association (IA)

A DHCPv6 server allocates IAs to DHCPv6 clients. Each IA is uniquely identified by an identity association identifier (IAID). IAIDs are generated by DHCPv6 clients. A one-to-one mapping is established between IAs and clients. An IA may contain several addresses, which can be allocated by the client to other interfaces. An IA may contain one of the following types of addresses:

- Non-temporary Addresses (NAs), namely, globally unique addresses.
- Temporary Addresses (TAs), which are hardly used.
- Prefix Delegation (PD).

Based on the address type, IAs are classified into IA_NA, IA_TA, and IA_PD (three IA-Types). FS DHCPv6 servers support only IA_NA and IA_PD.

Binding

A DHCPv6 binding is a manageable address information structure. The address binding data on a DHCPv6 server records the IA and other configurations of every client. A client can request multiple bindings. The address binding data on a server is present in the form of an address binding table with DUID, IA-Type and IAID as the indexes. A binding containing configurations uses DUID as the index.

DHCPv6 Conflict

When an address allocated by a DHCPv6 client is in conflict, the client sends a Decline packet to notify the DHCPv6 server that the address is rebound. Then, the server adds the address to the address conflict queue. The server will not allocate the addresses in the address conflict queue. The server supports viewing and clearing of address information in the address conflict queue.

Packet Type

RFC3315 stipulates that DHCPv6 uses UDP ports 546 and 547 for packet exchange. Specifically, a DHCPv6 client uses port 546 for receiving packets, while a DHCPv6 server and DHCPv6 relay agent use port 547 for receiving packets. RFC3315 defines the following types of packets that can be exchanged among the DHCPv6 server, client, and relay agent:

- Packets that may be sent by a DHCPv6 client to a DHCPv6 server include Solicit, Request, Confirm, Renew, Rebind, Release, Decline, and Information-request.
- Packets that may be sent by a DHCPv6 server to a DHCPv6 client include Advertise, Reply, and Reconfigure.
- Packets that may be sent by a DHCPv6 relay agent to another DHCPv6 relay agent or a DHCPv6 server include Relay-forward.
- Packets that may be sent by a DHCPv6 relay agent to another DHCPv6 relay agent or a DHCPv6 server include Relay-reply.

FS DHCPv6 servers do not support the Reconfigure packet.

FS DHCPv6 clients do not support the Confirm and Reconfigure packets.

Overview

Feature	Description
Requesting/Allocating Addresses	Dynamically obtains/allocates IPv6 addresses in a network in the client/server mode.
Requesting/Allocating Prefixes	Dynamically obtains/allocates IPv6 prefixes in a network in the client/server mode.
Stateless Service	Provides stateless configuration service for hosts in a network.
Relay Service	Provides the DHCPv6 server service for hosts in different networks by using the relay service.

5.3.1 Requesting/Allocating Addresses

A DHCPv6 client can request IPv6 addresses from a DHCPv6 server.

After being configured with available addresses, a DHCPv6 server can provide IPv6 addresses to hosts in the network, record the allocated addresses and improve the network manageability.

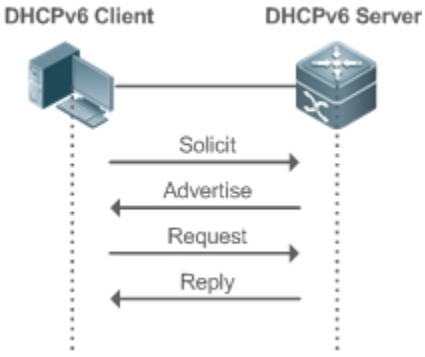
Working Principle

Network hosts serve as DHCPv6 clients and DHCPv6 servers to implement address allocation, update, confirmation, release and other operations through message exchange.

Four-Message Exchange

Figure 5-5 shows the four-message exchange process.

Figure 5- 5

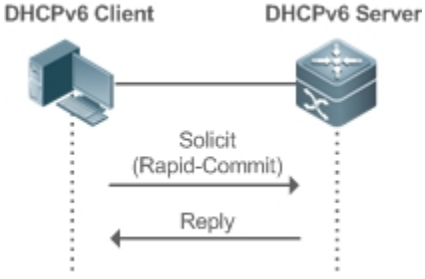


- A DHCPv6 client sends a Solicit message whose destination address is FF02::1:2 and destination port number is 547 within the local link to request address, prefix and configuration parameter allocation. All DHCPv6 servers or DHCPv6 relay agents within the link will receive the Solicit message.
- After receiving the Solicit message, a DHCPv6 server will send an Advertise message in the unicast mode if it can provide the information requested in the Solicit message. The Advertise message includes the address, prefix and configuration parameters.
- The DHCPv6 client may receive the Advertise message from multiple DHCPv6 servers. After selecting the most suitable DHCPv6 server, the DHCPv6 client sends a Request message whose destination address is FF02::1:2 and destination port number is 547 to request address, prefix and configuration parameter allocation.
- After receiving the Request message, the DHCPv6 server creates a binding locally and sends a Reply message in the unicast mode. The Reply message includes the address, prefix and configuration parameters that the DHCPv6 server will allocate to the DHCPv6 client. The DHCPv6 client obtains address, prefix or configuration parameters based on the information in the Reply message.

Two-Message Exchange

Two-message exchange can be used to complete address, prefix and parameter configuration for DHCPv6 clients more quickly.

Figure 5- 6

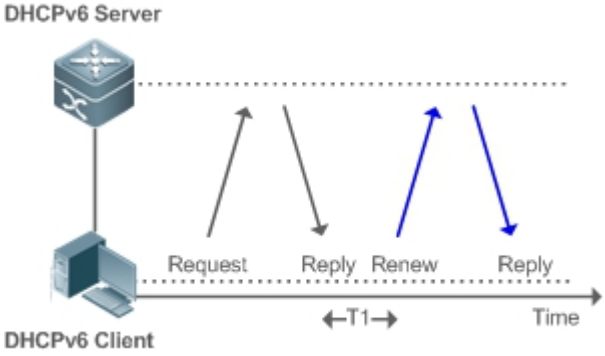


- A DHCPv6 client sends a Solicit message whose destination address is FF02::1:2 and destination port number is 547 within the local link to request address, prefix and configuration parameter allocation. The Solicit message contains Rapid Commit.
- If a DHCPv6 server supports the Rapid Commit option, the DHCPv6 server creates a binding locally and sends a Reply message in the unicast mode. The Reply message includes the address, prefix and configuration parameters to be allocated to the DHCPv6 client. The DHCPv6 client completes configuration based on the information in the Reply message.

Update and Rebinding

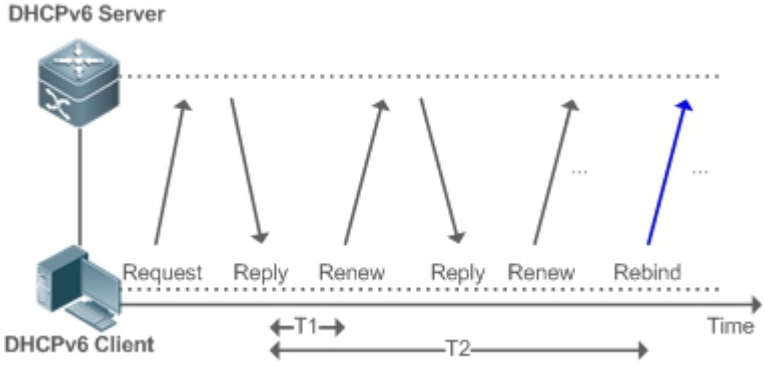
The DHCPv6 server provides the control address and the updated T1 and T2 in the IA of the message sent to the DHCPv6 client.

Figure 5- 7



- The DHCPv6 client will send a Renew multicast message to the DHCPv6 server for updating the address and prefix after T1 seconds. The Renew message contains the DUID of the DHCPv6 server and the IA information to be updated.
- After receiving the Renew message, the DHCPv6 server checks whether the DUID value in the Renew message is equal to the DUID value of the local device. If yes, the DHCPv6 server updates the local binding and sends a Reply message in the unicast mode. The Reply message contains the new T1 and other parameter s.

Figure 5- 8

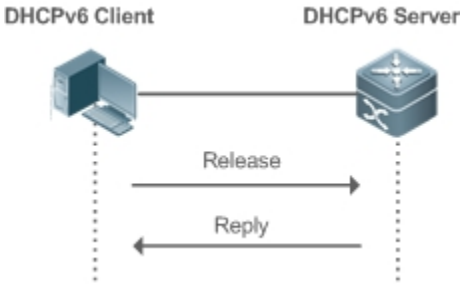


- If no response is received after the DHCPv6 client sends a Renew message to the DHCPv6 server, the DHCPv6 client will send a Rebind multicast message to the DHCPv6 server for rebinding the address and prefix after T2 expires.
- After receiving the Rebind message, the DHCPv6 server (perhaps a new DHCPv6 server) sends a Reply message according to the content of the Rebind message.

Release

If a DHCPv6 client needs to release an address or a prefix, the DHCPv6 client needs to send a Release message to a DHCPv6 server to notify the DHCPv6 server of the released addresses or prefixes. In this way, the DHCPv6 server can allocate these addresses and prefixes to other DHCPv6 clients.

Figure 5- 9



- After receiving the Release message, the DHCPv6 server removes the corresponding bindings based on the addresses or prefixes in the Release message, and sends a Reply message carrying the state option to the DHCPv6 client.

Confirmation

After moving to a new link (for example, after restart), a DHCPv6 client will send a Confirm message to the DHCPv6 server on the new link to check whether the original addresses are still available.

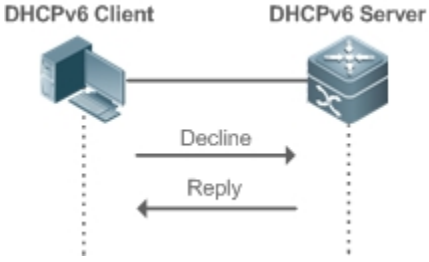
Figure 5- 10

- After receiving the Confirm message, the DHCPv6 server performs confirmation based on the address information in the Confirm message, and sends a Reply message carrying the state option to the DHCPv6 client. If the confirmation fails, the DHCPv6 client may initiate a new address allocation request.

DHCPv6 Conflict

If the DHCPv6 client finds that the allocated addresses have been used on the link after address allocation is completed, the DHCPv6 client sends a Decline message to notify the DHCPv6 server of the address conflict.

Figure 5- 11



- The DHCPv6 client includes the IA information of the conflicted addresses in the Decline message.
- After receiving the Decline message, the DHCPv6 server marks the addresses in the Decline message as "declined" and will not allocate these addresses. Then, the DHCPv6 server sends a Reply message carrying the state option to the DHCPv6 client. You can manually clear addresses marked as "declined" to facilitate re-allocation.

Related Configuration

Enabling the DHCPv6 Client Address Request Function on an Interface

- By default, an interface is not enabled with the DHCPv6 client address request function.
- You can run the **ipv6 dhcp client ia** command to enable the DHCPv6 client address request function for the interface.

The DHCPv6 client address request function is effective only on a layer-3 interface.

5.3.2 Requesting/Allocating Prefixes

Configure available prefixes on the DHCPv6 server. By using the prefix delegation of DHCPv6, uplink network devices can allocate address prefixes to downlink network devices, which implements flexible station-level automatic configuration and flexible control of station address space.

Working Principle

Downlink network devices serve as DHCPv6 clients to exchange messages with the DHCPv6 server to implement address allocation, update, release and other operations. Downlink network devices obtain, update, rebind and release prefixes by using the four-/two-message exchange mechanism similar to that for allocating addresses. However, prefix allocation is different from address allocation in the following aspects:

- In message exchange using the prefix delegation, the Confirm and Decline messages are not used.
- If a DHCPv6 client moves to a new link and needs to check whether the prefix information is available, it performs confirmation through Rebind and Reply message exchange.
- The IA type in various messages is IA_PD.

For the message exchange using the prefix delegation, refer to the section "Requesting/Allocating Addresses".

Related Configuration

Enabling the DHCPv6 Client Prefix Request Function on an Interface

By default, an interface is not enabled with the DHCPv6 client prefix request function.

You can run the **ipv6 dhcp client pd** command to enable or disable the DHCPv6 client prefix request function for the interface.

The DHCPv6 client prefix request function is effective only on a layer-3 interface.

5.3.3 Stateless Service

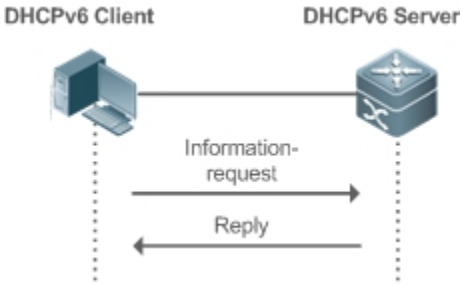
When a DHCPv6 client needs only configuration parameters, the DHCPv6 stateless service can be used to obtain related configuration parameters which cannot be obtained through a stateless automatic address configuration protocol, such as the DNS server address.

Working Principle

Network hosts serve as DHCPv6 clients to exchange messages with the DHCPv6 server to obtain and update configuration parameters.

Message Exchange Using the Stateless Service

Figure 5- 12



- A DHCPv6 client sends an Information-request message to a DHCPv6 server to request stateless messages. Usually, this message does not contain the DUID of the specified DHCPv6 server.
- The DHCPv6 server sends a Reply message containing the configuration parameters to the DHCPv6 client.

Related Configuration

Stateless Service of a DHCPv6 Client

- By default, an interface is not enabled with the stateless service of the DHCPv6 client.
- If a host receives an RA message containing the O flag, it will enable the stateless service.

5.3.4 Relay Service

When the DHCPv6 client and DHCPv6 server are on different links, the DHCPv6 client can relay related messages to the DHCPv6 server through the DHCPv6 relay agent. The DHCPv6 server also relays the response to the DHCPv6 client through the relay agent.

Working Principle

When receiving a message from the DHCPv6 client, the DHCPv6 relay agent creates a Relay-forward message. This message contains the original message from the DHCPv6 client and some options added by the relay agent. Then, the relay agent sends the Relay-forward message to a specified DHCPv6 server or a specified multicast address FF05::1:3.

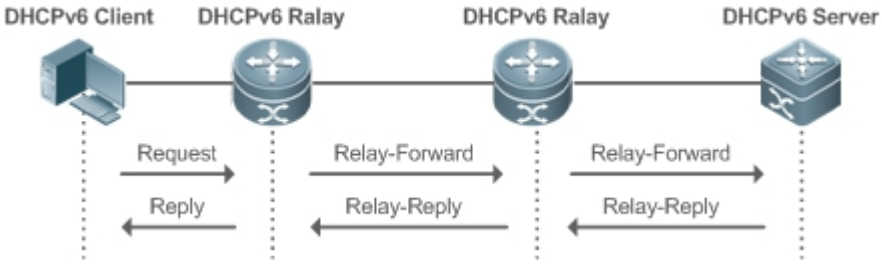
After receiving the Relay-forward message, the DHCPv6 server extracts the original message from the DHCPv6 client f for processing. Then, the DHCPv6 server constructs a response to the original message, encapsulates the response in a Relay-reply message, and then sends the Relay-reply message to the DHCPv6 relay agent.

After receiving the Relay-reply message, the DHCPv6 relay agent extracts the original message from the DHCPv6 server for processing, and forwards the message to the DHCPv6 client.

Multi-level relay agents are allowed between the DHCPv6 client and DHCPv6 server.

DHCPv6 Relay Agent

Figure 5- 13



- The DHCPv6 relay agent performs message encapsulation and decapsulation between the DHCPv6 client and DHCPv6 server to enable communication between the DHCPv6 client and DHCPv6 server on different links.

5.4 Configuration

Configuration	Description and Command	
Configuring the DHCPv6 Relay	(Mandatory) It is used to enable the DHCPv6 relay agent service.	
	ipv6 dhcp relay destination	Configures the DHCPv6 relay agent function.
Configuring the DHCPv6 Client	(Mandatory) It is used to request addresses or prefixes.	
	ipv6 dhcp client ia	Enables the DHCPv6 client and requests IANA addresses.
	ipv6 dhcp client pd	Enables the DHCPv6 client and requests address prefixes.
	(Optional) It is used to enable a host that receives an RA message to request stateless service through the DHCPv6 client.	
	ipv6 nd other-config-flag	Sets the O flag in the RA message on the device that sends the RA message so that the host that receives the RA message can request stateless service through the DHCPv6 client.

-

5.4.1 Configuring the DHCPv6 Relay

Configuration Effect

- A DHCPv6 relay agent can be configured for address allocation, prefix delegation and parameter allocation to enable communication between the DHCPv6 client and server on different links.

Notes

- A destination address must be specified. If the destination address is a multicast address (such as FF05::1:3), you also need to specify an egress interface.

Configuration Steps

Configuring the DHCPv6 Relay Agent Function

- Mandatory.
- Unless otherwise specified, you should configure the DHCPv6 relay agent function on all devices that need to provide the DHCPv6 relay agent service.

Verification

The DHCPv6 client and DHCPv6 server exchange messages through the relay agent.

- Check whether the interface is enabled with the DHCPv6 relay.

- Check whether the DHCPv6 relay agent can receive and send messages.

Related Commands

Configuring the DHCPv6 Relay Agent Function

Command	**ipv6 dhcp relay destination** *ipv6-address* [*interface-type interface-number*]
Parameter Description	*ipv6-address*: Specifies the destination address of the relay agent. *interface-type*: Specifies the type of the destination interface (optional). *interface-number*: Specifies the destination interface number (optional).
Command Mode	Interface configuration mode
Usage Guide	All DHCPv6 packets from clients received by an interface enabled with the DHCPv6 relay function will be encapsulated and sent to a specified destination address (or multiple destination addresses) through a specified interface (optional).

Configuration Example

Configuring the DHCPv6 Relay

Configuration Steps	Specify an interface enabled with the relay service to forward received DHCPv6 client packets to a specified destination address through the specified interface (optional).
	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface vlan 1 FS(config-if)#ipv6 dhcp relay destination 3001::2 FS(config-if)#ipv6 dhcp relay destination ff02::1:2 vlan 2
Verification	Run the **show ipv6 dhcp relay destination all** command to display the configured destination addresses.
	Interface:VLAN 1 Destination address(es) Output Interface 3001::2 ff02::1:2 VLAN 2

Common Errors

- The configuration is performed on other interfaces than the Switch Virtual Interface (SVI), routed port and L3 AP port.

5.4.2 Configuring the DHCPv6 Client

Configuration Effect

- Enable a device to automatically request IPv6 addresses or related parameters from a server.

Notes

- The configuration must be performed on layer-3 interfaces.

Configuration Steps

Enabling the DHCPv6 Client and Requesting IANA Addresses

- Mandatory.
- Unless otherwise specified, you should enable the DHCPv6 client address request function on all devices that need to request addresses.

Enabling the DHCPv6 Client and Requesting Address Prefixes

- Mandatory.
- Unless otherwise specified, you should enable the DHCPv6 client prefix request function on all devices that need to request prefixes.

Enabling the Stateless Service of the DHCPv6 Client

- It is mandatory if the DHCPv6 client needs to obtain configuration parameters.

Verification

Check whether the interface is enabled with the DHCPv6 client and check the addresses, prefixes and other configuration obtained on the interface.

Related Commands

Enabling the DHCPv6 Address Request Function

Command	**ipv6 dhcp client ia** [**rapid-commit**]
Parameter Description	**rapid-commit:** Permits the simplified message exchange process.
Command Mode	Interface configuration mode
Usage Guide	If the DHCPv6 client mode is not enabled, this command will enable the DHCPv6 client mode on the interface. After the **ipv6 dhcp client ia** command is configured, an IANA address request will be sent to the DHCPv6 server. The **rapid-commit** keyword permits the two-message exchange process between the client and server. If this keyword is configured, the Solicit message sent by the client contains the **rapid-commit** option.

Enabling the DHCPv6 Client Prefix Request

Command	**ipv6 dhcp client pd** *prefix-name* [**rapid-commit**]
Parameter Description	*prefix-name*: Indicates a IPv6 general prefix. **rapid-commit:** Permits the simplified message exchange process.
Command Mode	Interface configuration mode
Usage Guide	If the DHCPv6 client mode is not enabled, this command will enable the DHCPv6 client mode on the interface. After the **ipv6 dhcp client pd** command is configured, a prefix request will be sent to the DHCPv6 server. After receiving the prefix, the client will save the prefix in the IPv6 general prefix pool. Then, other commands and applications can use this prefix.

	The **rapid-commit** keyword permits the two-message exchange process between the client and server. If this keyword is configured, the Solicit message sent by the client contains the **rapid-commit** option.

Configuring Stateless Service

Command	**ipv6 nd other-config-flag**
Parameter Description	-
Command Mode	Interface configuration mode
Usage Guide	Configure this command on a host that sends the RA message. Then, the host that receives the RA message obtains stateless configurations through the DHCPv6 client.

Configuration Example

Enabling the DHCPv6 Address Request Function

Configuration Steps	● Configure the DHCPv6 client address request function on an interface.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if)# ipv6 dhcp client ia
Verification	● Run the **show ipv6 dhcp interface** command to display whether the interface is enabled with the DHCPv6 client.
	FS#show ipv6 dhcp interface GigabitEthernet 0/1 GigabitEthernet 0/1 is in client mode Rapid-Commit: disable

Enabling the DHCPv6 Client Prefix Request

Configuration Steps	● Configure the DHCPv6 client prefix request function on an interface.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if)# ipv6 dhcp client pd pd_name
Verification	● Run the **show ipv6 dhcp interface** command to display whether the interface is enabled with the DHCPv6 client.
	FS#show ipv6 dhcp interface GigabitEthernet 0/1 GigabitEthernet 0/1 is in client mode Rapid-Commit: disable

Enabling the DHCPv6 Stateless Service

Configuration Steps	● Configure this command on an interface that sends the RA message.
	FS# configure terminal FS(config)# interface GigabitEthernet 0/1 FS(config-if)# ipv6 nd other-config-flag
Verification	● Run the **show ipv6 dhcp interface** command to display whether an interface of the host obtains configuration parameters.
	FS#show ipv6 dhcp interface GigabitEthernet 0/2 GigabitEthernet 0/2 is in client mode DNS server: 2001::1 Rapid-Commit: disable

Common Errors

- The DHCPv6 client address request is enabled on non-layer-3 interfaces.
- The DHCPv6 address request is enabled on interfaces enabled with the DHCPv6 relay or DHCPV6 server.
- The DHCPv6 client prefix request is enabled on non-layer-3 interfaces.
- The DHCPv6 prefix request is enabled on interfaces enabled with the DHCPv6 relay or DHCPV6 server.

5.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the statistics on sent and received packets after the DHCPv6 relay is enabled on the current device.	**clear ipv6 dhcp relay statistics**
Restarts the DHCPv6 client.	**clear ipv6 dhcp client** *interface-type interface-number*

Displaying

Description	Command
Displays the DUID of a device.	**show ipv6 dhcp**
Displays DHCPv6 interface.	**show ipv6 dhcp interface** [*interface-name*]
Displays the destination address of the DHCPv6 relay agent.	**show ipv6 dhcp relay destination** { **all** | *interface-type interface-number* }
Displays the statistics on sent and received packets after the DHCPv6 relay is enabled on a device.	**show ipv6 dhcp relay statistics**
Displays the local IPv6 prefix pool.	**show ipv6 local pool** [*poolname*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs DHCPv6.	**debug ipv6 dhcp** [**detail**]

6 Configuring DNS

6.1 Overview

A Domain Name System (DNS) is a distributed database containing mappings between domain names and IP addresses on the Internet, which facilitate users to access the Internet without remembering IP strings that can be directly accessed by computers. The process of obtaining an IP address through the corresponding host name is called domain name resolution (or host name resolution).

Protocols and Standards

- RFC1034: DOMAIN NAMES - CONCEPTS AND FACILITIES
- RFC1035: DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION

6.2 Applications

Application	Description
Static Domain Name Resolution	Performs domain name resolution directly based on the mapping between a domain name and an IP address on a device.
Dynamic Domain Name Resolution	Obtains the IP address mapped to a domain name dynamically from a DNS server on the network.

6.2.1 Static Domain Name Resolution

Scenario

- Preset the mapping between a domain name and an IP address on a device.
- When you perform domain name operations (such as Ping and Telnet) through application programs, the system can resolve the IP address without being connected to a server on the network.

Deployment

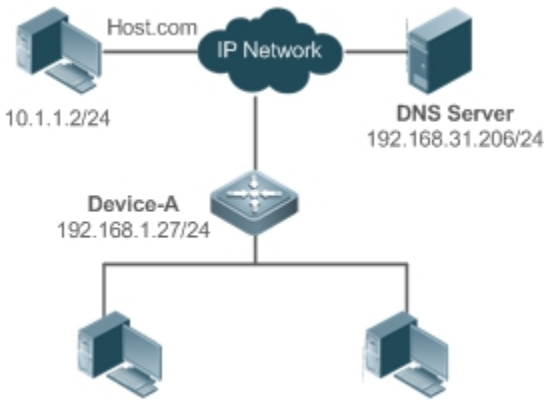
- Preset the mapping between a domain name and an IP address on a device.

6.2.2 Dynamic Domain Name Resolution

Scenario

- DNS Server is deployed on the network to provide the domain name service.
- Domain name "host.com" is deployed on the network.
- Device-A applies to DNS Server for domain name "host.com".

Figure 6- 1 Dynamic Domain Name Resolution



Deployment

- Deploy DNS Server as the DNS server of Device-A.

6.3 Features

Basic Concepts

DNS

The DNS consists of a resolver and a DNS server. The DNS server stores the mappings between domain names and IP addresses of all hosts on the network, and implements mutual conversion between the domain names and IP addresses. Both the TCP and UDP port IDs of DNS are 53, and generally a UDP port is used.

Features

Feature	Description
Domain Name Resolution	IP addresses are obtained based on domain names from a DNS server or a local database.

6.3.1 Domain Name Resolution

Working Principle

Static Domain Name Resolution

Static domain name resolution means that a user presets the mapping between a domain name and an IP address on a device. When you perform domain name operations (such as Ping and Telnet) through application programs, the system can resolve the IP address without being connected to a server on the network.

Dynamic Domain Name Resolution

Dynamic domain name resolution means that when a user perform domain name operations through application programs, the DNS resolver of the system queries an external DNS server for the IP address mapped to the domain name.

The procedure of dynamic domain name resolution is as follows:

1. A user application program (such as Ping or Telnet) requests the IP address mapped to a domain name from the DNS resolver of the system.
2. The DNS resolver queries the dynamic cache at first. If the domain name on the dynamic cache does not expire, the DNS resolver returns the domain name to the application program.

3. If all domain names expire, the DNS resolver initiates a request for domain name-IP address conversion to the external DNS server.
4. After receiving a response from the DNS server, the DNS resolver caches and transfers the response to the application program.

Related Configuration

Enabling Domain Name Resolution

- By default, domain name resolution is enabled.
- Run the **ip domain-lookup** command to enable domain name resolution.

Configuring the IP Address Mapped to a Static Domain Name

- By default, no mapping between a domain name and an IP address is configured.
- Run the **ip host** command to specify the IPv4 address mapped to a domain name.
- Run the **ipv6 host** command to specify the IPv6 address mapped to a domain name.

Configuring a DNS Server

- By default, no DNS server is configured.
- Run the **ip name-server** command to configure a DNS server.

6.4 Configuration

Configuration	Description and Command	
Configuring Static Domain Name Resolution	Optional.	
	ip domain-lookup	Enables domain name resolution.
	ip host	Configures the IPv4 address mapped to a domain name.
	ipv6 host	Configures the IPv6 address mapped to a domain name.
Configuring Dynamic Domain Name Resolution	Optional.	
	ip domain-lookup	Enables domain name resolution.
	ip name-server	Configures a DNS server.

6.4.1 Configuring Static Domain Name Resolution

Configuration Effect

The system resolver resolves the IP address mapped to a domain name on a local device.

Configuration Steps

Enabling Domain Name Resolution

- The domain name resolution function is enabled by default.
- If this function is disabled, static domain name resolution does not take effect.

Configuring the IP Address Mapped to a Domain Name

- (Mandatory) Domain names to be used must be configured with mapped IP addresses.

Verification

- Run the **show run** command to check the configuration.
- Run the **show hosts** command to check the mapping between the domain name and the IP address.

Related Commands

Configuring the IPv4 Address Mapped to a Domain Name

Command	**ip host** *host-name ip-address*
Parameter Description	*host-name*: indicates a domain name. *ip-address*: indicates a mapped IPv4 address.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the IPv6 Address Mapped to a Domain Name

Command	**ipv6 host** *host-name ipv6-address*
Parameter Description	*host-name*: indicates a domain name. *ipv6-address*: indicates a mapped IPv6 address.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring Static Domain Name Resolution

Configuration Steps	• Set the IP address of static domain name www.test.com to 192.168.1.1 on a device. • Set the IP address of static domain name www.testv6.com to 2001::1 on a device.
	FS#configure terminal FS(config)# ip host www.test.com 192.168.1.1 FS(config)# ipv6 host www.testv6.com 2001::1 FS(config)# exit
Verification	Run the **show hosts** command to check whether the static domain name entry is configured.
	FS#show hosts Name servers are: Host type Address TTL(sec) www.test.com static 192.168.1.1 --- www.testv6.com static 2001::1 ---

6.4.2 Configuring Dynamic Domain Name Resolution

Configuration Effect

The system resolver resolves the IP address mapped to a domain name through a DNS server.

Configuration Steps

Enabling Domain Name Resolution

- Domain name resolution is enabled by default.
- If this function is disabled, dynamic domain name resolution does not take effect.

Configuring a DNS Server

- (Mandatory) To use dynamic domain name resolution, you must configure an external DNS server.

Verification

- Run the **show run** command to check the configuration.

Related Commands

Configuring a DNS Server

Command	**ip name-server** { *ip-address* | *ipv6-address* }
Parameter Description	*ip-address*: indicates the IPv4 address of the DNS server. *Ipv6-address*: indicates the IPv6 address of the DNS server.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring Dynamic Domain Name Resolution

Scenario **Figure 6- 2**	
	Device resolves the domain name through the DNS server (192.168.10.1) on the network.
Configuration Steps	Set the IP address of the DNS server to 192.168.10.1 on the device.
	DEVICE#configure terminal DEVICE(config)# ip name-server 192.168.10.1 DEVICE(config)# exit
Verification	Run the **show hosts** command to check whether the DNS server is specified.

	FS(config)#show hosts Name servers are: 192.168.10.1 static Host type Address TTL(sec)

6.5 Monitoring

Clearing

Running the **clear** command during device operation may cause data loss or even interrupt services.

Description	Command
Clears the dynamic host name cache table.	**clear host** [*host-name*]

Displaying

Description	Command
Displays DNS parameters.	**show hosts** [*host-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the DNS function.	**debug ip dns**

7 Configuring FTP Server

7.1 Overview

The File Transfer Protocol (FTP) server function enables a device to serve as an FTP server. In this way, a user can connect an FTP client to the FTP server and upload files to and download files from the FTP server through FTP.

A user can use the FTP server function to easily obtain files such as syslog files from a device and copy files to the file system of the device through FTP.

Protocols and Standards

- RFC959: FILE TRANSFER PROTOCOL (FTP)
- RFC3659: Extensions to FTP
- RFC2228: FTP Security Extensions
- RFC2428: FTP Extensions for IPv6 and NATs
- RFC1635: How to Use Anonymous FTP

7.2 Applications

Application	Description
Providing FTP Services in a LAN	Provides the uploading and downloading services for a user in a Local Area Network (LAN).

7.2.1 Providing FTP Services in a LAN

Scenario

Provide the uploading and downloading services for a user in a LAN.

As shown in Figure 7- 1, enable the FTP server function only in a LAN.

- G and S are enabled with the FTP server function and layer-2 transparent transmission function respectively.
- A user initiates a request for FTP uploading and downloading services.

Figure 7- 1

Remarks	G is an egress gateway device. S is an access device.

Deployment

- G is enabled with the FTP server function.
- As a layer-2 switch, S provides the function of layer-2 transparent transmission.

7.3 Features

Basic Concepts

FTP

FTP is a standard protocol defined by the IETF Network Working Group. It implements file transfer based on the Transmission Control Protocol (TCP). FTP enables a user to transfer files between two networked computers and is the most important approach to transferring files on the Internet. A user can obtain abundant Internet for free through anonymous FTP. In addition, FTP provides functions such as login, directory query, file operation, and other session control. Among the TCP/IP protocol family, FTP is an application-layer protocol and uses TCP ports 20 and 21 for transmission. Port 20 is used to transmit data and port 21 is used to transmit control messages. Basic operations of FTP are described in RFC959.

User Authorization

To connect an FTP client to an FTP server, you should have an account authorized by the FTP server. That is, a user can enjoy services provided by the FTP server after logging in to the FTP server with a user name and password. A maximum of 10 accounts can be configured, a maximum of 2 connections are allowed for each account, and a maximum of 10 connections are supported by the server.

FTP File Transmission Modes

FTP provides two file transmission modes:

- Text transmission mode (ASCII mode): It is used to transfer text files (such as .txt, .bat, and .cfg files). This mode is different from the binary mode in carriage return and line feed processing. In ASCII mode, carriage return and line feed are changed to local CRC characters, for example, \n in Unix, \r\n in Windows, and \r in Mac. Assume that a file being copied contains ASCII text. If a remote computer does not run Unix, FTP automatically converts the file format to suit the remote computer.
- Binary transmission mode: It is used to transfer program files (for example, .app, .bin and .btm files), including executable files, compressed files and image files without processing data. Therefore, Binary mode facilitates faster transfer of all files and more reliable transfer of ASCII files.

FTP Working Modes

FTP provides two working modes:

Figure 7- 2

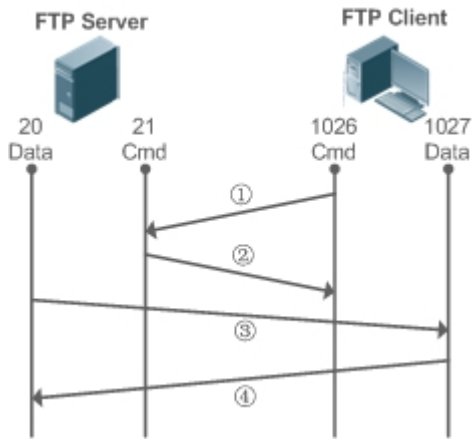
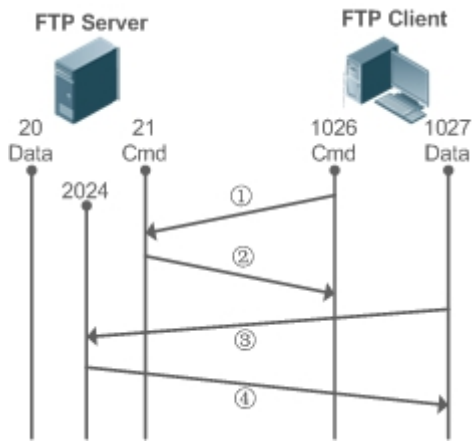


Figure 7- 3



- Figure 7-2 shows the active (PORT) mode. The FTP client uses port 1026 to connect to the FTP server through port 21. The client sends commands through this channel. Before receiving data, the client sends the **PORT** command on this channel. The **PORT** command contains information on the channel port (1027) of the client for receiving data. The server uses port 20 to connect to the client through port 1027 for establishing a data channel to receive and transmit data. The FTP server must establish a new connection with the client for data transmission.
- Figure 7-3 shows the passive (PASV) mode. The process for establishing a control channel is similar to that in the PORT mode. However, after the connection is established, the client sends the **PASV** command rather than the **PORT** command. After receiving the **PASV** command, the FTP server enables a high-end port (2024) at random and notifies the client that data will be transmitted on this port. The client uses port 1027 to connect the FTP server through port 2024. Then, the client and server can transmit and receive data on this channel. In this case, the FTP server does not need to establish a new connection with the client.

Supported FTP Commands

After receiving an FTP connection request, the FTP server requires the client to provide the user name and password for authentication.

If the client passes the authentication, the FTP client commands can be executed for operations. The available FTP client commands are listed as follows:

ascii	delete	mdelete	mput	quit	send
bin	dir	mdir	nlist	recv	size
bye		mget		rename	system
cd	get	mkdir	passive		type
cdup		mls	put	rmdir	user
close	ls		pwd		

For usage of these FTP client commands, please refer to your FTP client software document. In addition, many FTP client tools (such as CuteFTP and FlashFXP) provide the graphic user interface. These tools facilitate operations by freeing users from configuring FTP commands.

Overview

Feature	Description
Enabling the FTP Server Function	Provides the functions of uploading, downloading, displaying, creating and deleting files for an FTP client.

7.3.1 Enabling the FTP Server Function

Working Principle

The basic working principle is described in the previous chapter. FS devices provide FTP services after the user name, password, and top-level directory are configured.

Related Configuration

Enabling the FTP Server Function Globally

The FTP server function is disabled by default.

Run the **ftp-server enable** command to enable the FTP server function.

You must enable the FTP server function globally before using it.

Configuring a User Name, Password, and Top-Level Directory

There is no authorized user or top-level directory by default.

Run the **ftp-server usernamepassword** and **ftp-server topdir** commands to set an authorized user and top-level directory.

The three configurations above are mandatory; otherwise, the FTP server function cannot be enabled.

7.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions	(Mandatory) It is used to enable an FTP server.	
	ftp-server enable	Enables the FTP server function.
	ftp-server login timeout	Configures Login timeout for an FTP session.
	ftp-server login times	Configures the valid login count.
	ftp-server topdir	Configures the top-level directory of the FTP server.
	ftp-server username password	Configures a user name and password.
	Optional.	
	ftp-server timeout	Configures the idle timeout of an FTP session.

7.4.1 Configuring Basic Functions

Configuration Effect

- Create an FTP server to provide FTP services for an FTP client.

Notes

- The user name, password, and top-level directory need to be configured.
- To enable the server to close an abnormal session within a limited period, you need to configure the idle timeout of a session.

Configuration Steps

Enabling the FTP Server Function

- Mandatory.
- Unless otherwise noted, enable the FTP server function on every router.

Configuring a Top-Level Directory

- Mandatory.
- Unless otherwise noted, configure the top-level directory as the root directory on every router.

Configuring a User Name and Password for Login

- Mandatory.
- The lengths of the user name and password are restricted.

Configuring the Login Timeout for an FTP Session

- Optional.
- When the client is disconnected from the server due to an error or other abnormal causes, the FTP server may not know that the user is disconnected and continues to keep the connection. Consequently, the FTP connection is occupied for a long time and the server cannot respond to the login requests of other users. This configuration can ensure that other users can connect to the FTP server within a period of time upon an error.

Verification

Connect an FTP client to the FTP server.

- Check whether the client is connected.
- Check whether operations on the client are normal.

Related Commands

Enabling the FTP Server Function

Command	**ftp-server enable**
Parameter Description	**-**
Command Mode	Global configuration mode
Usage Guide	The client cannot access the FTP server unless the top-level directory, user name and password are configured. Therefore, it is recommended that you configure the top-level directory, user name and password for login by referring to the subsequent chapters before enabling the service for the first time.

Configuring the Valid Login Count

Command	**ftp-server login times** *times*
Parameter Description	*times:* Indicates the valid login count, ranging from 1 to 10.
Command	Global configuration mode

Mode	
Usage Guide	The valid login count refers to the number of times you can perform account verification during an FTP session. The default value is 3, which means that your session will be terminated if you enter an incorrect user name or password for three times and other users can go online.

Configuring the Login Timeout for an FTP Session

Command	**ftp-server login timeout** *timeout*
Parameter Description	*timeout*: Indicates the login timeout, ranging from 1 to 30 minutes.
Command Mode	Global configuration mode
Usage Guide	The login timeout refers to the maximum duration that the session lasts since being established. If you do not pass the password verification again during the login timeout, the session will be terminated to ensure that other users can log in.

Configuring the Top-Level Directory of the FTP Server

Command	**ftp-server topdir** *directory*
Parameter Description	*directory:* Indicates the user access path.
Command Mode	Global configuration mode
Usage Guide	If the top-level directory of the server is set to "/syslog", the FTP client can access only the files and directories in the "/syslog" directory on the device after login. Due to restriction on the top-level directory, the client cannot return to the upper directory of "/syslog".

Configuring a User Name and Password for Server Login

Command	**ftp-server username** *username* **password** [*type*] *password*
Parameter Description	***Username***: Indicates a user name. *type*: 0 or 7. 0 indicates that the password is not encrypted (plaintext) and 7 indicates that the password is encrypted (cipher text). **password**: Indicates a password.
Command Mode	Global configuration mode
Usage Guide	The FTP server does not support anonymous login; therefore, a user name must be configured. A user name consists of up to 64 characters including letters, half-width digits and symbols without spaces. A password consists of only letters or digits. Spaces at the beginning and end of the password are ignored. Spaces inside the password are viewed as part of the password. A plaintext password consists of 1 to 25 characters. A cipher text password consists of 4 to 52 characters. User names and passwords must match. A maximum of 10 users can be configured.

Configuring the Idle Timeout for an FTP Session

Command	**ftp-Server timeout** *time*
Parameter Description	*time*: Indicates the idle timeout, ranging from 1 to 3,600 minutes.
Command Mode	Global configuration mode
Usage Guide	The idle timeout of a session refers to the duration from the end of an FTP operation to the start of the next FTP operation in an FTP session. After the server responds to an FTP client command operation (for example, after a file is completely transferred), the server starts to count the idle time again, and stops when the next FTP client command operation arrives. Therefore, the configuration of the idle timeout has no effect on some time-consuming file transfer operations.

Displaying Server Status

Command	**show ftp-server**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	Run this command to display FTP server status.

Debugging

Command	**debug ftp-server pro/err**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	Run this command to debug message/error events of the FTP server.

Configuration Example

Creating an FTP Server on an IPv4 Network

Scenario	● A TCP connection is established for transmission from a server to a client.
Configuration Steps	● Enable the FTP server function. ● Configure the top-level directory/syslog. ● Set the user name **user** and password to **password**. ● Set the session idle timeout to 5 minutes.
	FS(config)#ftp-server username user FS(config)#ftp-server password password FS(config)#ftp-server timeout 5 FS(config)#ftp-server topdir / FS(config)#ftp-server enable
Verification	Run the **show ftp-server** command to check whether the configuration takes effect.

```
FS#show ftp-server
    ftp-server information
==================================
enable : Y
topdir : tmp:/
timeout: 10min
username:aaaa          password:(PLAINT)bbbb          connect num[2]
    [0]trans-type:BINARY (ctrl)server IP:192.168.21.100[21]
                                    client IP:192.168.21.26[3927]
    [1]trans-type:ASCII (ctrl)server IP:192.168.21.100[21]
                                    client IP:192.168.21.26[3929]
username:a1            password:(PLAINT)bbbb          connect num[0]
username:a2            password:(PLAINT)bbbb          connect num[0]
username:a3            password:(PLAINT)bbbb          connect num[0]
username:a4            password:(PLAINT)bbbb          connect num[0]
username:a5            password:(PLAINT)bbbb          connect num[0]
username:a6            password:(PLAINT)bbbb          connect num[0]
username:a7            password:(PLAINT)bbbb          connect num[0]
username:a8            password:(PLAINT)bbbb          connect num[0]
username:a9            password:(PLAINT)bbbb          connect num[0]
```

Common Errors

- No user name is configured.
- No password is configured.
- No top-level directory is configured.

7.5 Monitoring

Displaying

Description	Command
Displays the FTP server configuration.	**show ftp-server**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the FTP server error events.	**debug ftp-server err**

Debugs the FTP server message events.	**debug ftp-server pro**

8 Configuring FTP Client

8.1 Overview

The File Transfer Protocol (FTP) is an application of TCP/IP. By establishing a connection-oriented and reliable TCP connection between the FTP client and server, a user can access a remote computer that runs the FTP server program.

An FTP client enables file transfer between a device and the FTP server over the FTP protocol. A user uses the client to send a command to the server. The server responds to the command and sends the execution result to the client. By means of command interaction, the user can view files in the server directory, copy files from a remote computer to a local computer, or transfer local files to a remote computer.

FTP is intended to facilitate sharing of program/data files and encourage remote operation (by using programs). Users do not need to be concerned with differences of different files systems on different hosts. Data is transmitted in an efficient and reliable manner. FTP enables remote file operation securely.

FS FTP clients are different from standard FTP clients that run interactive commands. Instead, you enter the **copy** command in CLI to perform control-connection instructions such as **open**, **user**, and **pass**. After a control connection is established, the file transfer process starts, and then a data connection is established to upload or download files.

Old devices support TFTP. However, TFTP is used to transfer small files whereas FTP is used to transfer large files. Implementing FTP on a device enables the file transfer between the local device and other clients or servers.

Protocols and Standards

- RFC959: FILE TRANSFER PROTOCOL (FTP)

8.2 Applications

Application	Description
Uploading a Local File to a Remote Server	Local and remote files need to be shared, for example, uploading a local file to a remote server.
Downloading a File from a Remote Server to a Local Device	Local and remote files need to be shared, for example, downloading a file from a remote server to a local device.

8.2.1 Uploading a Local File to a Remote Server

Scenario

Local and remote files need to be shared, for example, uploading a local file to a remote server.

As shown in Figure 8- 1, resources are shared only on the Intranet.

Figure 8- 1

Deployment

- Implement only communication on the Intranet.
- Enable file uploading on the FTP client.
- Enable file uploading on the FTP server.

8.2.2 Downloading a File from a Remote Server to a Local Device

Scenario

Local and remote files need to be shared, for example, downloading a file from a remote server to a local device.

As shown in Figure 8- 2, resources are shared only on the Intranet.

Figure 8- 2

Deployment

- Implement only communication on the Intranet.
- Enable file downloading on the FTP client.
- Enable file downloading on the FTP server.

8.3 Features

Basic Concepts

Uploading FTP Files

Upload files from an FTP client to an FTP server.

Downloading FTP Files

Download files from an FTP server to an FTP client.

FTP Connection Mode

An FTP client and an FTP server can be connected in the active or passive mode.

FTP Transmission Mode

The transmission between an FTP client and an FTP server is available in two modes, namely, text (ASCII) and binary (Binary).

Specifying the Source Interface IP Address for FTP Transmission

An FTP client is configured with a source IP address for communication with an FTP server.

Overview

Feature	Description
Uploading FTP Files	Uploads files from an FTP client to an FTP server.
Downloading FTP Files	Downloads files from an FTP server to an FTP client.
FTP Connection Mode	Specifies the connection mode between an FTP client and an FTP server.
FTP Transmission Mode	Specifies the transmission mode between an FTP client and an FTP server.
Specifying the Source Interface IP Address for FTP Transmission	Configures a source IP address of an FTP client for communication with an FTP server.

8.3.1 Uploading FTP Files

FTP enables file uploading. Start the FTP client and FTP server simultaneously, and upload files from the FTP client to the FTP server.

8.3.2 Downloading FTP Files

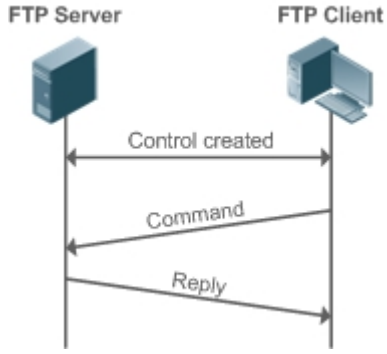
FTP enables file downloading. Start the FTP client and FTP server simultaneously, and download files from the FTP server to the FTP client.

8.3.3 FTP Connection Mode

FTP needs to use two TCP connections: one is a control link (command link) that is used to transfer commands between the FTP client and server; the other one is a data link that is used to upload or download data.

1. Control connection: Some simple sessions are enabled with the control connection only. A client sends a command to a server. After receiving the command, the server sends a response. The process is shown in Figure 8- 3.

Figure 8- 3 Control Connection



2. Control connection and data connection: When a client sends a command for uploading or downloading data, both the control connection and data connection need to be established.

FTP supports two data connection modes: active (PORT) and passive (PASC). The two modes are different in establishing a data connection.

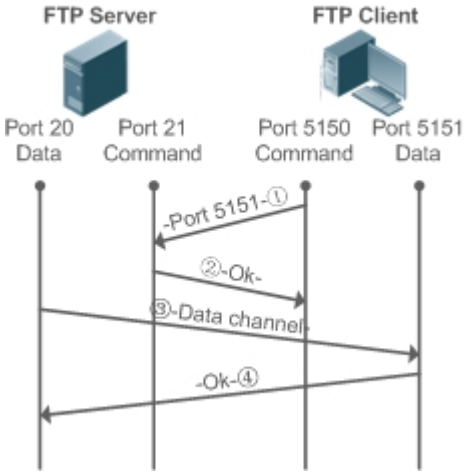
- Active mode

In this mode, an FTP server connects to an FTP client actively when a data connection is established. This mode comprises four steps:

1. The client uses source port 5150 to communicate with the server through port 21 as shown in Figure 8-4 to send a connection request and tell the server that the port to be used is port 5151.

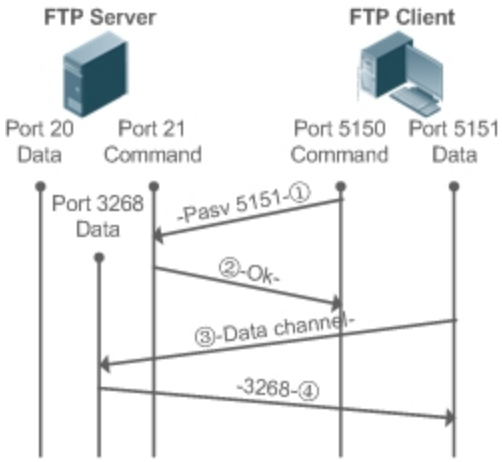
2. After receiving the request, the server sends a response OK(ACK). The client and server exchanges control signaling by console ports.
3. The server enables port 20 as the source port to send data to port 5151 of the client.
4. The client sends a response. Data transmission ends.

Figure 8- 4 Active (PORT) Mode



- Passive mode

Figure 8- 5 Passive (PASV) Mode



This mode is often set by the **passive** command. When a data connection is established, the FTP server is connected to the FPT client passively. This mode comprises four steps:

1. In the passive mode, the client initializes the control signaling connection. The client uses source port 5150 to connect to the server through port 21 as shown in Figure 8-5, and runs the **passive** command to request the server to enter the PASV mode.
2. The server agrees to enter the PASV mode, selects a port number greater than 1024 at random, and tells the port number to the client.
3. After receiving the message, the client uses port 5151 as shown in Figure 8-5 to communicate with the server through port 3268. Here, port 5151 is the source port and port 3268 is the destination port.

4. After receiving the message, the server sends data and responds an ACK(OK) response.

After the data connection is established, you can perform file uploading and downloading. Besides, you can perform some operations on the server file from the client.

The control connection for command and feedback transmission is always present whereas the data connection is established as required. Only an FTP client has the right to select and set the PASV or PORT mode. The FTP client sends a command to establish a data connection. FS FTP clients use the PASV mode by default.

8.3.4 FTP Transmission Mode

FTP provides two transmission modes: text (ASCII) and binary (Binary). At present, FS FTP clients support both the ASCII and Binary modes and use the BINARY mode by default.

- ASCII mode

The difference between the ASCII and Binary modes lies in carriage return and line feed processing. In ASCII mode, carriage return and line feed are changed to a local Carriage Return Character (CRC), for example, \n in Unix, \r\n in Windows, and \r in Mac.

- Binary mode

The Binary mode can be used to transfer executable files, compressed files and image files without processing data. For example, a text file needs to be transferred from Unix to Windows. When the Binary mode is used, the line breaks in Unix will not be converted from \r to \r\n; therefore in Windows, this file has no line feeds and displays many black squares. Therefore, Binary mode facilitates faster transfer of all files and more reliable transfer of ASCII files.

8.3.5 Specifying the Source Interface IP Address for FTP Transmission

An FTP client is configured with a source IP address for communication with an FTP server. In this way, the FTP client connects to the server and shares files with the server through the specified source IP address.

8.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions	(Mandatory) It is used to configure the functions of an FTP client.	
	copy flash	Uploads a file.
	copy ftp	Downloads a file.
Configuring Optional Functions	(Optional) It is used to configure the working mode of the FTP client.	
	ftp-client port	Sets the connection mode to active (port).
	ftp-client ascii	Sets the transmission mode to ASCII.
	ftp-client source-address	Configures the source IP address of the FTP client.
	default ftp-client	Restores the default settings, namely, connection mode set to passive (PASV), transmission mode to Binary and source IP address removed.

8.4.1 Configuring Basic Functions

Configuration Effect

- Implement file uploading and downloading.

Notes

- Pay attention to the command formats for uploading and downloading.

Configuration Steps

Uploading a File

- This configuration is mandatory when a file needs to be uploaded.
- Configure the FTP URL as the destination address of **copy** in Privileged EXEC mode.

Downloading a File

- This configuration is mandatory when a file needs to be downloaded.
- Configure the FTP URL as the source address of **copy** in Privileged EXEC mode.

Verification

- Check whether the uploaded file exists on the FTP server.
- Check whether the downloaded file exists at the destination address.

Related Commands

Uploading a File

Command	**copy flash:**[*local-directory/*]*local-file* **ftp:** *//username:password@dest-address*[*/remote-directory*]*/remote-file*
Parameter Description	*local-directory:* Specifies a directory on the local device. If it is not specified, it indicates the current directory. *local-file:* Specifies a local file to be uploaded. *username*: Specifies a user name for accessing the FTP server, consisting of no more than 32 bytes and excluding delimiters such as /, :, @ and space. This parameter is mandatory. *password:* Specifies a password for accessing the FTP server, consisting of no more than 32 bytes and excluding delimiters such as /, :, @ and space. This parameter is mandatory. *dest-address:* Specifies an IP address for the FTP server. *remote-directory:* Specifies a directory on the server. *remote-file:* Renames the file on the server. The directory specified by the *local-directory* field must have been created on the device. This command will not automatically create a directory.
Command Mode	Global configuration mode
Usage Guide	Run this command to upload a file from the flash of a local device to an FTP server.

Downloading an FTP File

Command	**copy ftp:**//*username:password@dest-address*[/*remote-directory*]/*remote-file* **flash:**[*local-directory/*]*local-file*
Parameter Description	*username*: Specifies a user name for accessing the FTP server, consisting of no more than 32 bytes and excluding delimiters such as /, :, @ and space. This parameter is mandatory. *password:* Specifies a password for accessing the FTP server, consisting of no more than 32 bytes and excluding delimiters such as /, :, @ and space. This parameter is mandatory. *dest-address:* Specifies an IP address for the FTP server. *remote-directory:* Specifies a directory on the server. *remote-file:* Specifies a file to be downloaded. *local-directory:* Specifies a directory on the local device. If it is not specified, it indicates the current directory. *local-file:* Renames the file in the local flash. The directory specified by the *local-directory* field must have been created on the device. This command will not automatically create a directory.
Command Mode	Global configuration mode
Usage Guide	Run this command to download a file from an FTP server to the flash of a local device.

Configuration Example

Uploading a File

Configuration Steps	Upload the **local-file** file in the **home** directory of a device to the **root** directory of an FTP server whose user name is **user**, password is **pass** and IP address is **192.168.23.69** and name the file as **remote-file**.
	FS# copy flash: home/local-file ftp://user:pass@192.168.23.69/root/remote-file
Verification	Check whether the **remote-file** file exists on the FTP server.

Downloading a File

Configuration Steps	Download the **remote-file** file from the **root** directory of an FTP server whose user name is **user**, password is **pass** and IP address is **192.168.23.69** to the **home** directory of a device and save the file as **local-file**.
	FS# copy ftp://user:pass@192.168.23.69/root/remote-file flash: home/local-file
Verification	Check whether the **remote-file** file exists in the **home** directory of the flash.

Common Errors

- The command formats for uploading and downloading are incorrect.
- The user name or password is incorrect.

8.4.2 Configuring Optional Functions

Configuration Effect

- Set the connection and transmission modes and configure a source IP address of the client for file uploading and download.

Notes

N/A Configuration Steps

Setting the Connection Mode to Active (Port)

- Optional.
- Configure the connection mode of FTP.

Setting the Transmission Mode to ASCII

- Optional.
- Configure the transmission mode of FTP.

Configuring the Source IP Address of the FTP Client

- Optional.
- Configure the source IP address of the FTP client.

Restoring the Default Settings

- Optional.
- Restore the default settings of the FTP client.

Verification

Run the **show run** command to check whether the configuration takes effect.

Related Commands

Setting the Connection Mode to Active (Port)

Command	**ftp-client port**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Run this command to set the connection mode to active (port). The default connection mode is passive (PASV).

Configuring the Source IP Address of the FTP Client

Command	**ftp-client source-address** { *ip-address* | *ipv6-address* }
Parameter Description	*ip-address:* Specifies the IPv4 address of a local interface. *ipv6-address:* Specifies the IPv6 address of a local interface.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure an interface IP address of the client for connection to the server. By default, the client is not configured with a local IP address. Instead, the route selects an IP address for the client.

Setting the Transmission Mode to ASCII

Command	**ftp-client ascii**

Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Run this command to set the transmission mode to ASCII. The default transmission mode is Binary.

Restoring the Default Settings

Command	**default ftp-client**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Run this command to restore the default settings, namely, connection mode set to passive (PASV), transmission mode to Binary and source IP address removed.

Configuration Example

Configuring Optional Functions

Configuration Steps	● Set the connection mode of FTP to port. ● Set the transmission mode to ASCII. ● Set the source IP address to 192.168.23.167.
	FS# configure terminal FS(config)# ftp-client ascii FS(config)# ftp-client port FS(config)# ftp-client source-address 192.168.23.167 FS(config)# end
Verification	Run the **show run** command on the device to check whether the configuration takes effect. FS# show run ! ftp-client ascii ftp-client port ftp-client source-address 192.168.23.167 !

Common Errors

- The source IP address is not a local IP address.

8.5 Monitoring

Displaying

Description	Command
Displays the FTP client configuration.	**show run**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the FTP Client.	**debug ftp-client**

9 Configuring TFTP

9.1 Overview

The Trivial File Transfer Protocol (TFTP) service enables a device to be configured as a TFTP server. Then the client can be connected to the TFTP server to upload files to or download files from the device using the TFTP protocol.

Users can easily obtain files such as upgrade package files from the device or copy files to the file system of the device using the TFTP service.

Protocols and Standards

- RFC1350: The TFTP Protocol (revision 2)
- RFC2347: TFTP Option Extension
- RFC2348: TFTP Blocksize Option
- RFC2349: TFTP Timeout Interval and Transfer Size Options

9.2 Applications

Application	Description
Providing the TFTP Service in a LAN	Enables users in a LAN to upload and download files.

9.2.1 Providing the TFTP Service in a LAN

Scenario

Enable users in a LAN to upload and download files.

In the following figure:

- Device G serves as a TFTP server.
- The User sends a TFTP uploading or downloading request.

Figure 9- 1

Remarks G is a network device on which the TFTP server is enabled.

Deployment

- Enable the TFTP server on the device G.
- The user uploads files to or download files from the device G.

9.3 Features

Basic Concepts

TFTP

TFTP is a set of standard protocols defined by the IETF Network Working Group, and operates at the application layer. Implemented on the top of the User Datagram Protocol (UDP), TFTP is a simple protocol to transfer files. TFTP provides only the file uploading and downloading functions instead of many common FTP functions. It does not support the directory list and the authentication function, and does not provide any security mechanism. TFTP uses the way of acknowledged retransmission upon timeout to ensure data transmission, which covers three transmission modes: netascii in the form of an eight-bit ASCII code, eight-bit octet of the source data type, and mail (which is no longer supported). TFTP uses UDP port 69. A description of TFTP can be found in RFC 1350.

TFTP Packet

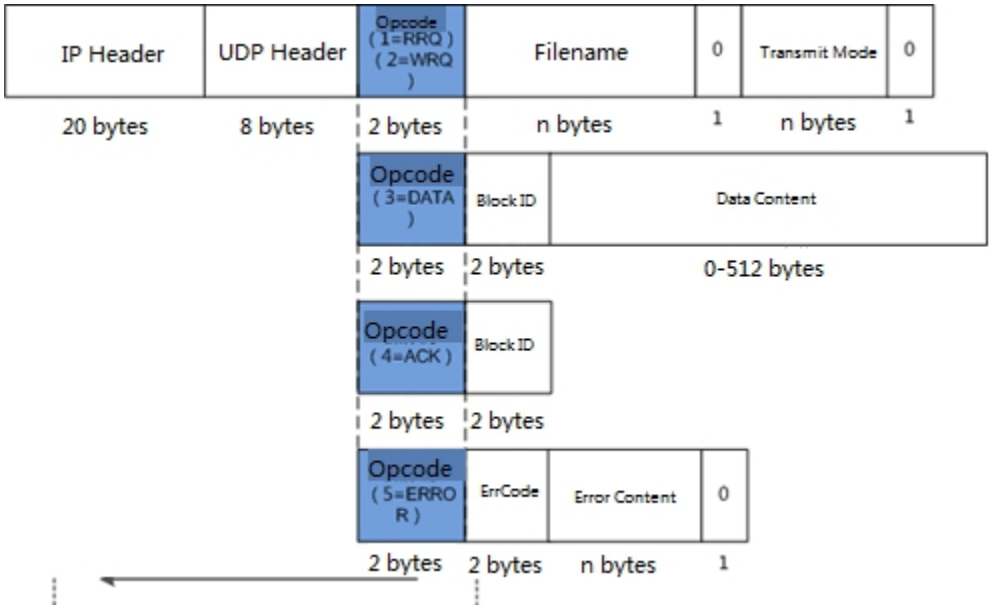
Any transfer begins with a request to read or write a file from a TFTP client. After the TFTP server grants the request, the file is sent in fixed length blocks of 512 bytes. A data packet of less than 512 bytes indicates the termination of a transfer.

Each data packet contains a block of data, and must be acknowledged by an acknowledgement packet before the next data packet can be sent. If no acknowledgement packet is received within specified time, the last sent data packet is retransmitted.

The TFTP packet header includes an opcode field, which indicates the packet type. TFTP supports the following five types of packets:

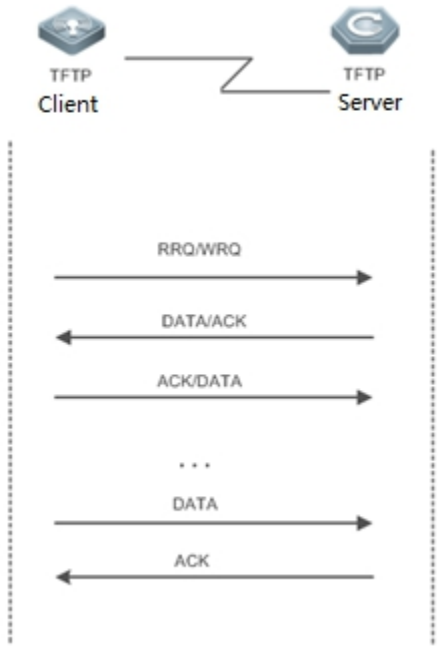
- Read Request (RRQ)
- Write Request (WRQ)
- DATA
- Acknowledgment (ACK)
- ERROR

Figure 9- 2



Working Principle

Figure 9- 3



- The TFTP client initiates an RRQ or WRQ to the TFTP server.
- Upon receipt of the RRQ, the TFTP server first determines whether the read condition is met (for example, whether the file exists or whether the client has the access permission), and returns a DATA packet to the TFTP client if yes; upon receipt of the WRQ, the TFTP server first determines whether the write condition is met (for example, whether there is a sufficient space or whether the client has the write permission), and returns an ACK packet to the TFTP client if yes.
- The TFTP client receives the DATA packet in the case of file downloading, and replies with an ACK packet; or receives the ACK packet in the case of file uploading, and then sends a DATA packet.
- The process of transmission acknowledgement repeats till the last DATA packet is less than 512 bytes, which indicates the end of the transmission.
- If errors occur during the transmission, an ERROR packet is returned.

9.3.1 Enabling the TFTP Service

Working Principle

The working principle of TFTP is as described in the previous chapter. After the TFTP service is enabled on the device, configure a top directory so that the TFTP service is available for users.

Related Configuration

Enabling the TFTP Service

- By default, the TFTP service is disabled.
- Run the **tftp-server enable** command to enable the TFTP service.
- You must enable the TFTP service; otherwise, the TFTP server is unavailable.

Configuring the Top Directory

- By default, no top directory is configured.

- Run the **tftp-server topdir** command to configure the top directory.
- You must configure the top directory; otherwise, file uploading and downloading functions are unavailable.

9.4 Configuration

Configuration	Description and Command	
Configuring the Basic Functions of the TFTP Service	Mandatory configuration, which is used to enable the TFTP service.	
	tftp-server enable	Enables the TFTP service.
	tftp-server topdir	Configures the top directory of the TFTP server.

9.4.1 Basic Functions

Networking Requirements

- Establish a TFTP server to provide the TFTP client with uploading and downloading functions.

Configuration Tips

- Top directory configuration is required.

Configuration Steps

Enabling the TFTP Service

- Mandatory configuration.
- Enable the TFTP service on each device unless otherwise stated.

Configuring the Top Directory

- Mandatory configuration.
- Configure a top directory as the root directory on each device unless otherwise stated.

Verification

Connect the TFTP server to the TFTP client.

- Check whether the client is connected to the server.
- Check whether the client can normally download files from and upload files to the server.

Related Commands

Enabling the TFTP Service

Command	tftp-server enable
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	The client cannot access the TFTP server before a top directory is correctly configured for the server. Therefore, it is recommended that you configure the top directory of the server first if it is the first time for you to enable the TFTP

	server. For details about how to configure the top directory, see the description to immediately follow below.

Configuring the Top Directory of the TFTP Server

Command	**tftp-servertopdir** *directory*
Parameter Description	*directory*: access path
Command Mode	Global configuration mode
Usage Guide	For example, you can set the top directory of the server to **/dir**. Then the TFTP client can access files and folders in only the **/dir** directory on the device after logging in, and the TFTP client cannot return to the parent directory of the **/dir** directory due to the restrictions of the top directory.

Enabling the TFTP Server Debugging Switch

Command	**debug tftp-server**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	You can run this command to enable the TFTP server debugging switch, so that the process or error information of the TFTP server can be output as necessary.

Configuration Example

Establishing the TFTP Service on an IPv4 Network

Scenario	● Enable the TFTP service. ● Set the top directory of the TFTP server to /dir.
	FS(config)#tftp-server topdir /dir FS(config)#tftp-server enable
Verification	● Run the **show run** command to display the configuration. FS#show run … tftp-server enable tftp-server topdir /dir …

Common Errors

No top directory is configured.

9.5 Monitoring

Displaying

System resources are occupied when debugging information is output. Therefore, disable the debugging switch immediately after use.

Function	Command
Enables the TFTP server debugging switch.	**debug tftp-server**

10 Configuration Tunnel

10.1 Overview

Tunnel interfaces are used to implement the tunneling function and are virtual interfaces of the system. tunnel interfaces do not specifically specify a transport protocol or load protocol, they provide a transport link used to implement a standard transport link. Each Tunnel interface represents one transport link.

Tunnel function implementation consists of three main components：

- Load Protocol: The encapsulation protocol for the load (network data) transmitted through Tunnel. If IPv4 and IPv6 protocols are used as load protocols, for GRE tunnels, they can be either IPv4 or IPv6 protocols.
- Carrier protocol: the protocol used to encapsulate and identify the load to be transmitted in a secondary way. In the TUNNEL described in this paper, only GRE tunnels have a carrier protocol, i.e., the GRE protocol (i.e., encapsulation protocol). All other types of tunnels are IPv4/IPv6 outside the direct encapsulation of IPv4/IPv6.
- Transport protocol: The network protocol that actually transmits the load after the secondary encapsulation of the carrier protocol. Riptide products use the most widely used IPv4 and IPv6 protocols as transport protocols.

In fact, if two private networks of the same protocol need to realize intercommunication through a heterogeneous public network, they can be realized by using Tunnel mode.

Tunnel transport is suitable for the following cases：

- Allowing communication between local networks running non-IP protocols over a single network (IP network), since Tunnel supports many different load protocols; allowing protocols that have a limit on the number of routing hops to work over a wider range, since Tunnel works with Transport Protocol (IP) routing.
- Allowing intermittent subnets to be connected through a single network (IP network).
- Allows for VPN (Virtual Private Network) functionality over WANs.

Since Tunnel encapsulates the load and then transmits it, this introduces processing complexity and in some cases the following changes need to be noted.

- Since Tunnel is a logical link, it may appear to be a single hop when routing, but in reality it may take more than one hop to route. When using Tunnel, it is important to note that the routing of the Tunnel link does not match the actual route.
- Since Tunnel transport encapsulates the load in the transport protocol, this needs to be taken into account when setting up firewalls, especially access control chains (ACLs); it is also important to note that the transmission bandwidth of the load protocol at this time (e.g. MTU) is smaller than the theoretical value.

The following section only describes the content of Tunnel.

Protocol Specification

- RFC2784：Generic Routing Encapsulation (GRE)
- RFC2890：Key and Sequence Number Extensions to GRE
- RFC3056：Connection of IPv6 Domains via IPv4 Clouds
- RFC3068：An Anycast Prefix for 6to4 Relay Routers
- RFC3964：Security Considerations for 6to4
- RFC4023：Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)

- RFC4087：IP Tunnel MIB
- RFC4213：Basic Transition Mechanisms for IPv6 Hosts and Routers
- RFC4797： Use of Provider Edge to Provider Edge (PE-PE) Generic Routing Encapsulation (GRE) or IP in BGP/MPLS IP Virtual Private Networks
- RFC5158：6to4 Reverse DNS Delegation Specification
- RFC5214：Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)
- RFC5332：MPLS Multicast Encapsulations
- RFC5579：Transmission of IPv4 Packets over Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) Interfaces
- RFC5845：Generic Routing Encapsulation (GRE) Key Option for Proxy Mobile IPv6
- RFC5969：IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) -- Protocol Specification
- RFC6245：Generic Routing Encapsulation (GRE) Key Extension for Mobile IPv4
- RFC6343：Advisory Guidelines for 6to4 Deployment
- RFC6372：6to4 Provider Managed Tunnels
- RFC6654：Gateway-Initiated IPv6 Rapid Deployment on IPv4 Infrastructures (GI 6rd)
- draft-zhou-dhc-gre-option-00 DHCPv4 and DHCPv6 options for GRE
- draft-cai-softwire-6rd-mib-03 Definitions of Managed Objects for 6rd
- draft-howard-isp-ip6rdns-05 Reverse DNS in IPv6 for Internet Service Providers
- draft-tsou-softwire-6rd-multicast-02 IPv6 Multicast Using Native IPv4 Capabilities in a 6rd Deployment
- draft-templin-v6ops-isops-18 Operational Guidance for IPv6 Deployment in IPv4 Sites using ISATAP

10.2 Typical Applications

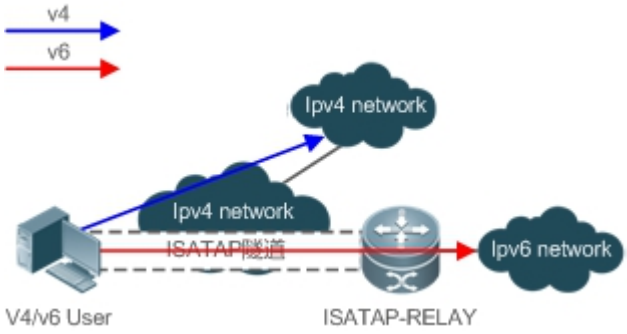
Typical Applications	Scenario Description
Accessing an IPv6 site inside the campus network	Accessing an IPv6 site within the campus network
Campus network access to IPv6 backbone network	The campus network is connected to the IPv6 backbone.

Accessing an IPv6 site inside the campus network

Typical Applications

Some campus networks have IPv6 servers deployed and PCs need to be able to access these servers. This application can be implemented using ISATAP automatic tunneling.

Figure 10- 1



NOTE ISATAP-RELAY is a tunnel-enabled device. When campus network users access IPv4 servers, they can directly access through IPv4 network; when they access IPv6 servers, they need to access through ISATAP tunnel.

Function Deployment

- V4/V6 users access the IPv4 network through V4 addresses.
- V4/V6 users access the IPv6 network through ISATAP tunnels.
- ISATAP tunnels are established between the PC and the ISATAP-RELAY router.

Campus Network Access to IPv6 Backbone Network

Application Scenarios

Upgrading an old campus network to support IPv6, or creating a new campus network that supports IPv6. This can be achieved using 6to4 automatic tunneling.

10.3 Function Details

Basic Concept

- Tunnel MTU

The MTU (Maximum Transmission Unit) of a tunnel interface is generally lower than that of a normal interface. For example, for a normal IPv4 over IPv4 tunnel, assuming transmission over Ethernet, the actual maximum amount of data that can be transmitted per message, excluding the Layer 2 IP header, is 1460 bytes. So the MTU of a tunnel interface is 1480 bytes instead of the 1500 bytes of a standard Ethernet interface.

- Tunnel Path MTU

A tunnel interface actually represents a virtual link. Take an IPv4 over IPv4 tunnel as an example. The MTU of this tunnel interface is actually the MTU of the tunnel link, i.e. the path MTU between the local end and the peer end of the tunnel. when the routing path between the local end and the peer end changes, the path MTU may change, thus affecting the MTU of the tunnel interface

- Tunnel Nesting Levels

Tunnels can be nested. Take an IPv4 over IPv4 tunnel as an example. The original IPv4 message can be encapsulated with two or more layers of IPv4 prefixes before being sent out. The scenario where tunnel nesting is required is generally uncommon, mainly for security reasons. When tunnels are nested, less data can be transmitted. After the opposite address of a tunnel interface is routed by FIB, if the exit is another tunnel interface, tunnel nesting encapsulation will occur; if the exit is this tunnel again after routing, then infinite nesting will occur (Riptide products can detect this situation and give hints).

Functional features

Functional Characteristics	Function

Tunnel nesting encapsulation layer limit	Increase the tunnel nesting layer limit according to actual needs (rarely used).
Tunnel data validity check	Add checksum to the data transmitted by GRE tunnels to check if errors occur during data transmission.

10.4 Configuration details

Configuration items	Configuration Suggestions & Related Commands	
Configure basic tunnel interface functions	Must be configured. Used to establish a tunnel.	
	Interface tunnel	Creates a new tunnel interface
	tunnel source	Sets the local address of the tunnel
Configure the tunnel mode	Optional configuration to set the tunnel mode	
	tunnel mode	Configure the tunnel encapsulation mode
Configure the address of the opposite end of the tunnel	Optional configuration to set the tunnel destination address	
	tunnel destination	Configure the tunnel destination address
Configure the tunnel transport network TOS	Optional configuration to set the TOS of the transport network	
	tunnel tos	Configure the tunnel transport network TOS
Configure the tunnel transport network TTL	Optional configuration to set the TTL of the transport network	
	tunnel ttl	Configures the TTL of the transport network

10.4.1 Configuring the tunnel interface

Configuration Effect

- A new tunnel interface is created.

Caution

None

Configuration Method

New tunnel interface function

- In global configuration mode, use interface tunnel number to create a new tunnel interface.
- Only when a tunnel interface is created can the tunnel service be used.

Verification method

- Use show interfaces tunnel number to see if the configuration is successful.

Related commands

Configure a tunnel interface

Command Fomat	interface tunnel number
Parameter Description	number：Tunnel interface number。
Command mode	Global Mode
Usage guide	-

Check the tunnel configuratio

Command Fomat	show interfaces tunnel number
Parameter Description	number：Tunnel interface number。
Command mode	Global Mode
Usage guide	-

Configuration examples

Create a tunnel interface

Configuration method	Create a tunnel interface FS# configure terminal FS(config)# interface tunnel 1 FS(config-if-Tunnel 1)# end
Configuration method	Check tunnel interface configuration FS# show interfaces tunnel 1 …… Tunnel attributes: …… Tunnel protocol/transport is gre ip

Common Errors

- Insufficient memory causes the tunnel interface creation to fail.
- Insufficient hardware chip resources cause the tunnel interface creation to fail.

10.4.2 Configuring tunnel encapsulation mode

Configuration Effect

- When you need to use a tunnel with a non-default encapsulation mode
- Configure in tunnel interface mode

Caution

None

Configuration method

Sets the encapsulation mode of the tunnel

- Optional configuration
- The default encapsulation mode for the switch and wireless products is tunnel mode ipv6ip.
- The default encapsulation mode for router, gateway products is tunnel mode gre ip.
- The tunnel mode command allows you to change the tunnel mode to a different mode in the tunnel interface mode.

Configuration method

Use show interfaces tunnel number to observe whether the configuration is successful.

Related commands

Configuring Tunnel Encapsulation Mode

Command Fomat	tunnel mode { gre {ip | ipv6} | ipv6| ipip | ipv6ip [6to4 | isatap] }
Parameter Description	Each mode represents the encapsulation format of the messages sent from this tunnel interface gre ip means that it goes through a GRE header encapsulation, then an IPv4 encapsulation, and then is sent on the new IPv4 network gre ipv6 means that the message is first encapsulated in a GRE header, then in an IPv6 header, and then sent on the new IPv6 network ipv6 means that messages from this interface are sent on the new IPv6 network directly after IPv6 encapsulation ipip means that this interface can only carry IPv4 packets, and the packets are then encapsulated in IPv4 and sent on the new IPv4 network ipv6ip means that the interface can only carry IPv6 messages, which are encapsulated in IPv4 and sent on the new IPv4 network All the above tunnels are manual tunnels, while ipv6ip 6to4/isatap is an automatic tunnel where the destination IPv4 address is mapped through the destination IPv6 address during the packet encapsulation process.
Command mode	Interface mode
Usage guide	The 2 ends of the tunnel generally need to be configured with the same mode, otherwise they will not work properly.

Configuration examples

Configure the encapsulation mode of the tunnel to IPv4 over IPv4.

Configuration method	Configure the encapsulation mode of the tunnel interface to IPv4 over IPv4.

```
FS# configure terminal
FS(config)# interface tunnel 1
FS(config-if-Tunnel 1)# tunnel mode ipip
FS(config)# end
```

Configuration method	Check tunnel interface configuration FS# show interfaces tunnel 1 …… Tunnel attributes: …… Tunnel protocol/transport is ipv6ip

Common errors

- The same vrf already has a 6to4 or isatap tunnel, continue to configure these 2 tunnels.

10.4.3 Configure the tunnel local address

Configuration effect

- Set the local address of the tunnel.

Caution

- The local address of the tunnel must match the transport protocol of the tunnel. Otherwise, the tunnel port will not UP (be disabled).
- When specifying the local address indirectly by specifying another interface, it refers to the IPv4 home address or the IPv6 first global public address.

Configuration method

Set the local address of the tunnel

- Must be configured
- Use the tunnel source command in tunnel interface mode to specify the local address of the tunnel.

Configuration method

- The show interfaces tunnel number command allows you to observe the local address of the tunnel

Related commands

Configure the tunnel local address

Command Fomat	tunnel source { ip-address | interface-name interface-number }
Parameter Description	ip-address: can be an IPv4 address or an IPv6 address. interface-name interface-number: can be various Layer 3 interfaces.

Command mode	Interface mode
Usage guide	If you specify an IPv4 or IPv6 address directly, you need to set it to the device's own address

Configuration examples

Configuration	Tunnel local address
Configuration method	Configure the local address of the tunnel to 1.1.1.1

```
FS# configure terminal
FS(config)# interface tunnel 1
FS(config-if-Tunnel 1)# tunnel mode ipv6ip
FS(config-if-Tunnel 1)# tunnel source 1.1.1.1
```

Configuration method	Check Tunnel interface Configuration

```
FS# show interfaces tunnel 1
……
Tunnel attributes:
  Tunnel source 1.1.1.1, destination UNKNOWN, unrouteable
  Tunnel TOS/Traffic Class not set, Tunnel TTL 254
  Tunnel config nested limit is 0, current nested number is 0
  Tunnel protocol/transport ipv6ip
  Tunnel transport VPN is no set
……
```

Common configuration errors

None

10.4.4 Configuring the tunnel peer address

Configuration Effect

- Manual tunnels, which must be configured with a peer address before they can be used (the interface will only UP).

Caution

- Automatic tunnels cannot be configured with a peer address.

Configuration method

- Configure the end address of the tunnel
- 6to4, isatap tunnels cannot be configured, other tunnels must be configured
- Configure the destination address in interface mode with the tunnel destination command.

Configuration method

- The show interfaces tunnel command allows you to see if the configuration is successful.

Related commands

Configuring the peer address

Command Fomat	tunnel destination { ip-address }
Parameter Description	ip-address: can be an IPv4 address or an IPv6 address.
Command mode	Interafce mode
Usage guide	The manual tunnel must be configured with a peer address. The protocol family type of the configured peer address must be the same as the transport protocol, otherwise the tunnel interface is disabled (no UP).

Configuration examples

Configure the address of the tunnel peer.

Configuration method	Configure the tunnel peer address as 2.2.2.2 FS# configure terminal FS(config)# interface tunnel 1 FS(config-if-Tunnel 1)# tunnel mode ipv6ip FS(config-if-Tunnel 1)# tunnel destination 2.2.2.2
Configuration method	Check Tunnel interface Configuration FS# show interfaces tunnel 1 …… Tunnel attributes: Tunnel source: UNKNOWN, destination 2.2.2.2, unrouteable Tunnel TOS/Traffic Class not set, Tunnel TTL 254 Tunnel config nested limit is 0, current nested number is 0 Tunnel protocol/transport ipv6ip ……

Common Errors

- Configuring a peer address for an autotunnel.
- Same butt-end address as another tunnel.

10.4.5 Configuring the tunnel transport network TOS

Configuration Effect

- Specify the tos or traffic class field in the transport protocol header.

Caution

- If not specified, the tos and traffic class fields corresponding to the bearer protocol are copied.

Configuration method

Configure the transport network TOS

- Optional configuration
- If you need to revise the priority of tunnel data in the network
- Use the tunnel tos command in interface mode

Configuration method

- Use the show interfaces tunnel command to see if the settings are successful.

Related commands

Configuring Transport Network TOS

Command Fomat	tunnel tos number
Parameter Description	number：The value of tos to be set。
Command mode	Interafce mode
Usage guide	By default, if the tunnel inner bearer and outer encapsulation are both IPv4 protocols, the tos bytes of the inner IPv4 header are copied to the outer IPv4 header by default. If the tunnel inner bearer and outer encapsulation are both IPv6 protocols, the traffic class8 bits of the inner IPv6 header are copied to the outer IPv6 header by default. In other cases, the outer IPv4 tos / IPv6 traffic class,is 0.

Configuration examples

Configure the transport network TOS.

Configuration method: Configure the transport network TOS.

```
FS# configure terminal
FS(config)# interface tunnel 1
FS(config-if-Tunnel 1)# tunnel mode ipv6ip
FS(config-if-Tunnel 1)# tunnel tos 2
```

Configuration method: Check Tunnel interface Configuration

```
FS# show interfaces tunnel 1
……
Tunnel attributes:
  Tunnel source 1.1.1.1, destination UNKNOWN, unrouteable
  Tunnel TOS/Traffic Class 0x2, Tunnel TTL 254
  Tunnel config nested limit is 0, current nested number is 0
  Tunnel protocol/transport ipv6ip
  Tunnel transport VPN is VPN1
……
```

Common Errors

None

10.4.6 Configure tunnel transmission network TTL

Configuration Effect

- Specify the value of TTL or hoplimit for the tunnel encapsulation protocol header.

Caution

None

Configuration method

Configure the transport network TTL

- Optional configuration
- Default 254
- If you need to change the length of the tunnel link to a specific value
- use the tunnel ttl command to set it.

Configuration method

- Use the show interfaces tunnel command to see if the configuration is successful.

Related commands

Configure the transport network TTL

Command Fomat	tunnel ttl hop-limit
Parameter Description	hop-limit：Hop limit of transmission network。
Command mode	Global Mode
Usage guide	The default is 254. If you need to change it to a specific value, you can use this command.

Configuration examples

Configure transmission network TTL

Configuration method	Configure transmission network TTL
	FS# configure terminal FS(config)# interface tunnel 1 FS(config-if-Tunnel 1)# tunnel mode ipv6ip FS(config-if-Tunnel 1)# tunnel ttl 3
Configuration method	Check Tunnel interface Configuration

Configuration method	Configure transmission network TTL
	FS# configure terminal FS(config)# interface tunnel 1 FS(config-if-Tunnel 1)# tunnel mode ipv6ip FS(config-if-Tunnel 1)# tunnel ttl 3
Configuration method	Check Tunnel interface Configuration
	FS# show interfaces tunnel 1 …… Tunnel attributes: Tunnel source 1.1.1.1, destination UNKNOWN, unrouteable Tunnel TOS/Traffic Class 0x2, Tunnel TTL 3 Tunnel config nested limit is 0, current nested number is 0 Tunnel protocol/transport ipv6ip Tunnel transport VPN is VPN1

Common Errors

None

10.5 Monitoring

Clear all kinds of information

None

View operation

Role	Command
View information about the tunnel interface.	show interfaces tunnel number

Viewing debugging information

Outputting debugging information will occupy system resources. Please turn off the debug switch immediately after you finish using it.

Role	Command
Turn on the debug switch for the tunnel interface	debug tunnel

11 Configuring Network Communication Test Tools

11.1 Overview

Network communication test tools can be used to check the connectivity of a network and helps you analyze and locate network faults. Network communication test tools include Packet Internet Groper (PING) and Traceroute. Ping is used to check the connectivity and delay of a network. A greater delay indicates a slower network speed. Traceroute helps you learn about the topology of physical and logical links and transmission rate. On a network device, you can run the **ping** and **traceroute** commands to use the two tools respectively.

Protocols and Standards

- RFC792: Internet Control Message Protocol
- RFC4443: Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

11.2 Applications

Application	Description
End-to-End Connectivity Test	Both the network device and the destination host are connected to the IP network and configured with IP addresses.
Host Route Test	Both the network device and the destination host are connected to the IP network and configured with IP addresses.

11.2.1 End-to-End Connectivity Test

Scenario

As shown in Figure 10- 1, Network Device A and Target Host B are connected to the IP network.

If both the network device and the target host are connected to the IP network, the end-to-end connectivity test aims to check whether IP packets can be transmitted between the two ends. The target host can be the network device itself. In this case, the connectivity test aims to check the network interface and TCP/IP configurations on the device.

Figure 10- 1

Deployment

Execute the ping function on the network device.

11.2.2 Host Route Test

Scenario

As shown in Figure 10- 2, Network Device A and Target Host B are connected to the IP network.

If both the network device and the target host are connected to the IP network, the host route test aims to check gateways (or routers) that IP packets pass through between the two ends. Generally, the target host is not within the same IP network segment as the network

device.

Figure 10- 2

Deployment

Execute the traceroute function on the network device.

11.3 Features

Overview

Feature	Description
Ping Test	Test whether the specified IPv4 or IPv6 address is reachable and display the related information.
Traceroute Test	Display the gateways that IPv4 or IPv6 packets pass through when transmitted from the source to the destination.

11.3.1 Ping Test

Working Principle

The ping tool sends an Internet Control Message Protocol (ICMP) Request message to the destination host to request the for an ICMP Echo Reply message. In this way, the ping tool determines the delay and the connectivity between the two network devices.

Related Configuration

- Run the **ping** command.

11.3.2 Traceroute Test

Working Principle

The traceroute tool uses the Time To Live (TTL) field in the headers of the ICMP and IP messages for the test First, the traceroute tool on the network device sends an ICMP Request message with TTL 1 to the destination host. After receiving the message, the first router on the path decreases the TTL by 1. As the TTL becomes 0, the router drops the packets and returns an ICMP time exceeded message to the network device. After receiving this message, the traceroute tool learns that this router exists on this path, and then sends an ICMP Request packet with TTL 2 to the destination host to discover the second router. Each time the traceroute tool increases the TTL in the ICMP Request message by 1 to discover one more router. This process is repeated until a data packet reaches the destination host. After the packet reaches the destination host, the host returns an ICMP Echo message instead of an ICMP time exceeded message to the network device. Then, the traceroute tool finishes the test and displays the path from the network device to the destination host.

Related Configuration

- Run the **traceroute** command.

11.4 Configuration

Configuration	Description and Command	
Ping Test	(Optional) It is used to check whether an IPv4 or IPv6 address is reachable.	
	ping	Executes the Ping function.
Traceroute Test	(Optional) It is used to display the gateways that IPv4 or IPv6 packets pass through when transmitted from the source to the destination.	
	traceroute	Executes the traceroute function.

11.4.1 Ping Test

Configuration Effect

After conducting a ping test on a network device, you can learn whether the network device is connected to the destination host and whether packets can be transmitted between the network device and the destination host.

Notes

The network device must be configured with an IP address.

Configuration Steps

- To check whether an IPv4 address is reachable, use the **ping IPv4** command.
- To check whether an IPv6 address is reachable, use the **ping IPv6** command.

Verification

Run the **ping** command to display related information on the command line interface (CLI) window.

Related Commands

Ping IPv4

Command	**ping** [**ip**] { *address*] [**length** *length*] [**ntimes** *times*] [**timeout** *seconds*] [**data** *data*] [**source** *source*] [**df-bit**] [**validate**] [**detail**] [**interval** millisecond] }
Parameter Description	*address*: Specifies the destination IPv4 address or domain name. *length*: Specifies the length of the data packet. The value ranges from 36 to 18,024. The default length is 100. *times*: Specifies the number of probes. The value ranges from 1 to 4,294,967,295 *seconds*: Specifies the timeout. The value ranges from 1s to 10s. *data*: Specifies the data in the packet. The data is a string of 1 to 255 bytes. By default, the string is "abcd". *source*: Specifies the source IPv4 address or source port of the packet. The loopback interface address, for example, 127.0.0.1, cannot be used as the source address. **df-bit**: Configures the DF bit of the IP address. When the DF bit is set to 1, the packet is not fragmented. By default, the DF bit is 0. **validate**: Configures whether to verify the response packet. **detail**: Configures whether to display the Echo Reply message in detail. By default, only the exclamation mark (!) and dot (.) are displayed. millisecond: Specifies the interval at which the ping packet is sent. The value ranges from 10 ms to 300,000 ms. The

	default interval is 100 ms.
Command Mode	In User EXEC mode, you can execute only the basic ping function. In Privileged EXEC mode, you can execute the extended ping function. In other configuration modes, you can run the **do** command to execute the extended ping function. For details about the configuration, see the description about the **do** command.
Configuration Usage	When the ping function is executed, information about the response (if any) will be displayed, and then related statistics will be output. Using the extended ping function, you can specify the number, length and timeout of packets to be sent. Like the basic ping function, related statistics will be output. To use the domain name, you must first configure the domain name server (DNS). For details about the configuration, see *Configuring DNS*.

Ping IPv6

Command	**ping** [**ipv6**] [*address* [**length** *length*] [**ntimes** *times*] [**timeout** *seconds*] [**data** *data*] [**source** *source*] [**df-bit**] [**validate**] [**detail**] [**interval** millisecond]]
Parameter Description	*address*: Specifies the destination IPv6 address or domain name. *length*: Specifies the length of data packet. The value ranges from 16 to 18,024. The default length is 100. *times*: Specifies the number of probes. The value ranges from 1 to 4,294,967,295. *seconds*: Specifies the timeout. The value ranges from 1s to 10s. *data*: Specifies the data in the packet. The data is a string of 1 to 255 bytes. *source*: Specifies the source IPv6 address or source port of the packet. The loopback interface address, for example, ::1, cannot be used as the source address. **Detail**: Configures whether to display the Echo Reply message in detail. By default, only the exclamation mark (!) and dot (.) are displayed. millisecond: Specifies the interval at which the ping packet is sent. The value ranges from 10 ms to 300,000 ms. The default interval is 100 ms.
Command Mode	In User EXEC mode, you can execute only the basic ping IPv6 function. In Privileged EXEC mode, you can execute the extended ping IPv6 function. In other configuration modes, you can run the **do** command to execute the extended ping function. For details about the configuration, see the description about the **do** command.
Configuration Usage	When the ping IPv6 function is executed, information about the response (if any) will be displayed, and then related statistics will be output. Using the extended ping IPv6 function, you can specify the number, length and timeout of packets to be sent. Like the basic ping IPv6 function, related statistics will be output. To use the domain name, you must first configure the DNS. For details about the configuration, see *Configuring DNS*.

Configuration Example

Executing the Common Ping Function

Configuration Steps	In Privileged EXEC mode, run the **ping 192.168.21.26** command.
	Common ping command: FS# ping 192.168.21.26

	Sending 5, 100-byte ICMP Echoes to 192.168.21.26, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/10 ms Detailed ping command: FS#ping 192.168.21.26 detail Sending 5, 100-byte ICMP Echoes to 192.168.21.26, timeout is 2 seconds: < press Ctrl+C to break > Reply from 192.168.21.26: bytes=100 time=4ms TTL=64 Reply from 192.168.21.26: bytes=100 time=3ms TTL=64 Reply from 192.168.21.26: bytes=100 time=1ms TTL=64 Reply from 192.168.21.26: bytes=100 time=1ms TTL=64 Reply from 192.168.21.26: bytes=100 time=1ms TTL=64 Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/4 ms.
Verification	Send five 100-byte packets to the specified IP address, and the response information will be displayed in the specified time (2s by default). Finally the statistics is output.

Executing the Extended Ping Function

Configuration Steps	In Privileged EXEC mode, run the **ping 192.168.21.26** command. In addition, specify the length, number, and timeout of the packets.
	Common ping command: FS# ping 192.168.21.26 length 1500 ntimes 100 data ffff source 192.168.21.99 timeout 3 Sending 100, 1500-byte ICMP Echoes to 192.168.21.26, timeout is 3 seconds: < press Ctrl+C to break > !! Success rate is 100 percent (100/100), round-trip min/avg/max = 2/2/3 ms Detailed ping command: ping 192.168.21.26 length 1500 ntimes 20 data ffff source 192.168.21.99 timeout 3 detail Sending 20, 1500-byte ICMP Echoes to 192.168.21.26, timeout is 3 seconds: < press Ctrl+C to break > Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64

	Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=2ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=3ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64 Success rate is 100 percent (20/20), round-trip min/avg/max = 1/1/3 ms.
Verification	Send twenty 1,500-byte packets to the specified IP address, and the response information (if any) will be displayed in the specified time (3s by default). Finally the statistics is output.

↘ Executing the Common Ping IPv6 Function

Configuration Steps	In Privileged EXEC mode, run the **ping ipv6 2001::1** command.
	Common ping command: FS# ping ipv6 2001::1 Sending 5, 100-byte ICMP Echoes to 2001::1, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/10 ms Detailed ping command: FS#ping 2001::1 detail Sending 5, 100-byte ICMP Echoes to 2001::1, timeout is 2 seconds:

	< press Ctrl+C to break > Reply from 2001::1: bytes=100 time=1ms Reply from 2001::1: bytes=100 time=1ms Reply from 2001::1: bytes=100 time=1ms Reply from 2001::1: bytes=100 time=1ms Reply from 2001::1: bytes=100 time=1ms Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms.
Verification	Send five 100-byte packets to the specified IP address, and the response information will be displayed in the specified time (2s by default). Finally the statistics is output.

➘ Executing the Extended Ping IPv6 Function

Configuration Steps	In Privileged EXEC mode, run the **ping ipv6 2001::5** command. In addition, specify the length, number, and timeout of the packets.
	Common ping command: FS# ping ipv6 2001::5 length 1500 ntimes 100 data ffff source 2001::9 timeout 3 Sending 100, 1500-byte ICMP Echoes to 2000::1, timeout is 3 seconds: < press Ctrl+C to break > !! Success rate is 100 percent (100/100), round-trip min/avg/max = 2/2/3 ms Detailed ping command: FS#ping 2001::5 length 1500 ntimes 10 data ffff source 2001::9 timeout 3 Sending 10, 1500-byte ICMP Echoes to 2001::5, timeout is 3 seconds: < press Ctrl+C to break > Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms Reply from 2001::5: bytes=1500 time=1ms

	Success rate is 100 percent (10/10), round-trip min/avg/max = 1/1/1 ms.
Verification	Send one hundred 1500-byte packets to the specified IPv6 address, and the response information (if any) will be displayed in the specified time (3s by default). Finally the statistics is output.

11.4.2 Traceroute Test

Configuration Effect

After conducting a traceroute test on a network device, you can learn about the routing topology between the network device and the destination host, and the gateways through which packets are sent from the network device to the destination host.

Notes

The network device must be configured with an IP address.

Configuration Steps

- To trace the route an IPv4 packet would follow to the destination host, run the **traceroute IPv4** command.
- To trace the route an IPv6 packet would follow to the destination host, run the **traceroute IPv6** command.

Verification

Run the **traceroute** command to display related information on the CLI window.

Related Commands

Traceroute IPv4

Command	**traceroute** [**ip**] { *adress* [**probe** *number*] [**source** *source*] [**timeout** *seconds*] [**ttl** *minimum maximum*] }
Parameter Description	*address*: Specifies the destination IPv4 address or domain name. *number*: Specifies the number of probes. The value ranges from 1 to 255. *source*: Specifies the source IPv4 address or source port of the packet. The loopback interface address, for example, 127.0.0.1, cannot be used as the source address. *seconds*: Specifies the timeout. The value ranges from 1s to 10s. *minimum maximum*: Specifies the minimum and maximum TTL values. The value ranges from 1 to 255.
Command Mode	In User EXEC mode, you can execute only the basic traceroute function. In privileged EXEC mode, you can execute the extended traceroute function.
Configuration Usage	The **traceroute** command is used to test the network connectivity and accurately locate a fault when the fault occurs. To use the domain name, you must first configure the DNS. For details about the configuration, see *Configuring DNS*.

Traceroute IPv6

Command	**traceroute** [**ipv6**] { *address* [**probe** *number*] [**timeout** *seconds*] [**ttl** *minimum maximum*] }
Parameter Description	*address*: Specifies the destination IPv6 address or domain name. *number*: Specifies the number of probes. The value ranges from 1 to 255. *seconds*: Specifies the timeout. The value ranges from 1s to 10s. *minimum maximum*: Specifies the minimum and maximum TTL values. The value ranges from 1 to 255.

Command Mode	In User EXEC mode, you can execute only the basic traceroute IPv6 function. In privileged EXEC mode, you can execute the extended traceroute IPv6 function.
Configuration Usage	The traceroute IPv6 command is used to test the network connectivity and accurately locate a fault when the fault occurs. To use the domain name, you must first configure the DNS. For details about the configuration, see *Configuring DNS*.

Configuration Example

Executing the Traceroute Function on a Properly Connected Network

Configuration Steps	In Privileged EXEC mode, run the **traceroute ipv6 3004::1** command.
	FS# FS# traceroute ipv6 3004::1 < press Ctrl+C to break > Tracing the route to 3004::1 1 3000::1 0 msec 0 msec 0 msec 2 3001::1 4 msec 4 msec 4 msec 3 3002::1 8 msec 8 msec 4 msec
	The preceding test result indicates that the network device accesses host 3004::1 by transmitting packets through gateways 1–3. In addition, the time required to reach each gateway is displayed.

Executing the Traceroute Function on a Faulty Network

Configuration Steps	In Privileged EXEC mode, run the **traceroute 202.108.37.42** command.

```
FS# traceroute 202.108.37.42
< press Ctrl+C to break >
Tracing the route to 202.108.37.42
1        192.168.12.1        0 msec    0 msec    0 msec
2        192.168.9.2         0 msec    4 msec 4 msec
3        192.168.110.1      16 msec   12 msec   16 msec
4        *    *    *
5        61.154.8.129       12 msec   28 msec   12 msec
6        61.154.8.17         8 msec   12 msec   16 msec
7        61.154.8.250       12 msec   12 msec   12 msec
8        218.85.157.222     12 msec   12 msec   12 msec
9        218.85.157.130     16 msec   16 msec   16 msec
10       218.85.157.77      16 msec   48 msec   16 msec
11       202.97.40.65       76 msec   24 msec   24 msec
12       202.97.37.65       32 msec   24 msec   24 msec
13       202.97.38.162      52 msec   52 msec   224 msec
14       202.96.12.38       84 msec   52 msec   52 msec
15       202.106.192.226    88 msec   52 msec   52 msec
16       202.106.192.174    52 msec   52 msec   88 msec
17       210.74.176.158    100 msec   52 msec   84 msec
18       202.108.37.42      48 msec   48 msec   52 msec
```

The preceding test result indicates that the network device accesses host 202.108.37.42 by transmitting packets through gateways 1–17, and Gateway 4 is faulty.

Executing the Traceroute IPv6 Function on a Properly Connected Network

Configuration Steps	In Privileged EXEC mode, run the **traceroute ipv6 3004::1** command.

```
FS# traceroute ipv6 3004::1
  < press Ctrl+C to break >
Tracing the route to 3004::1
1        3000::1        0 msec    0 msec    0 msec
2        3001::1        4 msec    4 msec 4 msec
3        3002::1        8 msec    8 msec    4 msec
4        3004::1        4 msec   28 msec   12 msec
```

	The preceding test result indicates that the network device accesses host 3004::1 by transmitting packets through gateways 1-4. In addition, the time required to reach each gateway is displayed.

Executing the Traceroute IPv6 Function on a Faulty Network

Configuration Steps	In Privileged EXEC mode, run the **traceroute ipv6 3004::1** command.
	FS# traceroute ipv6 3004::1 < press Ctrl+C to break > Tracing the route to 3004::1 1 3000::1 0 msec 0 msec 0 msec 2 3001::1 4 msec 4 msec 4 msec 3 3002::1 8 msec 8 msec 4 msec 4 * * * 5 3004::1 4 msec 28 msec 12 msec
	The preceding test result indicates that the network device accesses host 3004::1 by transmitting packets through gateways 1–5, and Gateway 4 is faulty.

12 Configuring TCP

12.1 Overview

The Transmission Control Protocol (TCP) is a transport-layer protocol providing reliable connection-oriented and IP-based services to for the application layer.

Internetwork data flows in 8-bit bytes are sent from the application layer to the TCP layer, and then fragmented into packet segments of a proper length via the TCP. The Maximum Segment Size (MSS) is usually limited by the Maximum Transmission Unit (MTU) of the data link layer. After that, the packets are sent to the IP layer and then to the TCP layer of a receiver through the network.

To prevent packet loss, every byte is identified by a sequence number via the TCP, and this ensures that packets destined for the peer are received in order. Then, the receiver responds with a TCP ACK packet upon receiving a packet. If the sender does not receive ACK packets in a reasonable Round-Trip Time (RTT), the corresponding packets (assumed lost) will be retransmitted.

- TCP uses the checksum function to check data integrity. Besides, MD5-based authentication can be used to verify data.
- Timeout retransmission and piggyback mechanism are adopted to ensure reliability.
- The Sliding Window Protocol is adopted to control flows. As documented in the Protocol, unidentified groups in a window should be retransmitted.

Protocols and Standards

- RFC 793: Transmission Control Protocol
- RFC 1122: Requirements for Internet Hosts -- Communication Layers
- RFC 1191: Path MTU Discovery
- RFC 1213: Management Information Base for Network Management of TCP/IP-based Internets: MIB-II
- RFC 4022: Management Information Base for the Transmission Control Protocol (TCP)

12.2 Applications

Application	Description
Optimizing TCP Performance	To avoid TCP packet fragmentation on a link with a small MTU, Path MTU Discovery (PMTUD) is enabled.
Detecting TCP Connection Exception	TCP checks whether the peer works normally.

12.2.1 Optimizing TCP Performance

Scenario

For example, TCP connection is established between A and D, as shown in the following figure. The MTU of the link between A and B is 1500 bytes, 1300 bytes between B and C, and 1500 bytes between C and D. To optimize TCP transmission performance, packet fragmentation should be avoided between B and C.

Figure 11- 1

Remarks:	A, B, C and D are routers.

Deployment

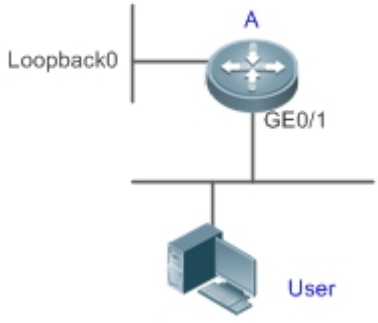
- Enable PMTUD on A and D.

12.2.2 Detecting TCP Connection Exception

Scenario

For example, in the following figure, User logs in to A through telnet but is shut down abnormally, as shown in the following figure. In case of TCP retransmission timeout, the User's TCP connection remains for a long period. Therefore, TCP keepalive can be used to rapidly detect TCP connection exception.

Figure 11- 2



Remarks:	A is a router.

Deployment

- Enable TCP keepalive on A.

12.3 Features

Basic Concepts

TCP Header Format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |       Destination Port        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
|                             Sequence Number                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Acknowledgment Number                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Data |           |U|A|P|R|S|F|                                       |
| Offset| Reserved  |R|C|S|S|Y|I|            Window                     |
|       |           |G|K|H|T|N|N|                                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Checksum            |         Urgent Pointer                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             data                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- **Source Port** is a 16-bit source port number.
- **Destination Port** is a 16-bit destination port number.
- **Sequence Number** is a 32-bit sequence number.
- **Acknowledgment Number** is a 32-bit number that identifies the next sequence number that the receiver is expecting to receive.
- **Data Offset** is a 4-bit number that indicates the total number of bytes in the TCP header (option included) divided by 4.
- A flag bit is 6-bit. URG: the urgent pointer field is significant; ACK: the acknowledgment field is significant; PSH: indicates the push function; RST: resets TCP connection; SYN: synchronizes the sequence number (establishing a TCP connection); FIN: no more data from the sender (closing a TCP connection).
- A 16-bit Window value is used to control flows. It specifies the amount of data that may be transmitted from the peer between ACK packets.
- **Checksum** is a 16-bit checksum.
- **Urgent Pointer** is 16-bit and shows the end of the urgent data so that interrupted data flows can continue. When the URG bit is set, the data is given priority over other data flows.

TCP Three-Way Handshake

- The process of TCP three-way handshake is as follows:

5. A client sends a SYN packet to the server.
6. The server receives the SYN packet and responds with a SYN ACK packet.
7. The client receives the SYN packet from the server and responds with an ACK packet.

- After the three-way handshake, the client and server are connected successfully and ready for data transmission.

Overview

Feature	Description
Configuring SYN Timeout	Configure a timeout waiting for a response packet after an SYN or SYN ACK packet is sent.
Configuring Window Size	Configure a window size.
Configuring Reset Packet Sending	Configure the sending of TCP reset packets after receiving port unreachable messages.
Configuring MSS	Configure an MSS for TCP connection.
Path MTU Discovery	Discover the smallest MTU on TCP transmission path, and adjust the size of TCP packets based on this MTU to avoid fragmentation.
TCP Keepalive	Check whether the peer works normally.

12.3.1 Configuring SYN Timeout

Working Principle

A TCP connection is established after three-way handshake: The sender sends an SYN packet, the receiver replies with a SYN ACK packet, and then the sender replies with an ACK packet.

- If the receiver does not reply with a SYN ACK packet after the sender sends an SYN packet, the sender keeps retransmitting the SYN packet for certain times or until timeout period expires.
- If the receiver replies with a SYN ACK packet after the sender sends an SYN packet but the sender does not reply with an ACK packet, the receiver keeps retransmitting the SYN ACK packet for certain times or until timeout period expires. (This occurs in the case of SYN flooding.)

Related Configuration

Configuring TCP SYN Timeout

- The default TCP SYN timeout is 20 seconds.
- Run the **ip tcp synwait-time** *seconds* command in global configuration mode to configure an SYN timeout ranging from 5 to 300 seconds.
- In case of SYN flooding, shortening SYN timeout reduces resource consumption. However, it does not work in continuous SYN flooding. When a device actively makes a request for a connection with an external device, through telnet for example, shortening SYN timeout reduces user's wait time. You may prolong SYN timeout properly on a poor network.

The **ip tcp syntime-out** command in version 10.x is disused but compatible in version 11.0. If this command is executed, it will be converted to the **ip tcp synwait-time** command.

In version 10.x, the configuration applies to only IPv4 TCP. In version 11.0 or later, it applies to both IPv4 TCP and IPv6 TCP.

12.3.2 Configuring Window Size

Working Principle

Data from the peer is cached in the TCP receiving buffer and subsequently read by applications. The TCP window size indicates the size of free space of the receiving buffer. For wide-bandwidth bulk-data connection, enlarging the window size dramatically promotes TCP transmission performance.

Related Configuration

Configuring Window Size

- Run the **ip tcp window-size** *size* command in global configuration mode to configure a window size ranging from 128 to (65535<< 14) bytes. The default is 65535 bytes. If the window size is greater than 65535 bytes, window enlarging will be enabled automatically.
- The window size advertised to the peer is the smaller value between the configured window size and the free space of the receiving buffer.

In version 10.x, the configuration applies to only IPv4 TCP. In version 11.0 or later, it applies to both IPv4 TCP and IPv6 TCP.

12.3.3 Configuring Reset Packet Sending

Working Principle

When TCP packets are distributed to applications, if the TCP connection a packet belongs to cannot be identified, the local end sends a reset packet to the peer to terminate the TCP connection. Attackers may use port unreachable messages to attack the device.

Related Configuration

Configuring the Sending of TCP Reset Packets After Receiving Port Unreachable Messages

By default, TCP reset packet sending upon receiving port unreachable messages is enabled.

Run the **no ip tcp send-reset** command in global configuration mode to disable TCP reset packet sending upon receiving port unreachable messages.

After this function is enabled, attackers may use port unreachable messages to attack the device.

The **ip tcp not-send-rst** command in version 10.x is disused but compatible in version 11.0. If this command is executed, it will be converted to the **no ip tcp send-reset** command.

In version 10.x, the configuration applies to only IPv4 TCP. In version 11.0 or later, it applies to both IPv4 TCP and IPv6 TCP.

12.3.4 Configuring MSS

Working Principle

The MSS refers to the total amount of data contained in a TCP segment t excluding TCP options.

Three-way handshake is implemented through MSS negotiation. Both parties add the MSS option to SYN packets, indicating the largest amount of data that the local end can handle, namely, the amount of data allowed from the peer. Both parties take the smaller MSS between them as the advertised MSS.

The MSS value is calculated as follows:

- IPv4 TCP: MSS = Outgoing interface MTU –IP header size (20-byte)–TCP header size (20-byte).
- IPv6 TCP: MSS = IPv6 Path MTU –IPv6 header size (40-byte)–TCP header size (20-byte).

In version 10.x, the configuration applies to only IPv4 TCP. In version 11.0 or later, it applies to both IPv4 TCP and IPv6 TCP.

The effective MSS is the smaller one between the calculated MSS and the configured MSS.

If a connection supports certain options, the option length (with **data offset** taken into consideration) should be deducted from an MSS value. For example, 20 bytes for MD5 digest (with **data offset** taken into consideration) should be subtracted from the MSS.

Related Configuration

Configuring MSS

- Run the **ip tcp mss** *max-segment-size* command in global configuration mode to set an MSS. It ranges from 68 to 1000 bytes. By default, the MSS is calculated based on MTU. If an MSS is configured, the effective MSS is the smaller one between the calculated MSS and the configured MSS.
- An excessively small MSS reduces transmission performance. You can promote TCP transmission by increasing the MSS. Choose an MSS value by referring to the interface MTU. If the former is bigger, TCP packets will be fragmented and transmission performance will be reduced.

12.3.5 Path MTU Discovery

Working Principle

The Path MTU Discovery f stipulated in RFC1191 is used to discover the smallest MTU in a TCP path to avoid fragmentation, enhancing network bandwidth utilization. The process of TCPv4 Path MTU Discovery is described as follows:

1. The source sends TCP packets with the Don't Fragment (DF) bit set in the outer IP header.
2. If the outgoing interface MTU value of a router in the TCP path is smaller than the IP packet length, the packet will be discarded and an ICMP error packet carrying this MTU will be sent to the source.
3. Through parsing the ICMP error packet, the source knows the smallest MTU in the path (path MTU) is.
4. The size of subsequent data segments sent by the source will not surpass the MSS, which is calculated as follows: TCP MSS = Path MTU – IP header size – TCP header size.

Related Configuration

Enabling Path MTU Discovery

By default, Path MTU Discovery is disabled.

Run the **ip tcp path-mtu-discovery** command to enable PMTUD in global configuration mode.

In version 10.x, the configuration applies to both IPv4 TCP and IPv6 TCP. In version 11.0 or later, it applies to only IPv4 TCP. TCPv6 PMTUD is enabled permanently and cannot be disabled.

12.3.6 TCP Keepalive

Working Principle

You may enable TCP keepalive to check whether the peer works normally. If a TCP end does not send packets to the other end for a period of time (namely idle period), the latter starts sending keepalive packets successively to the former for several times. If no response packet is received, the TCP connection is considered inactive and then closed.

Related Configuration

Enabling Keepalive

- By default, TCP keepalive is disabled.
- Run the **ip tcp keepalive** [**interval** *num1*] [**times** *num2*] [**idle-period** *num3*] command to in global configuration mode to enable

TCP keepalive. See **Configuration** for parameter description.

In version 10.x, the configuration applies to only IPv4 TCP. In version 11.0 or later, it applies to both IPv4 TCP and IPv6 TCP.

The **service tcp-keepalives-in** command is used in version 10.x to enable keeplive on the TCP server. It is disused but compatible in version 11.0. If this command is executed, it will be converted to the **ip tcp keepalive** [**interval** *num1*] [**times** *num2*] [**idle-period** *num3*] command.

The **service tcp-keepalives-out** command is used in version 10.x to enable keeplive on the TCP client. It is disused but compatible in version 11.0. If this command is executed, it will be converted to the **ip tcp keepalive** [**interval** *num1*] [**times** *num2*] [**idle-period** *num3*] command.

This command applies to both TCP server and client.

12.4 Configuration

Configuration	Description and Command	
Optimizing TCP Performance	(Optional) It is used to optimize TCP connection performance.	
	ip tcp synwait-time	Configures a timeout for TCP connection.
	ip tcp window-size	Configures a TCP window size.
	ip tcp send-reset	Configures the sending of TCP reset packets after receiving port unreachable messages.
	ip tcp mss	Configures an MSS for TCP connection.
	ip tcp path-mtu-discovery	Enables Path MTU Discovery.
Detecting TCP Connection Exception	(Optional) It is used to detect whether the peer works normally.	
	ip tcp keepalive	Enables TCP keepalive.

12.4.1 Optimizing TCP Performance

Configuration Effect

- Ensure optimal TCP performance and prevent fragmentation.

Notes

N/A

Configuration Steps

Configuring SYN Timeout

- Optional.
- Configure this on the both ends of TCP connection.

Configuring TCP Window Size

- Optional.
- Configure this on the both ends of TCP connection.

Configuring the Sending of TCP Reset Packets After Receiving Port Unreachable Messages.

- Optional.
- Configure this on the both ends of TCP connection.

Configuring MSS

- Optional.
- Configure this on the both ends of TCP connection.

Enabling Path MTU Discovery

- Optional.
- Configure this on the both ends of TCP connection.

Verification

N/A

Related Commands

Configuring SYN Timeout

Command	**ip tcp synwait-time** *seconds*
Parameter Description	*seconds*: Indicates SYN packet timeout. It ranges from 5 to 300 seconds. The default is 20 seconds.
Command Mode	Global configuration mode
Usage Guide	In case of SYN flooding, shortening SYN timeout reduces resource consumption. However, it does not work in continuous SYN flooding. When a device actively makes a request for a connection with an external device, through telnet for example, shortening SYN timeout reduces user's wait time. You may prolong SYN timeout properly on a poor network.

Configuring TCP Window Size

Command	**ip tcp window-size** *size*
Parameter Description	*size*: Indicates a TCP window size. It ranges from 128 to (65535 << 14) bytes. The default is 65535 bytes.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Sending of TCP Reset Packets After Receiving Port Unreachable Messages

Command	**ip tcp send-reset**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	By default, TCP reset packet sending upon receiving port unreachable messages is enabled.

Configuring MSS

Command	**ip tcp mss** *max-segment-size*
Parameter Description	*max-segment-size*: Indicates the maximum segment size. It ranges from 68 to 10000 bytes. By default, the MSS is calculated based on MTU.
Command Mode	Global configuration mode
Usage Guide	This command defines the MSS for a TCP communication to be established. The negotiated MSS for a new connection should be smaller than this MSS. If you want to reduce the MSS, run this command. Otherwise, do not perform the configuration.

Configuring Path MTU Discovery

Command	**ip tcp path-mtu-discovery** [**age-timer** *minutes* **| age-timer infinite**]
Parameter Description	**age-timer** *minutes*: Indicates the interval for a new probe after a path MTU is discovered. It ranges from 10 to 30 minutes. The default is 10 minutes. **age-timer infinite**: No probe is implemented after a path MTU is discovered.
Command Mode	Global configuration mode
Usage Guide	The PMTUD is an algorithm documented in RFC1191 aimed to improve bandwidth utilization. When the TCP is applied to bulk data transmission, this function may facilitate transmission performance. If the MSS used for the connection is smaller than what the peer connection can handle, a larger MSS is tried every time the age timer expires. The age timer is a time interval for how often TCP estimates the path MTU with a larger MSS. The discovery process is stopped when either the send MSS is as large as the peer negotiated, or the user has disabled the timer on the router. You may turn off the timer by setting it to **infinite**.

Configuration Example

Enabling Path MTU Discovery

Configuration Steps	Enable PMTUD for a TCP connection. Adopt the default age timer settings.
	FS# configure terminal FS(config)# ip tcp path-mtu-discovery FS(config)# end
Verification	Run the **show tcp pmtu** command to display the IPv4 TCP PMTU.
	FS# show tcp pmtu Number Local Address Foreign Address PMTU 1 192.168.195.212.23 192.168.195.112.13560 1440
	Run the **show ipv6 tcp pmtu** command to display the IPv6 TCP PMTU.
	FS# show ipv6 tcp pmtu Number Local Address Foreign Address PMTU 1 1000::1:23 1000::2.13560 1440

Common Errors

N/A

12.4.2 Detecting TCP Connection Exception

Configuration Effect

- Check whether the peer works normally.

Notes

N/A

Configuration Steps

Enabling TCP Keepalive

- Optional.

Verification

N/A

Related Commands

Enabling TCP Keepalive

Command	**ip tcp keepalive** [**interval** *num1*] [**times** *num2*] [**idle-period** *num3*]
Parameter Description	**interval** *num1*: Indicates the interval to send keepalive packets. Ranging from 1 to120 seconds. The default is 75 seconds. **times** *num2*: Indicates the maximum times for sending keepalive packets. It ranges from 1 to 10. The default is 6. **idle-period** *num3*: Indicates the time when the peer sends no packets to the local end, It ranges from 60 to 1800 seconds. The default is15 minutes.
Command Mode	Global configuration mode
Usage Guide	You may enable TCP keepalive to check whether the peer works normally. The function is disabled by default. Suppose a user enables TCP keepalive function with the default interval, times and idle period settings. The user does not receive packets from the other end within 15 minutes and then starts sending Keepalive packets every 75 seconds for 6 times. If the user receives no TCP packets, the TCP connection is considered inactive and then closed.

Configuration Example

Enabling TCP Keepalive

Configuration Steps	Enable TCP keepalive on a device with **interval** and **idle-period** set to 3 minutes and 60 seconds respectively. If the user receives no TCP packets from the other end after sending keepalive packets four times, the TCP connection is considered inactive.

	FS# configure terminal FS(config)# ip tcp keepalive interval 60 times 4 idle-period 180 FS(config)# end
Verification	A user logs in to a device through telnet, and then shuts down the local device. Run the **show tcp connect** command on the remote device to observe when IPv4 TCP connection is deleted.

Common Errors

N/A

12.5 Monitoring

Displaying

Description	Command
Displays basic information on IPv4 TCP connection.	**show tcp connect** [**local-ip** *a.b.c.d*] [**local-port** *num*] [**peer-ip** *a.b.c.d*] [**peer-port** *num*]
Displays IPv4 TCP connection statistics.	**show tcp connect statistics**
Displays IPv4 TCP PMTU.	**show tcp pmtu** [**local-ip** *a.b.c.d*] [**local-port** *num*] [**peer-ip** *a.b.c.d*] [**peer-port** *num*]
Displays IPv4 TCP port information.	**show tcp port** [*num*]
Displays IPv4 TCP parameters.	**show tcp parameter**
Displays IPv4 TCP statistics.	**show tcp statistics**
Displays basic information on IPv6 TCP connection.	**show ipv6 tcp connect** [**local-ipv6** *X:X:X:X::X*] [**local-port** *num*] [**peer-ipv6** *X:X:X:X::X*] [**peer-port** *num*]
Displays IPv6 TCP connection statistics.	**show ipv6 tcp connect statistics**
Displays IPv6 TCP PMTU.	**show ipv6 tcp pmtu** [**local-ipv6** *X:X:X:X::X*] [**local-port** *num*] [**peer-ipv6** *X:X:X:X::X*] [**peer-port** *num*]
Displays IPv6 TCP port information.	**show ipv6 tcp port** [*num*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Displays the debugging information on IPv4 TCP packets.	**debug ip tcp packet** [**in** | **out**] [**local-ip** *a.b.c.d*] [**peer-ip** *a.b.c.d*] [**global**] [**local-port** *num*] [**peer-port** *num*] [**deeply**]
Displays the debugging information on IPv4 TCP connection.	**debug ip tcp transactions** [**local-ip** *a.b.c.d*] [**peer-ip** *a.b.c.d*] [**local-port** *num*] [**peer-port** *num*]
Displays the debugging information on IPv6 TCP packets.	**debug ipv6 tcp packet** [**in** | **out**] [**local-ipv6** *X:X:X:X::X*] [**peer-ipv6** *X:X:X:X::X*] [**global**] [**local-port** *num*] [**peer-port** *num*] [**deeply**]
Displays the debugging information on IPv6 TCP connection.	**debug ipv6 tcp transactions** [**local-ipv6** *X:X:X:X::X*] [**peer-ipv6** *X:X:X:X::X*] [**local-port** *num*] [**peer-port** *num*]

13 Configuring IPv4/IPv6 REF

13.1 Overview

On products incapable of hardware-based forwarding, IPv4/IPv6 packets are forwarded through the software. To optimize the software-based forwarding performance, FS introduces IPv4/IPv6 express forwarding through software (FS Express Forwarding, namely REF).

REF maintains two tables: forwarding table and adjacency table. The forwarding table is used to store route information. The adjacency table is derived from the ARP table and IPv6 neighbor table, and it contains Layer 2 rewrite(MAC) information for the next hop..

REF is used to actively resolve next hops and implement load balancing.

Protocols and Standards

N/A

13.2 Applications

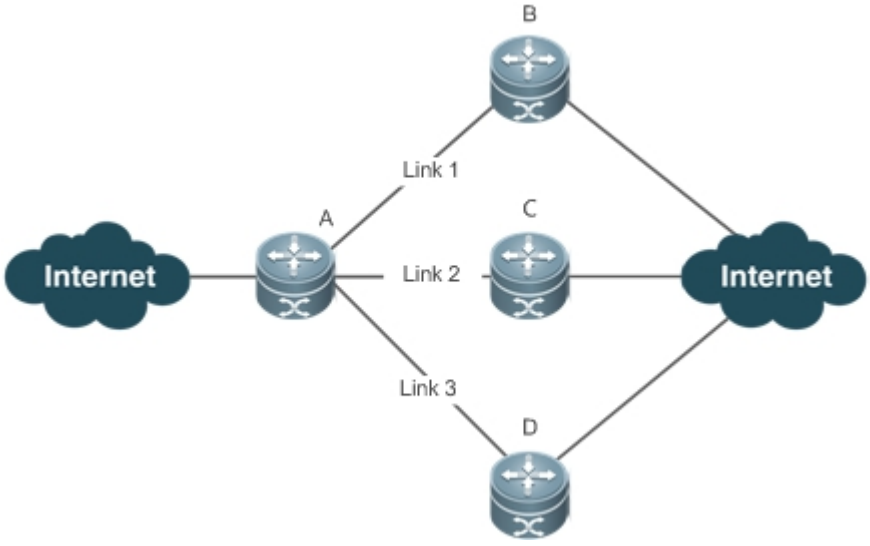
Application	Description
Load Balancing	During network routing, when a route prefix is associated with multiple next hops, REF can implement load balancing among the multiple next hops.

13.3 Load Balancing

Scenario

As shown in Figure 12- 1, a route prefix is associated with three next hops on router A, namely, link 1, link 2, and link 3. By default, REF implements load balancing based on the destination IP address. Load balancing can be implemented based on the source IP address and destination IP address as well.

Figure 12- 1



Remarks	A is a router that runs REF. B, C and D are forwarding devices.

Deployment

- Run REF on router A.

13.4 Features

Basic Concepts

IPv4/IPv6 REF involves the following basic concepts:

Routing table

An IPv4/IPv6 routing table stores routes to the specific destinations and contains the topology information. During packet forwarding, IPv4/IPv6 REF selects packet transmission paths based on the routing table.

Adjacent node

An adjacent node contains output interface information about routed packets, for example, the next hop, the next component to be processed, and the link layer encapsulation. When a packet is matched with an adjacent node, the packet is directly encapsulated and then forwarded. For the sake of query and update, an adjacent node table is often organized into a hash table. To support routing load balancing, the next hop information is organized into a load balance entry. An adjacent node may not contain next hop information. It may contain indexes of next components (such as other line cards and multi-service cards) to be processed.

Active resolution

REF supports next hop resolution. If the MAC address of the next hop is unknown, REF will actively resolve the next hop. IPv4 REF requests the ARP module for next hop resolution while IPv6 REF applies the ND module to resolution.

Packet forwarding path

Packets are forwarded based on their IPv4/IPv6 addresses. If the source and destination IPv4/IPv6 addresses of a packet are specified, the forwarding path of this packet is determined.

13.4.1 Load Balancing Policies

Load balancing is configured to distribute traffic load among multiple network links.

Working Principle

REF supports two load balancing modes. In the REF model, a route prefix is associated with multiple next hops, in other words, it is a multi-path route. The route will be associated with a load balance table and implement weight-based load balancing. When an IPv4/IPv6 packet is matched with a load balance entry based on the longest prefix match, REF performs hash calculation based on the IPv4/IPv6 address of the packet and selects a path to forward the packet.

IPv4/IPv6 REF supports two kinds of load balancing policies: load balancing based on destination IP address, and load balancing based on the source and destination IP addresses.

Related Configuration

13.5 Monitoring

Displaying REF Packet Statistics

REF packet statistics includes the number of forwarded packets and the number of packets discarded due to various causes. You can determine whether packets are forwarded as expected by displaying and clearing REF packet statistics.

Command	Description
show ip ref packet statistics	Displays IPv4 REF packet statistics.
clear ip ref packet statistics	Clears IPv4 REF packet statistics.
show ipv6 ref packet statistics	Displays IPv6 REF packet statistics.
clear ipv6 ref packet statistics	Clears IPv6 REF packet statistics.

Displaying Adjacency Information

You can run the following commands to display adjacency information:

Command	Description
show ip ref adjacency [**glean** | **local** | *ip-address* | {**interface** *interface_type interface_number* } **| discard | statistics**]	Displays the gleaned adjacencies, local adjacencies, adjacencies of a specified IP address, adjacencies associated with a specified interface, and all adjacent nodes in IPv4 REF.
show ipv6 ref adjacency [**glean | local |** *ipv6-address* **| (interface** *interface_type interface_number*} **| discard | statistics**]	Displays the gleaned adjacencies, local adjacencies, adjacencies of a specified IPv6 address, adjacencies associated with a specified interface, and all adjacent nodes in IPv6 REF.

Displaying Active Resolution Information

You can run the following commands to display next hops to be resolved:

Command	Description
show ip ref resolve-list	Displays the next hop to be resolved .
show ipv6 ref resolve-list	Displays the next hop to be resolved.

Displaying Packet Forwarding Path Information

Packets are forwarded based on their IPv4/IPv6 addresses. If the source and destination IPv4/IPv6 addresses of a packet are specified, the forwarding path of this packet is determined. Run the following commands and specify the IPv4/IPv6 source and destination addresses of a packet. The forwarding path of the packet is displayed, for example, the packet is discarded, submitted to a CPU, or forwarded. Furthermore, the interface that forwards the packet is displayed.

Command	Description
show ip ref exact-route *source-ipaddress dest_ipaddress*	Displays the forwarding path of a packet.
show ipv6 ref exact-route *src-ipv6-address dst-ipv6-address*	Displays the forwarding path of an IPv6 packet.

Displaying Route Information in an REF Table

Run the following commands to display the route information in an REF table:

Command	Description
show ip ref route [**default** | {*ip mask*}| **statistics**]	Displays route information in the IPv4 REF table. The parameter **default** indicates a default route.
show ipv6 ref route [**default** | **statistics** | *prefix/len*]	Displays route information in the IPv6 REF table. The parameter **default** indicates a default route.

Chapter4 IP Routing Configuration

1. Configuring RIP
2. Configuring OSPFv2
3. Configuring OSPFv3
4. Configuring IS-IS
5. Configuring BGP
6. Configuring PBR
7. Configuring VRF
8. Configuring RIPng
9. Managing Routes
10. Configuring Keys
11. Configuring Routing Policies

1 Configuring RIP

1.1 Overview

Routing Information Protocol (RIP) is a unicast routing protocol applied on IPv4 networks. RIP-enabled routers exchange routing information to obtain routes to remote networks.

As an Interior Gateway Protocol (IGP), RIP can run only within the autonomous system (AS) and is applicable to small-sized networks whose longest path involves less than 16 hops.

Protocols and Standards

- RFC1058: Defines RIPv1.
- RFC2453: Defines RIPv2.

1.2 Applications

Application	Description
Basic RIP Application	The routing information is automatically maintained through RIP on a small-sized network.

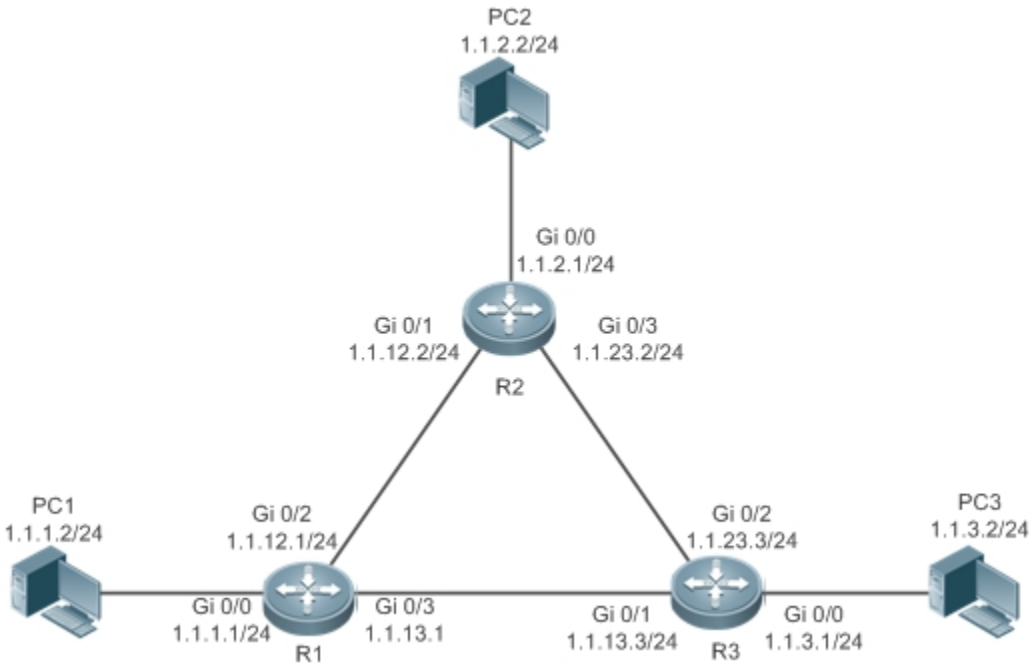
1.2.1 Basic RIP Application

Scenario

On a network with a simple structure, you can configure RIP to implement network interworking. Configuring RIP is simpler than configuring other IGP protocols like Open Shortest Path First (OSPF). Compared with static routes, RIP can dynamically adapt to the network structure changes and is easier to maintain.

As shown in Figure 1- 1, to implement interworking between PC1, PC2, and PC3, you can configure RIP routes on R1, R2, and R3.

Figure 1- 1



Deployment

- Configure IP addresses and gateways on three PCs.
- Configure IP addresses and subnet masks on three routers.
- Configure RIP on three routers.

1.3 Features

Basic Concepts

IGP and EGP

IGP runs within an AS. For example, RIP is a type of IGP.

Exterior Gateway Protocol (EGP) runs between ASs.

Classful Routing Protocol and Classless Routing Protocol

Protocols can be classified based on the type of routes supported:

- Classful routing protocol: It supports classful routes. For example, RIPv1 is a classful routing protocol.
- Classless routing protocol: It supports classless routes. For example, RIPv2 is a classless routing protocol.

Overview

Feature	Description
RIPv1 and RIPv2	RIP is available in two versions: RIPv1 and RIPv2.
Exchanging Routing Information	By exchanging routing information, RIP-enabled devices can automatically obtain routes to a remote network and update the routes in real time.
Routing Algorithm	RIP is a protocol based on the distance-vector algorithm. It uses the vector addition method to compute the routing information.
Avoiding Route Loops	RIP uses functions, such as split horizon and poison reverse, to avoid route loops.
Security Measures	RIP uses functions, such as authentication and source address verification, to ensure protocol security.
Reliability Measures	RIP uses functions, such as fast reroute, and graceful restart (GR), to enhance reliability of the protocol.

1.3.1 Routing Algorithm

RIP is a protocol based on the distance-vector algorithm. It uses the vector addition method to compute the routing information.

Working Principle

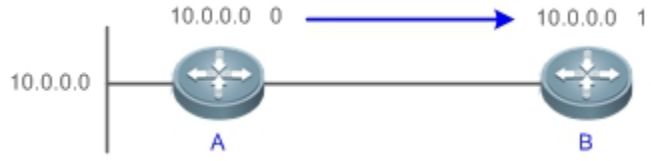
Distance-Vector Algorithm

RIP is a protocol based on the distance-vector algorithm. The distance-vector algorithm treats a route as a vector that consists of the destination network and distance (metric). The router obtains a route from its neighbor and adds the distance vector from itself to the neighbor to the route to form its own route.

RIP uses the hop count to evaluate the distance (metric) to the destination network. By default, the hop count from a router to its directly connected network is 0, the hop count from a router to a network that can be reached through the router is 1, and so on. That is, the metric is equal to the number of routers from the local network to the destination network. To restrict the convergence time, RIP stipulates that the metric must be an integer between 0 and 15. If the metric is equal to or greater than 16, the destination network or host is unreachable. For this reason, RIP cannot be applied on a large-scale network.

As shown in Figure 1- 2, Router A is connected to the network 10.0.0.0. Router B obtains the route (10.0.0.0,0) from Router A and adds the metric 1 to the route to obtain its own route ((10.0.0.0,1), and the next hop points to Router A.

Figure 1- 2

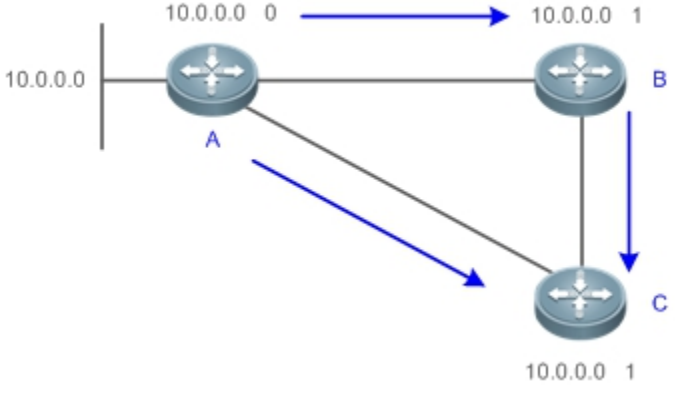


Selecting the Optimum Route

RIP selects an optimum route based on the following principle: If multiple routes to the same destination network is available, a router preferentially selects the route with the smallest metric.

As shown in Figure 1- 3, Router A is connected to the network 10.0.0.0. Router C obtains the route (10.0.0.0,0) from Router A and the route (10.0.0.0,1) from Router B. Router C will select the route that is obtained from Router A and add metric 1 to this route to form its own route (10.0.0.0,1), and the next hop points to Router A.

Figure 1- 3



When routes coming from different sources exist on a router, the route with the smallest distance is preferentially selected.

Route Source	Default Distance
Directly-connected network	0
Static route	1
OSPF route	110
RIP route	120
Unreachable route	255

Related Configuration

Modifying the Distance

By default, the distance of a RIP route is 120.

Run the **distance** command to modify the distance of a RIP route.

Modifying the Metric

For a RIP route that is proactively discovered by a device, the default metric is equal to the number of hops from the local network to the destination network. For a RIP router that is manually configured (default route or redistributed route), the default metric is 1.

Run the **offset-list in** command to increase the metric of a received RIP route.

Run the **offset-list out** command to increase the metric of a sent RIP route.

Run the **default-metric** command to modify the default metric of a redistributed route.

Run the **redistribute** command to modify the metric of a route when the route is redistributed.

Run the **default-information originate** command to modify the metric of a default route when the default route is introduced.

Run the **ip rip default-information** command to modify the metric of a default route when the default route is created.

1.3.2 RIPv1 and RIPv2

Two RIP versions are available: RIPv1 and RIPv2.

Working Principle

RIPv1

RIPv1 packets are broadcast. The broadcast address is 255.255.255.255, and the UDP port ID is 520. RIPv1 cannot identify the subnet mask, and supports only classful routes.

RIPv2

RIPv2 packets are multicast. The multicast address is 224.0.0.9, and the UDP port ID is 520. RIPv2 can identify the subnet mask, and supports classless routes, summarized route, and supernetting routes. RIPv2 supports plain text authentication and message digest 5 (MD5) authentication.

Related Configuration

Enabling the RIP Process

The RIP process is disabled by default.

Run the **router rip** command to enable the RIP process.

You must enable the RIP process on a device; otherwise, all functions related to RIP cannot take effect.

Running RIP on an Interface

By default, RIP does not run on an interface.

Run the **network** command to define an address range. RIP runs on interfaces that belong to this address range.

After RIP runs on an interface, RIP packets can be exchanged on the interface and RIP can learn routes to the network segments directly connected to the device.

Defining the RIP Version

By default, an interface receives RIPv1 and RIPv2 packets, and sends RIPv1 packets.

Run the **version** command to define the version of RIP packets sent or received on all interfaces.

Run the **ip rip send version** command to define the version of RIP packets sent on an interface.

Run the **ip rip receive version** command to define the version of RIP packets received on an interface.

If the versions of RIP running on adjacent routers are different, the RIPv1-enabled router will learn incorrect routes.

Preventing an Interface from Sending or Receiving Packets

By default, a RIP-enabled interface is allowed to send and receive RIP packets.

Run the **no ip rip receive enable** command to prevent an interface from receiving RIP packets.

Run the **no ip rip send enable** command to prevent an interface from sending RIP packets.

Run the **passive-interface** command to prevent an interface from sending broadcast or multicast RIP packets.

Configuring the Mode for Sending RIP Packets

By default, broadcast RIPv1 packets and multicast RIPv2 are sent.

Run the **ip rip v2-broadcast** command to send broadcast RIPv2 packets on an interface.

Run the **neighbor** command to send unicast RIP packets to a specified neighbor router.

1.3.3 Exchanging Routing Information

Compared with static routing, the dynamic routing protocol has a significant advantage, that is, by exchanging routing information, devices can automatically obtain routes to a remote network and update the routes in real time.

Working Principle

Initialization

After RIP is enabled on a router, the router sends a request packet to its neighbor router, requesting for all routing information, that is, the routing table. After receiving the request message, the neighbor router returns a response packet containing the local routing table. After receiving the response packet, the router updates the local routing table, and sends an update packet to the neighbor router, informing the neighbor router of the route update information. After receiving the update packet, the neighbor router updates the local routing table, and sends the update packet to other adjacent routers. After a series of updates, all routers can obtain and retain the latest routing information.

Periodical Update

By default, periodical update is enabled for RIP. Adjacent routers exchange complete routing information with each other every 30s (update timer), that is, the entire routing table is sent to neighbor routers. One update packet contains at most 25 routes. Therefore, a lot of update packets may be required to send the entire routing table. You can set the sending delay between update packets to avoid loss of routing information.

For every non-local route, if the route is not updated within 180s (invalid timer), the metric of the route is changed to 16 (unreachable). If the route is still not updated in the next 120s (flush timer), the route is deleted from the routing table.

Triggered Updates

After the triggered updates function is enabled, periodical update is automatically disabled. When routing information changes on a router, the router immediately sends routes related to the change (instead of the complete routing table) to the neighbor router, and use the acknowledgment and retransmission mechanisms to ensure that the neighbor router receives the routes successfully. Compared with periodical update, triggered updates help reduce flooding and accelerates route convergence.

Events that can trigger update include router startup, interface status change, changes in routing information (such as the metric), and reception of a request packet.

Route Summarization

When sending routing information to a neighbor router, the RIP-enabled router summarizes subnet routes that belong to the same classful network into a route, and sends the route to the neighbor router. For example, summarize 80.1.1.0/24 (metric=2) and 80.1.2.0/24 (metric=3) into 80.0.0.0/8 (metric=2), and set the metric of the summarized route to the optimum metric.

Only RIPv2 supports route summarization. Route summarization can reduce the size of the routing table and improve the efficiency of routing information exchange.

Supernetting Route

If the subnet mask length of a route is smaller than the natural mask length, this route is called supernetting route. For example, in the 80.0.0.0/6 route, as 80.0.0.0 is a Class A network address and the natural mask is 8 bits, 80.0.0.0/6 route is a supernetting route.

Only RIPv2 supports supernetting routes.

Default Route

In the routing table, a route to the destination network 0.0.0.0/0 is called default route.

The default route can be learned from a neighbor router, or sent to a neighbor router.

Route Redistribution

For RIP, other types of routes (such as direct routes, static routes, and routes of other routing protocols) are called external routes.

External routes (excluding the default route) can be redistributed to RIP and advertised to neighbors.

Route Filtering

Filtering conditions can be configured to limit the routing information exchanged between adjacent routers. Only the routing information that meets filtering conditions can be sent or received.

Related Configuration

Sending Delay Between Update Packets

By default, the update packets are sent continuously without any delay.
Run the **output-delay** command to set the sending delay between update packets.

RIP Timers

By default, the update timer is 30s, the invalid timer is 180s, and the flush timer is 120s.

Run the **timers basic** command to modify durations of the RIP timers.

Increasing the duration of the flush timer can reduce the route flapping. Decreasing the duration of the flush timer helps accelerate route convergence.

The durations of RIP timers must be consistent on adjacent routers. Unless otherwise required, you are advised not to modify the RIP timers.

Triggered Updates

By default, periodical update is enabled.

Run the **ip rip triggered** command to enable triggered updates on the interface and disable periodical update.

Run the **ip rip triggered retransmit-timer** command to modify the retransmission interval of update packets. The default value is 5s.

Run the **ip rip triggered retransmit-count** command to modify the maximum retransmission times of update packets. The default value is 36.

Route Summarization

By default, route summarization is automatically enabled if an interface is allowed to send RIPv2 packets.

Run the **no auto-summary** command to disable route summarization.

Run the **ip rip summary-address** command to configure route summarization on an interface.

Supernetting Route

By default, supernetting routes can be sent if an interface is allowed to send RIPv2 packets.

Run the **no ip rip send supernet-routes** command to prevent the sending of supernetting routes.

Default Route

Run the **ip rip default-information** command to advertise the default route to neighbors on an interface.

Run the **default-information originate** command to advertise the default route to neighbors from all interfaces.

Route Redistribution

Run the **redistribute** command to redistribute external routes (excluding the default route) to RIP and advertise them to neighbors.

Route Filtering

Run the **distribute-list out** command to set filtering rules to limit the routing information sent by the device.

Run the **distribute-list in** command to set filtering rules to limit the routing information received by the device.

1.3.4 Avoiding Route Loops

RIP uses functions, such as split horizon and poison reverse, to avoid route loops.

Working Principle

Route Loop

A RIP route loop occurs due to inherent defects of the distance-vector algorithm.

As shown in Figure 1- 4, Router A is connected to the network 10.0.0.0, and sends an update packet every 30s. Router B receives the route 10.0.0.0 from Router A every 30s. If Router A is disconnected from 10.0.0.0, the route to 10.0.0.0 will be deleted from the routing table on Router A. Next time, the update packet sent by Router A no longer contains this route. As Router B does not receive an update packet related to 10.0.0.0, Router B determines that the route to 10.0.0.0 is valid within 180s and uses the Update packet to send this route to Router A. As the route to 10.0.0.0 does not exist on Router A, the route learned from Router B is added to the routing table. Router B determines that data can reach 10.0.0.0 through Router A, and Router A determines that data can reach 10.0.0.0 through Router B. In this way, a route loop is formed.

Figure 1- 4

Split Horizon

Split horizon can prevent route loops. After split horizon is enabled on an interface, a route received on this interface will not be sent out from this interface.

As shown in Figure 1- 5, after split horizon is enabled on the interface between Router A and Router B, Router B will not send the route 10.0.0.0 back to Router A. Router B will learn 180s later that 10.0.0.0 is not reachable.

Figure 1- 5

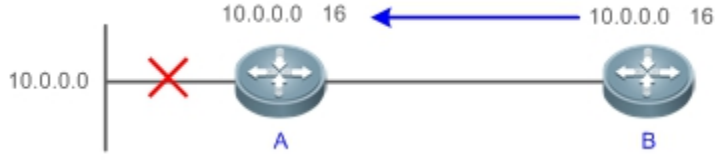
Poison Reverse

Poison reverse can also prevent route loops. Compared with slit horizon, poison reverse is more reliable, but brings more protocol packets, which makes network congestion more severe.

After poison reverse is enabled on an interface, a route received from this interface will be sent out from this interface again, but the metric of this router will be changed to 16 (unreachable).

As shown in Figure 1- 6, after learning the route 10.0.0.0 from Router A, Router B sets the metric of this route to 16 and sends the route back to Router A. After this route becomes invalid, Router B advertises the route 10.0.0.0 (metric = 16) to Router A to accelerate the process of deleting the route from the routing table.

Figure 1- 6



Related Configuration

Split Horizon

By default, split horizon is enabled.

Run the **no ip rip split-horizon** command to disable split horizon.

Poison Reverse

By default, poison reverse is disabled.

Run the **ip rip split-horizon poisoned-reverse** command to enable poison reverse. (After poison reverse is enabled, split horizon is automatically disabled.)

1.3.5 Security Measures

RIP uses functions, such as authentication and source address verification, to ensure protocol security.

Working Principle

Authentication

RIPv2 supports authentication, but RIPv1 does not.

After authentication is enabled on an interface, the routing information cannot be exchanged between adjacent devices if authentication fails. The authentication function is used to prevent unauthorized devices from accessing the RIP routing domain.

RIPv2 supports plain text authentication and MD5 authentication.

Source Address Verification

When a RIP-enabled device receives an Update packet, it checks whether the source IP address in the packet and the IP address of the inbound interface are in the same network segment. If not, the device drops the packet. Source address verification is used to ensure that RIP routing information is exchanged only between adjacent routing devices.

- On an unnumbered IP interface, source address verification is not performed (not configurable).
- If the triggered updates function is enabled, source address verification is automatically enabled (not configurable).
- If split horizon is disabled, source address verification is automatically enabled (not configurable).

Related Configuration

Authentication

By default, authentication is disabled.

Run the **ip rip authentication mode text** command to enable plain text authentication on an interface.

Run the **ip rip authentication mode md5** command to enable MD5 authentication on an interface.

Run the **ip rip authentication text-password** command to set the password for plain text authentication on an interface.

Run the **ip rip authentication key-chain** command to reference the key in the configured key chain as the authentication key on an interface.

Source Address Verification

By default, source address verification is enabled.

Run the **no validate-update-source** command to disable source address verification.

1.3.6 Reliability Measures

RIP uses GR to enhance reliability of the protocol.

Working Principle

GR

GR ensures uninterrupted data transmission when the protocol is restarted. If RIP is restarted on a GR-enabled device, the forwarding table before restart will be retained and a request packet will be sent to the neighbor so that the route can be learned again. During the

GR period, RIP completes re-convergence of the route. After the GR period expires, RIP updates the forwarding entry and advertises the routing table to the neighbor.

Related Configuration

GR

By default, GR is disabled.

Run the **graceful-restart** command to enable the GR function.

1.4 Configuration

Configuration	Description and Command	
Configuring RIP Basic Functions	(Mandatory) It is used to build a RIP routing domain.	
	router rip	Enables a RIP routing process and enters routing process configuration mode.
	network	Runs RIP on interfaces in the specified address range.
	version	Defines the RIP version.
	ip rip split-horizon	Enables split horizon or poison reverse on an interface.
	passive-interface	Configures a passive interface.
Controlling Interaction of RIP Packets	(Optional) This configuration is required if you wish to change the default mechanism for sending or receiving RIP packets.	
	neighbor	Sends unicast RIP packets to a specified neighbor.
	ip rip v2-broadcast	Sends broadcast RIPv2 packets on an interface.
	ip rip receive enable	Allows the interface to receive RIP packets.
	ip rip send enable	Allows the interface to send RIP packets.
	ip rip send version	Defines the version of RIP packets sent on an interface.
	ip rip receive version	Defines the version of RIP packets received on an interface.
Enabling Triggered Updates	Optional.	
	ip rip triggered	Enables triggered updates on an interface.
Enabling Source Address Verification	Optional.	
	validate-update-source	Enables source address verification.
Enabling Authentication	(Optional) Only RIPv2 supports authentication.	
	ip rip authentication mode	Enables authentication and sets the authentication mode on an interface.
	ip rip authentication text-password	Configures the password for plain text authentication on an interface.
	ip rip authentication key-chain	Configures the authentication key chain on an interface.

Configuration	Description and Command	
Enabling Route Summarization	(Optional) Only RIPv2 supports route summarization.	
	auto-summary	Enables automatic summarization of RIP routes.
	ip rip summary-address	Configures route summarization on an interface.
Enabling Supernetting Routes	(Optional) Only RIPv2 supports supernetting routes.	
	ip rip send supernet-routes	Enables advertisement of RIP supernetting routes on an interface
Advertising the Default Route or External Routes	Optional.	
	ip rip default-information	Advertises the default route to neighbors on an interface.
	default-information originate	Advertises the default route to neighbors.
	redistribute	Redistributes routes and advertises external routes to neighbors.
Setting Route Filtering Rules	Optional.	
	distribute-list in	Filters the received RIP routing information.
	distribute-list out	Filters the sent RIP routing information.
Modifying Route Selection Parameters	Optional.	
	distance	Modifies the administrative distance (AD) of a RIP route.
	offset-list	Increases the metric of a received or sent RIP route.
	default-metric	Configures the default metric of an external route redistributed to RIP.
Modifying Timers	Optional.	
	timers basic	Modifies the update timer, invalid timer, and flush timer.
	output-delay	Sets the sending delay between RIP route update packets.
Enabling GR	Optional.	
	graceful-restart	Configures the GR restarter capability.

1.4.1 Configuring RIP Basic Functions

Configuration Effect

- Build a RIP routing domain on the network.
- Routers in the domain obtain routes to a remote network through RIP.

Notes

- IPv4 addresses must be configured.
- IPv4 unicast routes must be enabled.

Configuration Steps

Enabling a RIP Routing Process

- Mandatory.
- Unless otherwise required, this configuration must be performed on every router in the RIP routing domain.

Associating with the Local Network

- Mandatory.
- Unless otherwise required, this configuration must be performed on every router in the RIP routing domain.
- Unless otherwise required, the local network associated with RIP should cover network segments of all L3 interfaces.

Defining the RIP Version

- If RIPv2 functions (such as the variable length subnet mask and authentication) are required, enable the RIPv2.
- Unless otherwise required, you must define the same RIP version on every router.

Enabling Split Horizon or Poison Reverse

- By default, split horizon is enabled and poison reverse is disabled.
- Unless otherwise required, enable split horizon on every interface connected to the broadcast network, such as the Ethernet. (Retain the default setting.)
- Unless otherwise required, enable split horizon on every interface connected to the point-to-point (P2P) network, such as the PPP and HDLC. (Retain the default setting.)
- It is recommended that split horizon and poison reverse be disabled on an interface connected to a non-broadcast multi-access (NBMA) network, such as FR and X.25; otherwise, some devices may fail to learn the complete routing information.
- If the secondary IP address is configured for an interface connected to a non-broadcast, it is recommended that split horizon and poison reverse be disabled.

Configuring a Passive Interface

- If you want to suppress Update packets on a RIP interface, configure the interface as a passive interface.
- Use the passive interface to set the boundary of the RIP routing domain. The network segment of the passive interface belongs to the RIP routing domain, but RIP packets cannot sent over the passive interface.
- If RIP routes need to be exchanged on an interface (such as the router interconnect interface) in the RIP routing domain, this interface cannot be configured as a passive interface.

Verification

- Check the routing table on a router to verify that the route to a remote network can be obtained through RIP.

Related Commands

Enabling a RIP Routing Process

Command Syntax	**router rip**
Parameter Description	N/A

Command Mode	Global configuration mode
Configuration Usage	This command is used to create a RIP routing process and enter routing process configuration mode.

Associating with the Local Network

Command Syntax	**network** *network-number* [*wildcard*]
Parameter Description	*network-number*: Indicates the number of a network. *wildcard*: Defines the IP address comparison bit. **0** indicates accurate matching, and **1** indicates that no comparison is performed.
Command Mode	Routing process configuration mode
Configuration Usage	RIP can run and learn direct routes and RIP packets can be exchanged only on an interface covered by **network**. If **network** 0.0.0.0 255.255.255.255 is configured, all interfaces are covered. If *wildcard* is not configured, the classful address range is used by default, that is, the interfaces whose addresses fall into the classful address range participate in RIP operations.

Defining the RIP Version

Command Syntax	**version** { **1** | **2** }
Parameter Description	**1**: Indicates RIPv1. **2**: Indicates RIPv2.
Command Mode	Global configuration mode
Configuration Usage	This command takes effect on the entire router. You can run this command to define the version of RIP packets sent or received on all interfaces.

Enabling Split Horizon

Command Syntax	**ip rip split-horizon** [**poisoned-reverse**]
Parameter Description	**poisoned-reverse**: Indicates poison reverse.
Command Mode	Interface configuration mode
Configuration Usage	After poison reverse is enabled, split horizon is automatically disabled.

Configuring a Passive Interface

Command Syntax	**passive-interface** { **default** | *interface-type interface-num* }
Parameter Description	**default:** Indicates all interfaces. ***interface-type interface-num*:** Specifies an interface.

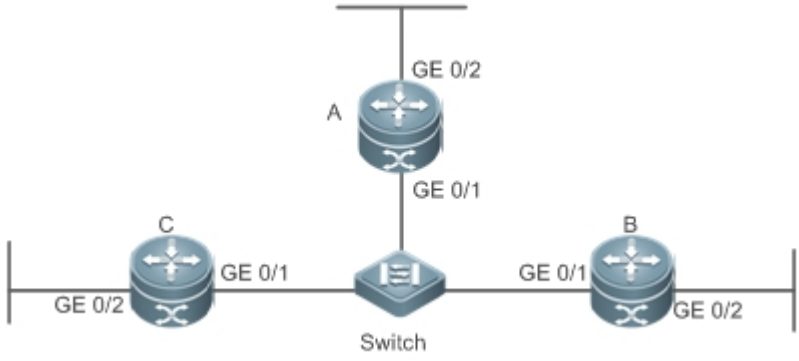
Command Mode	Routing process configuration mode
Configuration Usage	First, run the **passive-interface default** command to configure all interfaces as passive interfaces. Then, run the **no passive-interface** *interface-type interface-num* command to cancel the interfaces used for interconnection between routers in the domain.

Configuration Example

Building a RIP Routing Domain

Scenario

Figure 1- 7



Remarks	The interface IP addresses are as follows: A: GE0/1 110.11.2.1/24 GE0/2 155.10.1.1/24 B: GE0/1 110.11.2.2/24 GE0/2 196.38.165.1/24 C: GE0/1 110.11.2.3/24 GE0/2 117.102.0.1/16

Configuration Steps

- Configure the interface IP addresses on all routers.
- Configure the RIP basic functions on all routers.

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip address 110.11.2.1 255.255.255.0
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)# ip address 155.10.1.1 255.255.255.0
A(config)# router rip
A(config-router)# version 2
A(config-router)# network 0.0.0.0 255.255.255.255
A(config-router)# passive-interface default
A(config-router)# no passive-interface GigabitEthernet 0/1
```

B

```
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip address 110.11.2.2 255.255.255.0
```

	B(config-if-GigabitEthernet 0/1)# exit B(config)# interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)# ip address 196.38.165.1 255.255.255.0 B(config-if-GigabitEthernet 0/2)# exit B(config)# router rip B(config-router)# version 2 B(config-router)# network 0.0.0.0 255.255.255.255 B(config-router)# passive-interface default B(config-router)# no passive-interface GigabitEthernet 0/1
C	C# configure terminal C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 110.11.2.3 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# interface GigabitEthernet 0/2 C(config-if-GigabitEthernet 0/2)# ip address 117.102.0.1 255.255.0.0 C(config-if-GigabitEthernet 0/2)# exit C(config)# router rip C(config-router)# version 2 C(config-router)#no auto-summary C(config-router)# network 0.0.0.0 255.255.255.255 C(config-router)# passive-interface default C(config-router)# no passive-interface GigabitEthernet 0/1
Verification	Check the routing tables on Router A, Router B, and Router C. Verify that RIP learns the routes to remote networks (contents marked in blue).
A	A# show ip route Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default

	Gateway of last resort is no set C 110.11.2.0/24 is directly connected, GigabitEthernet 0/1 C 110.11.2.1/32 is local host. R 117.0.0.0/8 [120/1] via 110.11.2.2, 00:00:47, GigabitEthernet 0/1 C 155.10.1.0/24 is directly connected, GigabitEthernet 0/2 C 155.10.1.1/32 is local host. C 192.168.217.0/24 is directly connected, VLAN 1 C 192.168.217.233/32 is local host. R 196.38.165.0/24 [120/1] via 110.11.2.3, 00:19:18, GigabitEthernet 0/1
B	B# show ip route Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 110.11.2.0/24 is directly connected, GigabitEthernet 0/1 C 110.11.2.2/32 is local host. R 155.10.0.0/16 [120/1] via 110.11.2.1, 00:15:21, GigabitEthernet 0/1 C 196.38.165.0/24 is directly connected, GigabitEthernet 0/2 C 196.38.165.1/32 is local host. R 117.0.0.0/8 [120/1] via 110.11.2.2, 00:00:47, GigabitEthernet 0/1
C	C# show ip route Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default

	Gateway of last resort is no set C 110.11.2.0/24 is directly connected, GigabitEthernet 0/1 C 110.11.2.3/32 is local host. C 117.102.0.0/16 is directly connected, GigabitEthernet 0/2 C 117.102.0.1/32 is local host. R 155.10.0.0/16 [120/1] via 110.11.2.1, 00:20:55, GigabitEthernet 0/1 R 196.38.165.0/24 [120/1] via 110.11.2.3, 00:19:18, GigabitEthernet 0/1

Common Errors

- The IPv4 address is not configured on an interface.
- The RIP version is not defined on a device, or the RIP version on the device is different from that on other routers.
- The address range configured by the **network** command does not cover a specific interface.
- The **wildcard** parameter in the **network** command is not correctly configured. **0** indicates accurate matching, and **1** indicates that no comparison is performed.
- The interface used for interconnection between devices is configured as a passive interface.

1.4.2 Controlling Interaction of RIP Packets

Configuration Effect

Change the default running mechanism of RIP through configuration and manually control the interaction mode of RIP packets, including:

- Allowing or prohibiting the sending of unicast RIP packets to a specified neighbor on an interface
- Allowing or prohibiting the sending of unicast RIPv2 packets instead of broadcast packets to a specified neighbor on an interface
- Allowing or prohibiting the receiving of RIP packets on an interface
- Allowing or prohibiting the sending of RIP packets on an interface
- Allowing or prohibiting the receiving of RIP packets of a specified version on an interface
- Allowing or prohibiting the sending of RIP packets of a specified version on an interface

Notes

- The RIP basic functions must be configured.
- On an interface connecting to a neighbor device, the configured version of sent RIP packets must be the same as the version of received RIP packets.

Configuration Steps

Sending Unicast RIP Route Update Packets to a Specified Neighbor

- Configure this function if you wish that only some of devices connected to an interface can receive the updated routing information.

- By default, RIPv1 uses the IP broadcast address (255.255.255.255) to advertise the routing information, whereas RIPv2 uses the multicast address (224.0.0.9) to advertise the routing information. If you do not wish all devices on the broadcast network or NBMA network to receive routing information, configure the related interface as the passive interface and specify the neighbors that can receive the routing information. This command does not affect the receiving of RIP packets. RIPv2 packets are broadcast on an interface.
- Unless otherwise required, this function must be enabled on a router that sends the unicast Update packets.

Broadcasting RIPv2 Packets on an Interface

- This function must be configured if the neighbor router does not support the receiving of multicast RIPv2 packets.
- Unless otherwise required, this function must be configured on every router interface that broadcasts RIPv2 packets.

Allowing an Interface to Receive RIP Packets

- This function is enabled by default, and must be disabled if an interface is not allowed to receive RIP packets.
- Unless otherwise required, this function must be configured on every router interface that is not allowed to receive RIP packets.

Allowing an Interface to Send RIP Packets

- This function is enabled by default, and must be disabled if an interface is not allowed to send RIP packets.
- Unless otherwise required, this function must be configured on every router interface that is not allowed to send RIP packets.

Allowing an Interface to Send RIP Packets of a Specified Version

- This function must be configured if the version of RIP packets that can be sent on an interface is required to be different from the global configuration.
- Unless otherwise required, this function must be configured on every router interface that is allowed to send RIP packets of a specified version.

Allowing an Interface to Receive RIP Packets of a Specified Version

- This function must be configured if the version of RIP packets that can be received on an interface is required to be different from the global configuration.
- Unless otherwise required, this function must be configured on every router interface that is allowed to receive RIP packets of a specified version.

Verification

Run the **debug ip rip packet** command to verify the packet sending result and packet type.

Related Commands

Sending Unicast RIP Route Update Packets to a Specified Neighbor

Command Syntax	**neighbor** *ip-address*
Parameter Description	*ip-address:* Indicates the IP address of the neighbor. It should be the address of the network directly connected to the local device.
Command Mode	Routing process configuration mode

Configuration Usage	Generally, you can first run the **passive-interface** command in routing process configuration mode to configure the related interface as a passive interface, and then specify the neighbors that can receive the routing information. This command does not affect the receiving of RIP packets. After an interface is configured as a passive interface, the interface does not send the request packets even after the device is restarted.

Broadcasting RIPv2 Packets on an Interface

Command Syntax	**ip rip v2-broadcast**
Parameter Description	N/A
Command Mode	Interface configuration mode
Configuration Usage	The default behavior is determined by the configuration of the **version** command. The configuration result of this command can overwrite the default configuration of the **version** command. This command affects the behavior of sending RIP packets on the current interface, and the interface is allowed to send RIPv1 and RIPv2 packets simultaneously. If this command does not contain any parameter, the behavior of receiving RIP packets is determined by the configuration of the **version** command.

Allowing an Interface to Receive RIP Packets

Command Syntax	**ip rip receive enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Configuration Usage	To prohibit the receiving of RIP packets on an interface, use the **no** form of this command. This command takes effect only on the current interface. You can use the **default** form of the command to restore the default setting, that is, allowing the interface to receive RIP packets.

Allowing an Interface to Send RIP Packets

Command Syntax	**ip rip send enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Configuration Usage	To prohibit the sending of RIP packets on an interface, use the **no** form of this command in interface configuration mode. This command takes effect only on the current interface. You can use the **default** form of the command to restore the default setting, that is, allowing the interface to send RIP packets.

Allowing an Interface to Send RIP Packets of a Specified Version

Command Syntax	**ip rip send version** [**1**] [**2**]

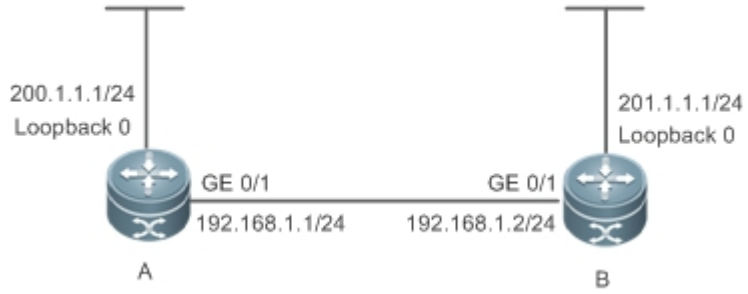
Parameter Description	**1**: Indicates that only RIPv1 packets are sent. **2**: Indicates that only RIPv2 packets are sent.
Command Mode	Interface configuration mode
Configuration Usage	The default behavior is determined by the configuration of the **version** command. The configuration result of this command can overwrite the default configuration of the **version** command. This command affects the behavior of sending RIP packets on the current interface, and the interface is allowed to send RIPv1 and RIPv2 packets simultaneously. If this command does not contain any parameter, the behavior of receiving RIP packets is determined by the configuration of the **version** command.

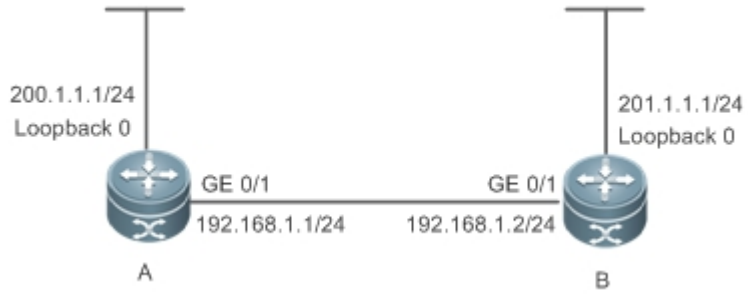
Allowing an Interface to Receive RIP Packets of a Specified Version

Command Syntax	**ip rip receive version** [**1**] [**2**]
Parameter Description	**1**: Indicates that only RIPv1 packets are received. **2**: Indicates that only RIPv2 packets are received.
Command Mode	Interface configuration mode
Configuration Usage	The default behavior is determined by the configuration of the **version** command. The configuration result of this command can overwrite the default configuration of the **version** command. This command affects the behavior of receiving RIP packets on the current interface, and the interface is allowed to receive RIPv1 and RIPv2 packets simultaneously. If this command does not contain any parameter, the behavior of receiving RIP packets is determined by the configuration of the **version** command.

Configuration Example

Prohibiting an Interface from Sending RIP Packets

Scenario **Figure 1-8**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Prohibit the sending of RIP packets on an interface of Router A.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# no ip rip send enable
Verification	Run the **debug ip rip packet send** command on Router A, and verify that packets cannot be sent.
A	A# debug ip rip packet recv

Scenario **Figure 1-8**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Prohibit the sending of RIP packets on an interface of Router A.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# no ip rip send enable
Verification	Run the **debug ip rip packet send** command on Router A, and verify that packets cannot be sent.
	*Nov 4 08:19:31: %RIP-7-DEBUG: [RIP] Prepare to send BROADCAST response... *Nov 4 08:19:31: %RIP-7-DEBUG: [RIP] Building update entries on GigabitEthernet 0/1 *Nov 4 08:19:31: %RIP-7-DEBUG: 117.0.0.0/8 via 0.0.0.0 metric 1 tag 0 *Nov 4 08:19:31: %RIP-7-DEBUG: [RIP] Interface GigabitEthernet 0/1 is disabled to send RIP packet!

Common Errors

A compatibility error occurs because the RIP version configured on the neighbor is different from that configured on the local device.

1.4.3 Enabling Triggered Updates

Configuration Effect

- Enable the RIP triggered updates function, after which RIP does not periodically send the route update packets.

Notes

- The RIP basic functions must be configured.
- It is recommended that split horizon with poisoned reverse be enabled; otherwise, invalid routing information may exist.
- Ensure that the triggered updates function is enabled on every router on the same link; otherwise, the routing information cannot be exchanged properly.

Configuration Steps

Enabling Triggered Updates

- This function must be enabled if demand circuits are configured on the WAN interface.
- The triggered updates function can be enabled in either of the following cases: (1) The interface has only one neighbor; (2) The interface has multiple neighbors but the device interacts with these neighbors in unicast mode.

- It is recommended that triggered updates be enabled on a WAN interface (running the PPP, Frame Relay, or X.25 link layer protocol) to meet the requirements of demand circuits.
- If the triggered updates function is enabled on an interface, source address verification is performed no matter whether the source address verification function is enabled by the **validate-update-source** command.
- Unless otherwise required, triggered updates must be enabled on demand circuits of every router.

Verification

When the RIP triggered updates function is enabled, RIP cannot periodically send the route update packets. RIP sends the route update packets to the WAN interface only in one of the following cases:

- A route request packet is received.
- The RIP routing information changes.
- The interface state changes.
- The router is started.

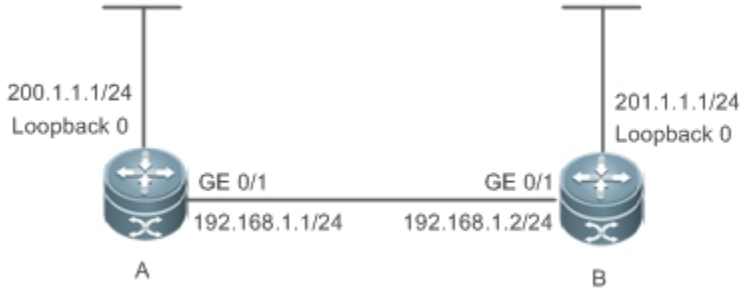
Related Commands

Enabling Triggered Updates

Command Syntax	**ip rip triggered** { **retransmit-timer** *timer* | **retransmit-count** *count* }
Parameter Description	**retransmit-timer** *timer***:** Configures the interval at which the update request or update response packet is retransmitted. The default value is 5s. The value ranges from 1 to 3,600. **retransmit-count** *count***:** Configures the maximum retransmission times of the update request or update response packet. The default value is 36. The value ranges from 1 to 3,600.
Command Mode	Interface configuration mode
Configuration Usage	You can run the **ip rip triggered** command to enable the RIP triggering function. When this function is enabled, the RIP periodical update function is automatically disabled. Therefore, the acknowledgment and retransmission mechanisms must be used to ensure that the Update packets are successfully sent or received on the WAN. You can use the **retransmit-timer** and **retransmit-count** parameters to specify the retransmission interval and maximum retransmission times of the request and update packets.

Configuration Example

Enabling Triggered Updates

Scenario **Figure 1-9**	

Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● On Router A, enable the RIP triggered updates function, and set the retransmission interval and maximum retransmission times of the request and update packets to 10s and 18, respectively.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# encapsulation ppp A(config-if-GigabitEthernet 0/1)# ip rip triggered A(config-if-GigabitEthernet 0/1)# ip rip triggered retransmit-timer 10 A(config-if-GigabitEthernet 0/1)# ip rip triggered retransmit-count 18 A(config-if-GigabitEthernet 0/1)# ip rip split-horizon poisoned-reverse A(config)# router rip A(config-router)# network 192.168.1.0 A(config-router)# network 200.1.1.0
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# encapsulation ppp B(config-if-GigabitEthernet 0/1)# ip rip triggered B(config-if-GigabitEthernet 0/1)# ip rip split-horizon poisoned-reverse B(config)# router rip B(config-router)# network 192.168.1.0 B(config-router)# network 201.1.1.0
Verification	On Router A and Router B, check the RIP database and verify that the corresponding routes are permanent.
A	A# sho ip rip database 201.1.1.0/24 auto-summary 201.1.1.0/24 [1] via 192.168.12.2 GigabitEthernet 0/1 06:25 permanent
B	B# sho ip rip database 200.1.1.0/24 auto-summary 200.1.1.0/24 [1] via 192.168.12.1 GigabitEthernet 0/1 06:25 permanent

Common Errors

- The triggered updates function is enabled when the RIP configurations at both ends of the link are consistent.

- The triggered updates function is not enabled on all routers on the same link.

1.4.4 Enabling Source Address Verification

Configuration Effect

- The source address of the received RIP route update packet is verified.

Notes

- The RIP basic functions must be configured.

Configuration Steps

Enabling Source Address Verification

- This function is enabled by default, and must be disabled when source address verification is not required.
- After split horizon is disabled on an interface, the RIP routing process will perform source address verification on the Update packet no matter whether the **validate-update-source** command is executed in routing process configuration mode.
- For an IP unnumbered interface, the RIP routing process does not perform source address verification on the Update packet no matter whether the **validate-update-source** command is executed in routing process configuration mode.
- Unless otherwise required, this function must be disabled on every router that does not requires source address verification.

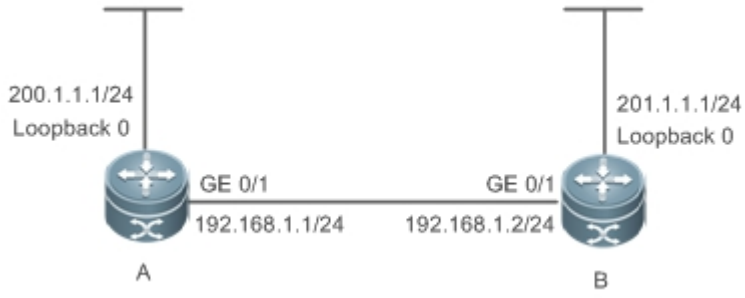
Verification

Only the route update packets coming from the same IP subnet neighbor are received.

Related Commands

Command Syntax	**validate-update-source**
Parameter Description	N/A
Command Mode	Routing process configuration mode
Configuration Usage	Source address verification of the Update packet is enabled by default. After this function is enabled, the source address of the RIP route update packet is verified. The purpose is to ensure that the RIP routing process receives only the route update packets coming from the same IP subnet neighbor.

Configuration Example

Scenario **Figure 1- 10**	

Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Disable source address verification of Update packets on all routers.
A	A# configure terminal A(config)# router rip A(config-router)# no validate-update-source
B	B# configure terminal B(config)# router rip B(config-router)# no validate-update-source
Verification	● On Router A, check the routing table and verify that the entry 201.1.1.0/24 is loaded. ● On Router B, check the routing table and verify that the entry 200.1.1.0/24 is loaded.
A	A# show ip route rip R 201.1.1.0/24 [120/1] via 192.168.2.2, 00:06:11, GigabitEthernet 0/1
B	B# show ip route rip R 200.1.1.0/24 [120/1] via 192.168.1.1, 00:06:11, GigabitEthernet 0/1

1.4.5 Enabling Authentication

Configuration Effect

- Prevent learning unauthenticated and invalid routes and advertising valid routes to unauthorized devices, ensuring stability of the system and protecting the system against intrusions.

Notes

- The RIP basic functions must be configured.
- Only RIPv2 supports authentication of RIP packets, and RIPv1 does not.

Configuration Steps

Enabling Authentication and Specifying the Key Chain Used for RIP Authentication

- This configuration is mandatory if authentication must be enabled.
- If the key chain is already specified in the interface configuration, run the **key chain** command in global configuration mode to define the key chain; otherwise, authentication of RIP packets may fail.
- Unless otherwise required, this configuration must be performed on every router that requires authentication.

Defining the RIP Authentication Mode

- This configuration is mandatory if authentication must be enabled.
- The RIP authentication modes configured on all devices that need to directly exchange RIP routing information must be the same; otherwise, RIP packets may fail to be exchanged.

- If plain text authentication is used, but the key chain for plain text authentication is not configured or associated, authentication is not performed. Similarly, if MD5 authentication is used, but the key chain is not configured or associated, authentication is not performed.
- Unless otherwise required, this configuration must be performed on every router that requires authentication.

Enabling RIP Plain Text Authentication and Configuring the Key Chain

- This configuration is mandatory if authentication must be enabled.
- If RIP plain text authentication should be enabled, use this command to configure the key chain for plain text authentication. Alternatively, you can obtain the key chain for plain text authentication by associating the key chain. The key chain obtained using the second method takes precedence over that obtained using the first method.
- Unless otherwise required, this configuration must be performed on every router that requires authentication.

Verification

- RIP plain text authentication provides only limited security because the password transferred through the packet is visible.
- RIP MD5 authentication can provide higher security because the password transferred through the packet is encrypted using the MD5 algorithm.
- Routes can be learned properly if the correct authentication parameters are configured.
- Routes cannot be learned if the incorrect authentication parameters are configured.

Related Commands

Enabling Source Address Verification

Command Syntax	**ip rip authentication key-chain** *name-of-keychain*
Parameter Description	*name-of-keychain*: Specifies the name of the key chain used for RIP authentication.
Command Mode	Interface configuration mode
Configuration Usage	The specified key chain must be defined by the **key chain** command in global configuration mode in advance.

Defining the RIP Authentication Mode

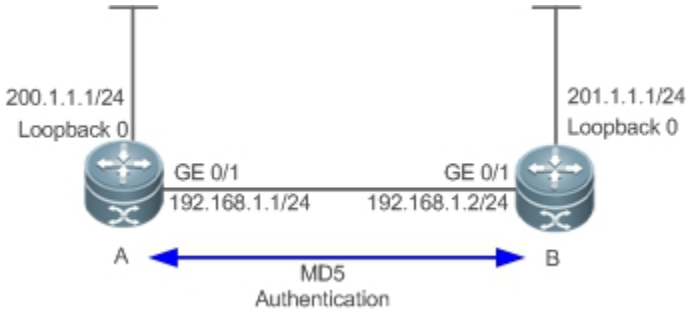
Command Syntax	**ip rip authentication mode** { **text** | **md5** }
Parameter Description	**text:** Indicates that the RIP authentication mode is plain text authentication. **md5:** Indicates that the RIP authentication mode is MD5 authentication.
Command Mode	Interface configuration mode
Configuration Usage	For all devices that need to directly exchange the RIP routing information, the RIP authentication mode of these devices must be the same.

Enabling RIP Plain Text Authentication and Configuring the Key Chain

Command Syntax	**ip rip authentication text-password** [**0** | **7**] *password-string*
Parameter Description	**0**: Indicates that the key is displayed in plain text. **7**: Indicates that the key is displayed in cipher text. *password-string*: Indicates the key chain used for plain text authentication. The key chain is a string of 1 to 16 bytes.
Command Mode	Interface configuration mode
Configuration Usage	This commands takes effect only in plain text authentication mode.

Configuration Example

Configuring RIP Basic Functions and Enabling MD5 Authentication

Scenario **Figure 1- 11**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Configure the authentication type and MD5 authentication key on all routers.
A	A# configure terminal A(config)# key chain hello A(config-keychain)# key 1 A(config-keychain-key)# key-string world A(config-keychain-key)# exit A(config-keychain)# exit A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip rip authentication mode md5 A(config-if-GigabitEthernet 0/1)# ip rip authentication key-chain hello
B	B# configure terminal B(config)# key chain hello B(config-keychain)# key 1 B(config-keychain-key)# key-string world B(config-keychain-key)# exit B(config-keychain)# exit

	B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip rip authentication mode md5 B(config-if-GigabitEthernet 0/1)# ip rip authentication key-chain hello
Verification	● On Router A, check the routing table and verify that the entry 201.1.1.0/24 is loaded. ● On Router B, check the routing table and verify that the entry 200.1.1.0/24 is loaded.
A	A# show ip route rip R 201.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
B	A# show ip route rip R 200.1.1.0/24 [120/1] via 192.168.1.1, 00:06:11, GigabitEthernet 0/1

Common Errors

- The keys configured on routers that need to exchange RIP routing information are different.
- The authentication modes configured on routers that need to exchange RIP routing information are different.

1.4.6 Enabling Route Summarization

Configuration Effect

Reduce the size of the routing table, improve the routing efficiency, avoid route flapping to some extent, and improve scalability and effectiveness of the network.

If a summarized route exists, subroutes included by the summarized route cannot be seen in the routing table, which greatly reduces the size of the routing table.

Advertising a summarized route is more efficient than advertising individual routes because: (1) A summarized route is processed first when RIP looks through the database; (2) All subroutes are ignored when RIP looks through the database, which reduces the processing time required.

Notes

- The RIP basic functions must be configured.
- The range of supernetting routes is larger than that of the classful network. Therefore, the automatic route summarization function is invalid for supernetting routes.
- RIPv1 always performs automatic route summarization. If the detailed routes should be advertised, you must set the RIP version to RIPv2.

Configuration Steps

Enabling Automatic Route Summarization

- This function is enabled by default.
- To learn specific subnet routes instead of summarized network routes, you must disable automatic route summarization.
- You can disable automatic route summarization only in RIPv2. RIPv1 always performs automatic route summarization.

Configuring RIP Route Summarization on an Interface

- This function must be configured if it is required to summarize classful subnets.
- The **ip rip summary-address** command is used to summarize an address or a subnet under a specified interface. RIP automatically summarizes to the classful network boundary. Each classful subnet can be configured only in the **ip rip summary-address** command.
- The summary range configured in this command cannot be supernetting routes, that is, the configured subnet mask length cannot be smaller than the natural mask length of the network.
- Unless otherwise required, this configuration should be performed on a router that requires classful subnet summarization.

Verification

Verify that the routes are summarized in the routing table of the peer end.

Related Commands

Enabling Automatic Route Summarization

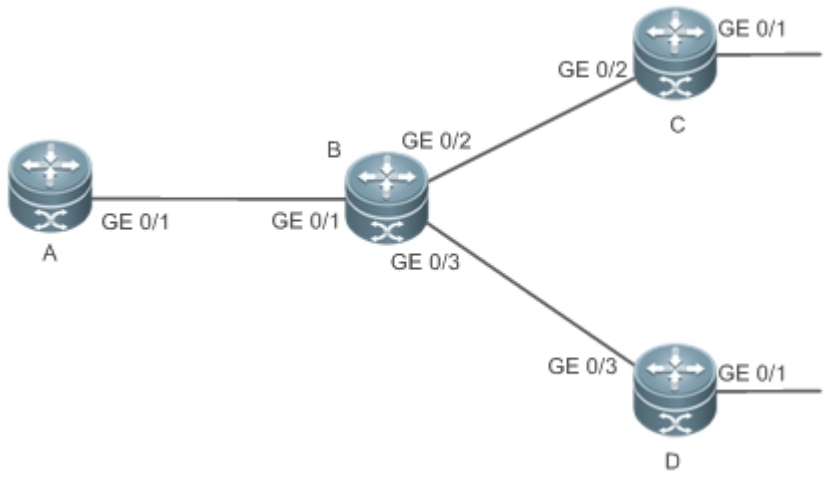
Command Syntax	**auto-summary**
Parameter Description	N/A
Command Mode	Routing process configuration mode
Configuration Usage	Route summarization is enabled by default for RIPv1 and RIPv2. You can disable automatic route summarization only in RIPv2. RIPv1 always performs automatic route summarization.

Configuring RIP Route Summarization on an Interface

Command Syntax	**ip rip summary-address** *ip-address ip-network-mask*
Parameter Description	*ip-address*: Indicates the IP address to be summarized. *ip-network-mask*: Indicates the subnet mask of the IP address to be summarized.
Command Mode	Interface configuration mode
Configuration Usage	This command is used to summarize an address or a subnet under a specified interface.

Configuration Example

Configuring Route Summarization

Scenario **Figure 1- 12**	 **Remarks**: The interface IP addresses are as follows: A: GE0/1 192.168.1.1 B: GE0/1 192.168.1.2 GE0/2 172.16.2.1 GE0/3 172.16.3.1 C: GE0/2 172.16.2.2 GE0/3 172.16.4.2 D: GE0/2 172.16.3.2 GE0/3 172.16.5.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Configure route summarization on Router B.
	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip rip summary-address 172.16.0.0 255.255.0.0 B(config)# router rip B(config-router)# version 2 B(config-router)# no auto-summary
Verification	Check the routing table on Router A, and verify that the entry 172.16.0.0/16 is generated.
	A# show ip route rip R 172.16.0.0/16 [120/2] via 192.168.1.2, 00:01:04, GigabitEthernet 0/1

Common Errors

- RIP basic functions are not configured or fail to be configured.

1.4.7 Enabling Supernetting Routes

Configuration Effect

- Allow RIP to send RIP supernetting routes on a specified interface.

Notes

- The RIP basic functions must be configured.

Configuration Steps

Enabling Supernetting Routes

- If a supernetting route is detected when a RIPv1-enabled router monitors the RIPv2 route response packets, the router will learn an incorrect route because RIPv1 ignores the subnet mask in the routing information of the packet. In this case, the **no** form of the command must be used on the RIPv2-enabled router to prohibit advertisement of supernetting routes on the related interface. This command takes effect only on the current interface.
- The command is effective only when RIPv2 packets are sent on the interface, and is used to control the sending of supernetting routes.

Verification

Verify that the peer router cannot learn the supernetting route.

Related Commands

Command Syntax	**ip rip send supernet-routes**
Parameter Description	N/A
Command Mode	Interface configuration mode
Configuration Usage	By default, an interface is allowed to send RIP supernetting routes.

Configuration Example

Disabling Supernetting Routes

Scenario **Figure 1- 13**	
Configuration Steps	• Configure the interface IP addresses on all routers. (Omitted) • Configure the RIP basic functions on all routers. (Omitted) • Prohibit the sending of RIP supernetting routes on the GigabitEthernet 0/1 interface of Router B.

	B# configure terminal B(config)# ip route 207.0.0.0 255.0.0.0 Null 0 B(config)# ip route 208.1.1.0 255.255.255.0 Null 0 B(config)# router rip B(config-router)# redistribute static B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# no ip rip send supernet-routes
Verification	Check the routing table on Router A, and verify that Router A can learn only the non-supernetting route 208.1.1.0/24, but not the supernetting route 207.0.0.0/8.
	A#show ip route rip R 208.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1

1.4.8 Advertising the Default Route or External Routes

Configuration Effect

- In the RIP domain, introduce a unicast route of another AS so that the unicast routing service to this AS can be provided for users in the RIP domain.
- In the RIP domain, inject a default route to another AS so that the unicast routing service to this AS can be provided for users in the RIP domain.

Notes

- The RIP basic functions must be configured.
- Route redistribution cannot introduce default routes of other protocols to the RIP routing domain.

Configuration Steps

Advertising the Default Route to Neighbors

This function must be enabled if it is required to advertise the default route to neighbors.

By default, a default route is not generated, and the metric of the default route is 1.

If the RIP process can generate a default route using this command, RIP does not learn the default route advertised by the neighbor.

Unless otherwise required, this configuration should be performed on a router that needs to advertise the default route.

Advertising the Default Route to Neighbors on an Interface

This function must be enabled if it is required to advertise the default route to neighbors on a specified interface.

By default, a default route is not configured and the metric of the default route is 1.

After this command is configured on an interface, a default route is generated and advertised through this interface.

Unless otherwise required, this configuration should be performed on a router that needs to advertise the default route.

Redistributes Routes and Advertises External Routes to Neighbors

This function must be enabled if routes of other protocols need to be redistributed.

By default,

- If OSPF redistribution is configured, redistribute the routes of all sub-types of the OSPF process.
- In other cases, redistribute all external routes.
- The metric of a redistributed route is 1 by default.
- The route map is not associated by default.

During route redistribution, it is not necessary to convert the metric of one routing protocol to the metric of another routing protocol because different routing protocols use completely different metric measurement methods. RIP measures the metric based on the hop count, and OSPF measures the metric based on the bandwidth. Therefore, the computed metrics cannot be compared with each other. During route redistribution, however, it is necessary to configure a symbolic metric; otherwise, route redistribution fails.

Unless otherwise required, this configuration should be performed on a router that needs to redistribute routes.

Verification

- On a neighbor device, verify that a default route exists in the RIP routing table.
- On the local and neighbor devices, verify that external routes (routes to other ASs) exist in the RIP routing table.

Related Commands

Advertising the Default Route to Neighbors

Command Syntax	**default-information originate** [**always**] [**metric** *metric-value*] [**route-map** *map-name*]
Parameter Description	**always:** Enables RIP to generate a default route no matter whether the local router has a default route. **metric** *metric-value:* Indicates the initial metric of the default route. The value ranges from 1 to 15. **route-map** *map-name:* Indicates the associated route map name. By default, no route map is associated.
Command Mode	Routing process configuration mode
Configuration Usage	If a default route exists in the routing table of a router, RIP does not advertise the default route to external entities by default. You need to run the **default-information originate** command in routing process configuration mode to advertise the default route to neighbors. If the **always** parameter is selected, the RIP routing process advertises a default route to neighbors no matter the default route exists, but this default route is not displayed in the local routing table. To check whether the default route is generated, run the **show ip rip database** command to check the RIP routing information database. To further control the behavior of advertising the RIP default route, use the **route-map** parameter. For example, run the **set metric** rule to set the metric of the default route. You can use the **metric** parameter to set the metric of the advertised default value, but the priority of this configuration is lower than that of the set metric rule of the **route-map** parameter. If the **metric** parameter is not configured, the default route uses the default metric configured for RIP. You still need to run the **default-information originate** command to introduce the default route generated by **ip default-network** to RIP.

Advertising the Default Route to Neighbors on an Interface

Command Syntax	**ip rip default-information** { **only** | **originate** } [**metric** *metric-value*]
Parameter Description	**only**: Indicates that only the default route is advertised. **originate**: Indicates that the default route and other routes are advertised. **metric** *metric-value:* Indicates the metric of the default route. The value ranges from 1 to 15.
Command Mode	Interface configuration mode
Configuration Usage	If you configure the **ip rip default-information** command for the interface, and the **default-information originate** command for the RIP process, only the default route configured for the interface is advertised. So far as **ip rip default-information** is configured for one interface, RIP does not learn the default route advertised by the neighbor.

Redistributes Routes and Advertises External Routes to Neighbors

Command Syntax	**redistribute** { **connected** | **ospf** *process-id* | **static** } [[**metric** *metric-value*] [**route-map** *route-map-name*]
Parameter Description	**connected**: Indicates redistribution from direct routes. **ospf** *process-id*: Indicates redistribution from OSPF. *process-id* indicates the OSPF process ID. The value ranges from 1 to 65535. **static**: Indicates redistribution from static routes. **metric** *metric-value*: Sets the metric of the redistributed route. The value ranges from 1 to 16. **route-map** *route-map-name*: Sets the redistribution filtering rules.
Command Mode	Routing process configuration mode
Configuration Usage	The configuration rules for the **no** form of the **redistribute** command are as follows: 1. If some parameters are specified in the **no** form of the command, default values of these parameters will be restored. 2. If no parameter is specified in the **no** form of the command, the entire command will be deleted.

Configuration Example

Redistributing Routes and Advertising External Routes to Neighbors

Scenario **Figure 1- 14**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● On Router B, configure redistribution of static routes.
B	B# configure terminal B(config)# router rip

	B(config-router)# redistribute static
Verification	On Router A, check the routing table and verify that the entry 172.10.10.0/24 is loaded.
	A# show ip route rip R 172.10.10.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1

1.4.9 Setting Route Filtering Rules

Configuration Effect

- Routes that do not meet filtering criteria cannot be loaded to the routing table, or advertised to neighbors. In this way, users within the network can be prevented from accessing specified destination networks.

Notes

- The RIP basic functions must be configured.
- In regard to the filtering rules of sent routes, you must configure route redistribution first, and then filter the redistributed routes.

Configuration Steps

Filtering the Received RIP Routing Information

- This function must be configured if it is required to filter received routing information.
- To refuse receiving some specified routes, you can configure the route distribution control list to process all the received route update packets. If no interface is specified, route update packets received on all interfaces will be processed.
- Unless otherwise required, this configuration should be performed on a router that requires route filtering.

Filtering the Sent RIP Routing Information

- This function must be configured if it is required to filter the redistributed routing information that is sent.
- If this command does not contain any optional parameter, route update advertisement control takes effect on all interfaces. If the command contains the interface parameter, route update advertisement control takes effect only on the specified interface. If the command contains other routing process parameters, route update advertisement control takes effect only on the specified routing process.
- Unless otherwise required, this configuration should be performed on a router that requires route filtering.

Verification

- Run the **show ip route rip** command to verify that the routes that have been filtered out are not loaded to the routing table.

Related Commands

Filtering the Received RIP Routing Information

Command Syntax	**distribute-list** { [*access-list-number* | *name*] | **prefix** *prefix-list-name* [**gateway** *prefix-list-name*] } **in** [*interface-type interface-number*]
Parameter Description	*access-list-number* | *name*: Specifies the access list. Only routes permitted by the access list can be received. **prefix** *prefix-list-name*: Uses the prefix list to filter routes.

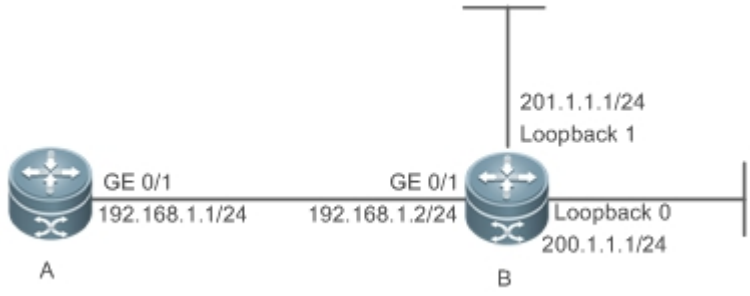
	gateway *prefix-list-name*: Uses the prefix list to filter the route sources. *interface-type interface-number*: Indicates that the distribution list is applied to the specified interface.
Command Mode	Routing process configuration mode
Configuration Usage	N/A

Filtering the Sent RIP Routing Information

Command Syntax	**distribute-list** { [*access-list-number* | *name*] | **prefix** *prefix-list-name* } **out** [*interface* **|** [**connected | ospf** *process-id* **| rip | static**]]
Parameter Description	*access-list-number* | *name*: Specifies the access list. Only routes permitted by the access list can be sent. **prefix** *prefix-list-name*: Uses the prefix list to filter routes. *Interface*: Applies route update advertisement control only on the specified interface. **connected**: Applies route update advertisement control only on direct routes introduced through redistribution. **ospf** *process-id*: Applies route update advertisement control only on the routes introduced from OSPF. *process-id* specifies an OSPF process. **rip**: Applies route update advertisement control only on RIP routes. **static**: Applies route update advertisement control only on static routes introduced through redistribution.
Command Mode	Routing process configuration mode
Configuration Usage	N/A

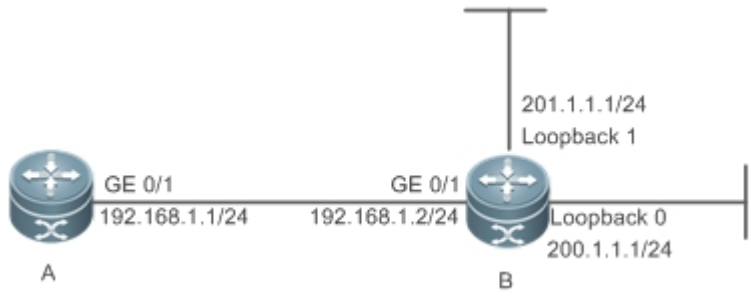
Configuration Example

Filtering the Received RIP Routing Information

Scenario **Figure 1- 15**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Enable the RIP routing process to control routes received over the GigabitEthernet 0/1 port and receive only the route 200.1.1.0.

A	A# configure terminal A(config)# router rip A(config-router)# distribute-list 10 in GigabitEthernet 0/1 A(config-router)# no auto-summary A(config)# access-list 10 permit 200.1.1.0 0.0.0.255
Verification	On Router A, check the routing table and verify that only the entry 200.1.1.0/24 exists.
A	A# show ip route rip R 200.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1

Filtering the Sent RIP Routing Information

Scenario **Figure 1- 16**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Enable the RIP routing process to advertise only the route 200.1.1.0/24.
B	B# configure terminal B(config)# router rip B(config-router)# redistribute connected B(config-router)# distribute-list 10 out B(config-router)# version 2 B(config)# access-list 10 permit 200.1.1.0 0.0.0.255
Verification	Check the routing table on Router A, and verify that route in the 200.1.1.0 network segment exists.
A	A# show ip route rip R 200.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1

Common Errors

- Filtering fails because the filtering rules of the access list are not properly configured.

1.4.10 Modifying Route Selection Parameters

Configuration Effect

- Change the RIP routes to enable the traffic pass through specified nodes or avoid passing through specified nodes.
- Change the sequence that a router selects various types of routes so as to change the priorities of RIP routes.

Notes

- The RIP basic functions must be configured.

Configuration Steps

Modifying the Administrative Distance of a RIP Route

- Optional.
- This configuration is mandatory if you wish to change the priorities of RIP routes on a router that runs multiple unicast routing protocols.

Increasing the Metric of a Received or Sent RIP Route

- Optional.
- Unless otherwise required, this configuration should be performed on a router where the metrics of routes need to be adjusted.

Configuring the Default Metric of an External Route Redistributed to RIP

- Optional.
- Unless otherwise required, this configuration must be performed on an ASBR to which external routes are introduced.

Verification

Run the **show ip rip** command to display the administrative distance currently configured. Run the **show ip rip data** command to display the metrics of redistributed routes to verify that the configuration takes effect.

Related Commands

Modifying the Administrative Distance of a RIP Route

Command Syntax	**distance** *distance* [*ip-address wildcard*]
Parameter Description	*distance*: Sets the administrative distance of a RIP route. The value is an integer ranging from 1 to 255. *ip-address*: Indicates the prefix of the source IP address of the route. *wildcard*: Defines the IP address comparison bit. **0** indicates accurate matching, and **1** indicates that no comparison is performed.
Command Mode	Routing process configuration mode
Configuration Usage	Run this command to configure the administrative distance of a RIP route.

Increasing the Metric of a Received or Sent RIP Route

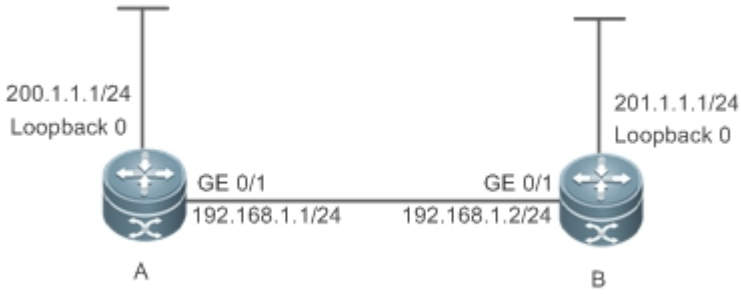
Command Syntax	**offset-list** { *access-list-number* | *name* } { **in** | **out** } *offset* [*interface-type interface-number*]
Parameter Description	*access-list-number* | *name*: Specifies the access list. **In**: Uses the ACL to modify the metric of a received route. **out**: Uses the ACL to modify the metric of a sent route. *offset*: Indicates the offset of the modified metric. The value ranges from 0 to 16. *interface-type*: Uses the ACL on the specified interface. *interface-number*: Specifies the interface number.
Command Mode	Routing process configuration mode
Configuration Usage	Run this command to increase the metric of a received or sent RIP route. If the interface is specified, the configuration takes effect only on the specified interface; otherwise, the configuration takes effect globally.

Configuring the Default Metric of an External Route Redistributed to RIP

Command Syntax	**default-metric** *metric-value*
Parameter Description	*metric-value:* Indicates the default metric. The valid value ranges from 1 to 16. If the value is equal to or greater than 16, the FSOS determines that this route is unreachable.
Command Mode	Routing process configuration mode
Configuration Usage	This command must be used together with the routing protocol configuration command **redistribute**.

Configuration Example

Increasing the Metric of a Received or Sent RIP Route

Scenario **Figure 1- 17**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Increase by 7 the metric of each RIP route in the range specified by ACL 7. ● Increase by 7 the metric of each learned RIP route in the range specified by ACL 8.

A	A# configure terminal A(config)# access-list 7 permit host 200.1.1.0 A(config)# access-list 8 permit host 201.1.1.0 A(config)# router rip A(config-router)# offset-list 7 out 7 A(config-router)# offset-list 8 in 7
Verification	Check the routing table on Router A and Router B to verify that the metrics of RIP routes are 8.
A	A# show ip route rip R 201.1.1.0/24 [120/8] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
B	B# show ip route rip R 200.1.1.0/24 [120/8] via 192.168.1.1, 00:06:11, GigabitEthernet 0/1

1.4.11 Modifying Timers

Configuration Effect

- Change the duration of RIP timers to accelerate or slow down the change of the protocol state or occurrence of an event.

Notes

- The RIP basic functions must be configured.
- Modifying the protocol control parameters may result in protocol running failures. Therefore, you are advised not to modify the timers.

Configuration Steps

Modifying the Update Timer, Invalid Timer, and Flush Timer

This configuration must be performed if you need to adjust the RIP timers.

By adjusting the timers, you can reduce the convergence time and fault rectification time of the routing protocol. For routers connected to the same network, values of the three RIP timers must be the same. Generally, you are advised not to modify the RIP timers unless otherwise required.

Setting timers to small values on a low-speed link brings risks because a lot of Update packets consume the bandwidth. You can set timers to small values generally on the Ethernet or a 2 Mbps (or above) link to reduce the convergence time of network routes.

Unless otherwise required, this configuration should be performed on a router where RIP timers need to be modified.

Setting the Sending Delay Between RIP Route Update Packets

This configuration must be performed if you need to adjust the sending delay between RIP Update packets.

Run the **output-delay** command to increase the sending delay between packets on a high-speed device so that a low-speed device can receive and process all Update packets.

Unless otherwise required, this configuration should be performed on a router where the sending delay needs to be adjusted.

Verification

Run the **show ip rip** command to display the current settings of RIP timers.

Related Commands

Modifying the Update Timer, Invalid Timer, and Flush Timer

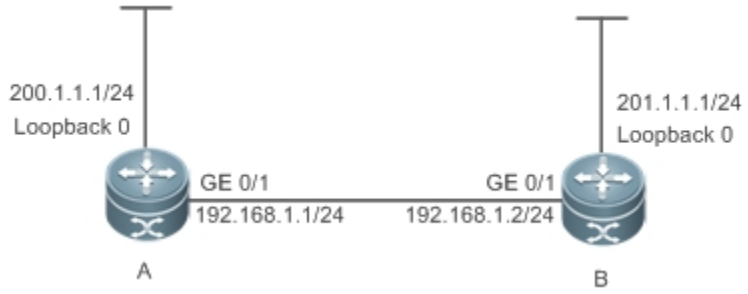
Command Syntax	**timers basic** *update invalid flush*
Parameter Description	*update*: Indicates the route update time in second. It defines the interval at which the device sends the route update packet. Each time an Update packet is received, the invalid timer and flush timer are reset. By default, a routing update packet is sent every 30s. *invalid*: Indicates the route invalid time in second, counted from the last time when a valid update packet is received. It defines the time after which the route in the routing list becomes invalid because the route is not updated. The duration of the invalid timer must be at least three times the duration of the update timer. If no Update packet is received before the invalid timer expires, the corresponding route enters the invalid state. If the Update packet is received before the invalid timer expires, the timer is reset. The default duration of the invalid timer is 180s. *flush*: Indicates the route flushing time in second, counted from the time when the RIP route enters the invalid state. When the flush timer expires, the route in the invalid state will be deleted from the routing table. The default duration of the flush timer is 120s.
Command Mode	Routing process configuration mode
Configuration Usage	By default, the update timer is 30s, the invalid timer is 180s, and the flush timer is 120s.

Setting the Sending Delay Between RIP Route Update Packets

Command Syntax	**output-delay** *delay*
Parameter Description	*delay:* Sets the sending delay between packets in ms. The value ranges from 8 to 50.
Command Mode	Interface configuration mode
Configuration Usage	Normally, a RIP route update packet is 512 bytes long and can contain 25 routes. If the number of routes to be updated exceeds 25, more than one update packet will be sent as fast as possible. When a high-speed device sends a lot of update packets to a low-speed device, the low-speed device may not be able to process all update packets in time, causing a loss of routing information. In this case, you need to run the **output-delay** command to increase the sending delay between packets on a high-speed device so that a low-speed device can receive and process all update packets.

Configuration Example

Setting the Sending Delay Between RIP Route Update Packets

Scenario **Figure 1- 18**	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● Configure the sending delay of update packets on Router A.
A	A# configure terminal A(config)# router rip A(config-router)# output-delay 30
Verification	Capture packets on Router A and compare the sending time of update packets before and after the configuration, and verify that a delay of 30 ms is introduced.

Common Errors

For routers connected to the same network, values of the three RIP timers are not the same.

1.4.12 Enabling GR

Configuration Effect

- When a distributed route switches services from the active board to the standby board, traffic forwarding continues and is not interrupted.
- When the RIP process is being restarted, traffic forwarding continues and is not interrupted.

Notes

- The RIP basic functions must be configured.
- The GR period is at least twice the RIP route update period.
- During the RIP GR process, ensure that the network environment is stable.

Configuration Steps

Configuring the GR Restarter Capability

This configuration must be performed if RIP needs to be gracefully restarted to ensure data forwarding during hot standby switchover.

The GR function is configured based on the RIP process. You can configure different parameters for different RIP processes based on the actual conditions.

The GR period is the maximum time from restart of the RIP process to completion of GR. During this period, the forwarding table before the restart is retained, and the RIP route is restored so as to restore the RIP state before the restart. After the restart period expires, RIP exits from the GR state and performs common RIP operations.

Unless otherwise required, this configuration should be performed on every router that needs to be gracefully restarted.

Verification

- Run the **show ip rip** command to display the GR state and configured time.
- Trigger a hot standby switchover, and verify that data forwarding is not interrupted.

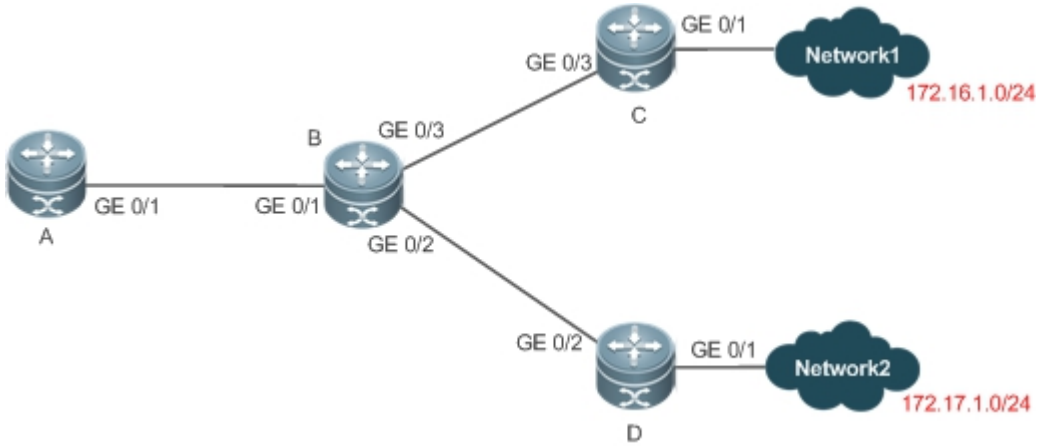
Related Commands

Configuring the GR Restarter Capability

Command Syntax	**graceful-restart** [**grace-period** *grace-period*]
Parameter Description	**graceful-restart**: Enables the GR function. **grace-period**: Explicitly configures the grace period. *grace-period*: Indicates the GR period. The value ranges from 1s to 1800s. The default value is twice the update time or 60s, whichever is the smaller.
Command Mode	Routing process configuration mode
Configuration Usage	This command allows you to explicitly modify the GR period. Note that GR must be completed after the update timer of the RIP route expires and before the invalid timer of the RIP route expires. An inappropriate GR period cannot ensure uninterrupted data forwarding during the GR process. A typical case is as follows: If the GR period is longer than the duration of the invalid timer, GR is not completed when the invalid timer expires. The route is not re-advertised to the neighbor, and forwarding of the route of the neighbor stops after the invalid timer expires, causing interruption of data forwarding on the network. Unless otherwise required, you are advised not to adjust the GR period. If it is necessary to adjust the GR period, ensure that the GR period is longer than the duration of the update timer but shorter than the duration of the invalid timer based on the configuration of the **timers basic** command.

Configuration Example

Configuring the GR Restarter Capability

Scenario **Figure 1- 19**	 **Remarks**: The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 B: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 GE 0/3 192.168.3.1 C: GE 0/1 192.168.4.2 GE 0/3 192.168.3.2 D: GE 0/1 192.168.5.2 GE 0/2 192.168.2.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the RIP basic functions on all routers. (Omitted) ● On Router B, enable the GR function.
	B# configure terminal B(config)# router rip B(config-router)# graceful-restart grace-period 90
Verification	● Trigger a hot standby switchover on Router B, and verify that the routing tables of destination Network 1 and Network 2 remain unchanged on Router A during the switchover. ● Trigger a hot standby switchover on Router B, ping destination Network 1 from Router A, and verify that traffic forwarding is not interrupted during the switchover.

1.5 Monitoring

Displaying

Description	Command
Displays the basic information about a RIP process.	**show ip rip**
Displays the RIP routing table.	**show ip rip database** [*network-number network-mask*] [**count**]
Displays information about external routes redistributed by RIP.	**show ip rip external** [**connected** | **ospf** *process-id* | **static**]
Displays the RIP interface information.	**show ip rip interface** [*interface-type interface-number*]
Displays the RIP neighbor information.	**show ip rip peer** [*ip-address*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs events that occur when the RIP process is running.	**debug ip rip event**
Debugs interaction with the NSM process.	**debug ip rip nsm**
Debugs the sent and received packets.	**debug ip rip packet** [**interface** *interface-type interface-number* | **recv** | **send**]
Debugs the RIP GR process.	**debug ip rip restart**
Debugs the route changes of the RIP process.	**debug ip rip route**

2 Configuring OSPFv2

2.1 Overview

Open Shortest Path First (OSPF) is an Interior Gateway Protocol (IGP) that is used within the Autonomous System (AS) to allow routers to obtain a route to a remote network.

OSPF Version 2 (OSPFv2) is applicable to IPv4, and OSPF Version 3 (OSPFv3) is applicable to IPv6. The protocol running mechanism and most configurations are the same.

OSPF has the following characteristics:

- Wide scope of application: OSPF is applicable to a larger-scale network that supports hundreds of routers.
- Fast convergence: Once the network topology changes, notifications can be quickly sent between routers to update routes.
- No self-loop: Only the link status information is synchronized between routers. Each router computes routes independently, and a self-loop will not occur.
- Area division: A large routing domain is divided into multiple small areas to save system resources and network bandwidth and ensure stability and reliability of routes.
- Route classification: Routes are classified into several types to support flexible control.
- Equivalent routes: OSPF supports equivalent routes.
- Authentication: OSPF supports packet authentication to ensure security of protocol interaction.
- Multicast transmission: Protocol packets are sent using the multicast address to avoid interfering with irrelevant entities and save system resources.

In this chapter, the term "router" refers to any network device that supports the routing function. These network devices can be L3 switches, routers, or firewall.

Unless otherwise specified, "OSPF" in the following descriptions refers to OSPFv2.

Protocols and Standards

RFC2328	This memo documents version 2 of the OSPFprotocol. OSPF is a link-state routing protocol.
RFC 2370	This memo defines enhancements to the OSPFprotocol to support a new class of link-stateadvertisements (LSA) called Opaque LSAs.Opaque LSAs provide a generalized mechanismto allow for the future extensibility of OSPF.
RFC3137	This memo describes a backward-compatibletechnique that may be used by OSPF (OpenShortest Path First) implementations to advertiseunavailability to forward transit traffic or to lowerthe preference level for the paths through such arouter.
RFC3623	This memo documents an enhancement to theOSPF routing protocol, whereby an OSPF routercan stay on the forwarding path even as its OSPFsoftware is restarted.
RFC3630	This document describes extensions to the OSPFprotocol version 2 to support intra-area TrafficEngineering (TE), using Opaque Link StateAdvertisements.
RFC3682	The use of a packet's Time to Live (TTL) (IPv4)or Hop Limit (IPv6) to protect a protocol stackfrom CPU-utilization based attacks has beenproposed in many settings.

RFC3906	This document describes how conventional hop-by-hop link-state routing protocols interact withnew Traffic Engineering capabilities to createInterior Gateway Protocol (IGP) shortcuts.
RFC4576	This document specifies the necessary procedure,using one of the options bits in the LSA (Link StateAdvertisements) to indicate that an LSA hasalready been forwarded by a PE and should beignored by any other PEs that see it.
RFC4577	This document extends that specification byallowing the routing protocol on the PE/CEinterface to be the OSPF protocol.
RFC4750	This memo defines a portion of the ManagementInformation Base (MIB) for use with networkmanagement protocols in TCP/IP-based Internets.In particular, it defines objects for managingversion 2 of the Open Shortest Path First RoutingProtocol. Version 2 of the OSPF protocol is specific to the IPv4 address family.

2.2 Applications

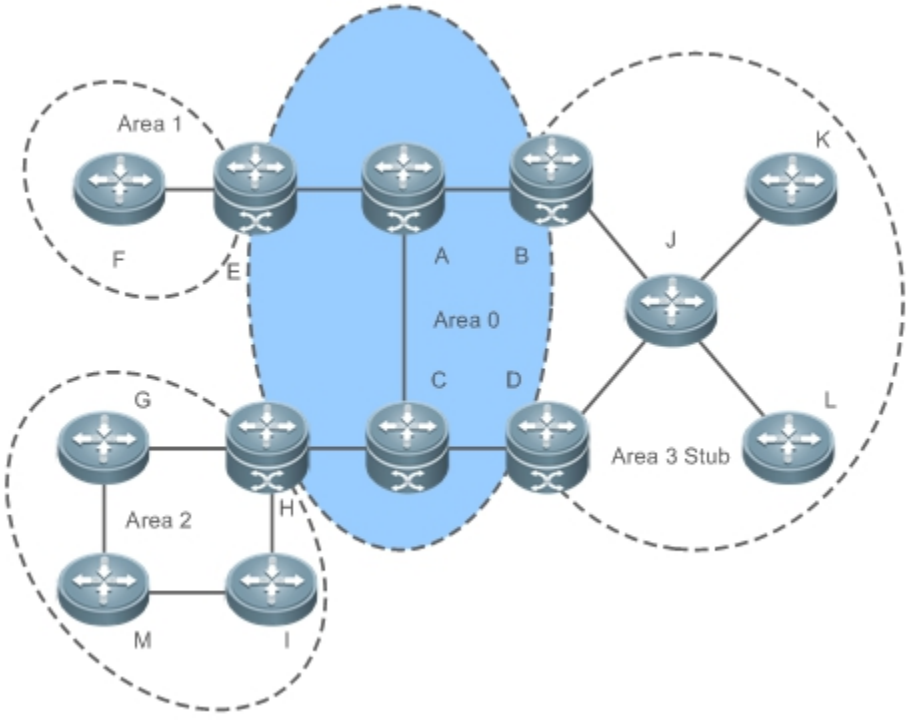
Application	Description
Intra-Domain Interworking	OSPF runs within the AS, which is divided into several areas.
Inter-Domain Interworking	Several ASs are interconnected. OSPF runs within each AS.

2.2.1 Intra-Domain Interworking

Scenario

OSPF runs within the AS. If the number of routers exceeds 40, it is recommended that the AS be divided into several areas. Generally, high-end devices featuring reliable performance and fast processing speed are deployed in a backbone area, and low-end or medium-range devices with relatively lower performance can be deployed in a normal area. All normal areas must be connected to the backbone area. It is recommended that a normal arealocated on the stub be configured as a stub area. As shown in Figure 2- 1, the network is divided into four areas. Communication between these areas must go through the backbone area, that is area 0.

Figure 2- 1 Division of the OSPF Areas



Remarks	A, B, C, D, E, and H are located in the backbone area, and are backbone routers. Area 3 is configured as a stub area.

Deployment

- OSPF runs on all routers within the AS to implement unicast routing.

2.3 Features

Basic Concepts

Routing Domain

All routers in an AS must be interconnected and use the same routing protocol. Therefore, the AS is also called routing domain.

An AS on which OSPF runs is also called OSPF routing domain, or OSPF domain for short.

OSPF Process

OSPF supports multiple instances, and each instance corresponds to an OSPF process.

One or more OSPF processes can be started on a router. Each OSPF process runs OSPF independently, and the processes are mutually isolated.

The process ID takes effect only on the local router, and does not affect exchange of OSPF packets on adjacent interfaces.

RouterID

The router ID uniquely identifies a router in an OSPF domain. Router IDs of any two routers cannot be the same.

If multiple OSPF processes exist on a router, each OSPF process uses one router ID. Router IDs of any two OSPF processes cannot be the same.

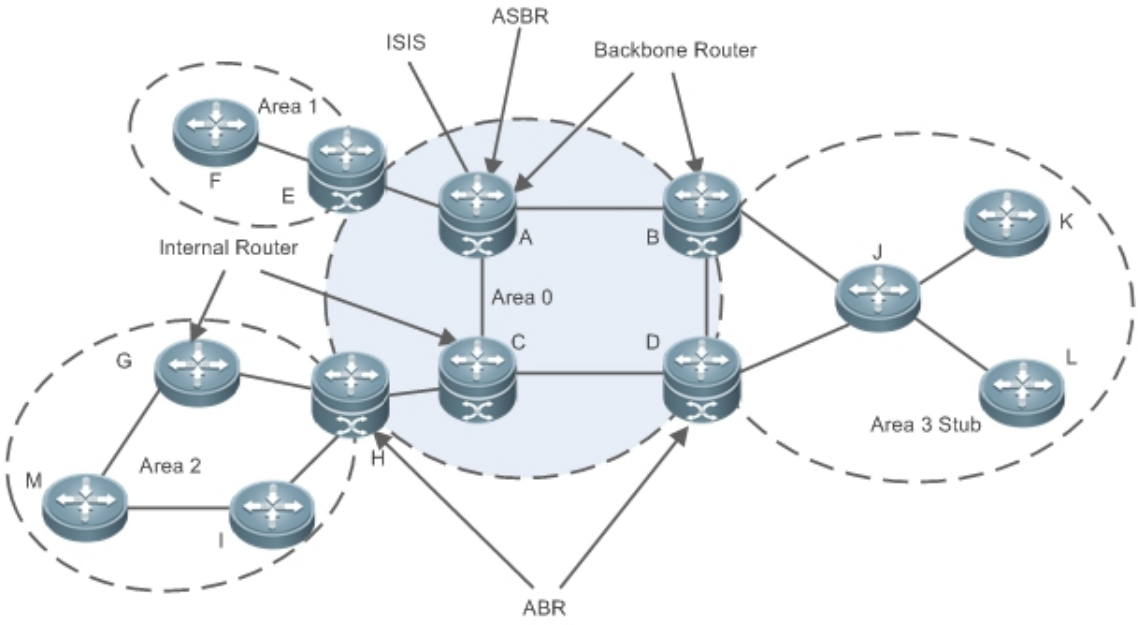
Area

OSPF supports multiple areas. An OSPF domain is divided into multiple areas to ease the computing pressure of a large-scale network.

An area is a logical group of routers, and each group is identified by an area ID. The border between areas is a router. A router may belong to one area or multiple areas. One network segment (link) can belong to only one area, or each OSPF-enabled interface must belong to a specified area.

Area 0 is the backbone area, and other areas are normal areas. Normal areas must be directly connected to the backbone area.

Figure 2- 2 Division of the OSPF Areas



OSPF Router

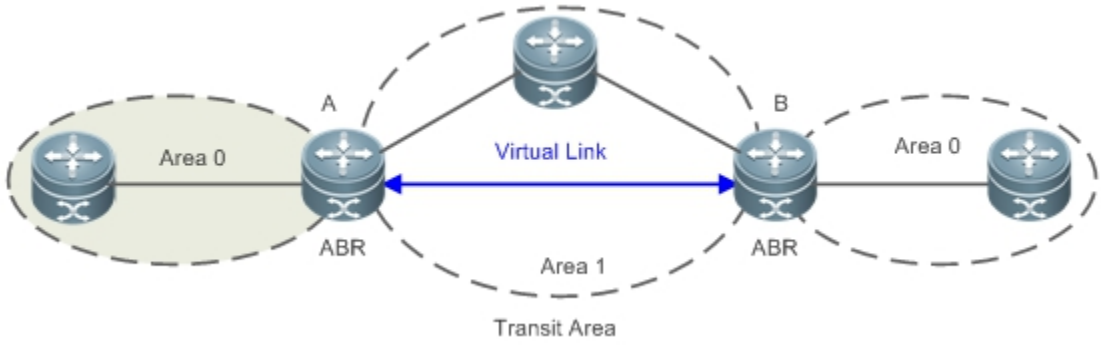
The following types of routers are defined in OSPF, and assigned with different responsibilities:

- Internal router
  All interface of an interval router belong to the same OSPF area. As shown in Figure 2- 2, A, C, F, G, I, M, J, K, and L are internal routers.
- Area border router (ABR)
  An ABR is used to connect the backbone area with a normal area. An ABR belongs to two or more areas, and one of the areas must be the backbone area. As shown in Figure 2- 2, B, D, E, and H are ABRs.
- Backbone router
  A backbone router has at least one interface that belongs to the backbone area. All ABRs and all routers in area 0 are backbone routers. As shown in Figure 2- 2, A, B, C, D, E, and H are backbone routers.
- AS boundary router (ASBR)
  An ASBR is used to exchange routing information with other ASs. An ASBR is not necessarily located on the border of an AS. It may be a router inside an area, or an ABR. As shown in Figure 2- 2, A is an ASBR.

Virtual Link

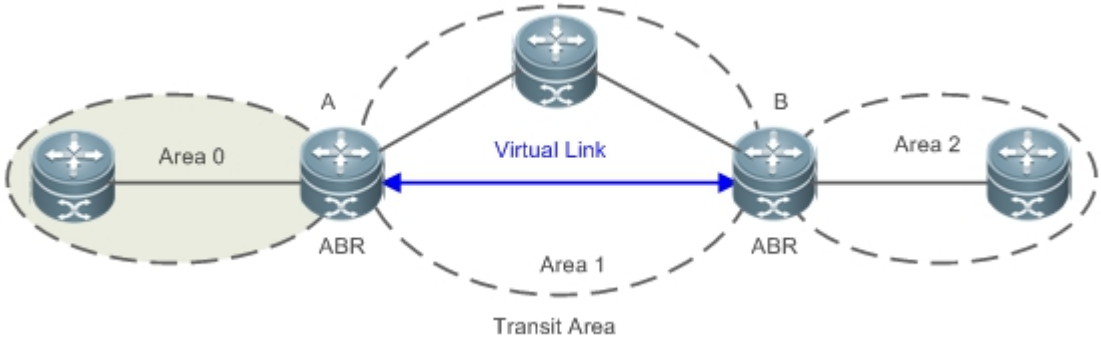
OSPF supports virtual links. A virtual link is a logical link that belongs to the backbone area. It is used to resolve the problems such as a discontinuous backbone area or a failure to directly connect a normal area to the backbone area on the physical network. A virtual link supports traversal of only one normal area, and this area is called transit area. Routers on both ends of a virtual link are ABRs.

Figure 2- 3 Discontinuous Backbone Area on the Physical Network



As shown in Figure 2- 3, a virtual link is set up between A and B to connect two separated area 0s. Area 1 is a transit area, and A and B are ABRs of area 1.

Figure 2- 4 Failure to Directly Connect a Normal Area to the Backbone Areaon the Physical Network



As shown in Figure 2- 4, a virtual link is set up between A and B to extend area 0 to B so that area 0 can be directly connected to area 2 on B. Area 1 is a transit area, A is an ABR of area 1, and B is an ABR of area 0 and area 2.

LSA

OSPF describes the routing information by means of Link State Advertisement (LSA).

LSA Type	Description
Router-LSA(Type 1)	This LSA is originated by every router. It describes the link state and cost of the router, and is advertised only within the area where the originating router is located.
Network-LSA(Type 2)	This LSA is originated by a designated routers (DR) on the NBMA network. It describes the link state in the current network segment, and is advertised only within the area where the DR is located.
Network-summary-LSA(Type 3)	This LSA is originated by an ABR. It describes a route to another area, and is advertised to areas except totally stub areas or Not-So-Stubby Area (NSSA) areas.
ASBR-summary-LSA(Type 4)	This LSA is originated by an ABR. It describes a route to an ASBR, and is advertised to areas except areas where the ASBR is located.
AS-external-LSA(Type 5)	This LSA is originated by an ABR. It describes a route to a destination outside the AS, and is advertised to all areas except the stub and NSSA areas.
NSSA LSA(Type 7)	This LSA is originated by an ABR. It describes a route to a destination outside the AS, and is advertised only within the NASSA areas.
Opaque LSA(Type 9/Type 10/Type 11)	Opaque LSAs provide a generalized mechanism to allow for the future extensibility of OSPF, wherein, ● Type 9 LSAs are only advertised within the network segment where interfaces resides. The

LSA Type	Description
	Grace LSA used to support graceful restart (GR) is one of Type 9 LSAs. ● Type 10 LSAs are advertised within an area. The LSA used to support Traffic Engineering (TE) is one of Type 10 LSAs. ● Type 11 LSAs are advertised within an AS. At present, there are no application examples of Type 11 LSAs.

Stub areas, NSSA areas, totally stub areas, and totally NSSA areas are special forms of normal areas and help reduce the load of routers and enhance reliability of OSPF routes.

OSPF Packet

The following table lists the protocol packets used by OSPF. These OSPF packets are encapsulated in IP packets and transmitted in multicast or unicast mode.

Packet Type	Description
Hello	Hello packets are sent periodically to discover and maintain OSPF neighbor relationships.
Database Description (DD)	DD packets carry brief information about the local Link-State Database (LSDB) and are used to synchronize the LSDBs between OSPF neighbors.
Link State Request (LSR)	LSR packets are used to request the required LSAs from neighbors. LSR packets are sent only after DD packets are exchanged successfully between OSPF neighbors.
Link State Update (LSU)	LSU packets are used to send the required LSAs to peers.
Link State Acknowledgment (LSAck)	LSAck packets are used to acknowledge the received LSAs.

Overview

Feature	Description
Link-State Routing Protocols	Run OSPF on the router to obtain routes to different destinations on the network.
OSPF Route Management	Plan or optimize OSPF routes through manual configuration to implement management of OSPF routes.
Enhanced Security and Reliability	Use functions such as authentication to enhance security, stability, and reliability of OSPF.
Network Management	Use functions such as the management information base (MIB) and Syslog to facilitate OSPF management.

2.3.1 Link-State Routing Protocols

OSPF is a type of link-state routing protocols. Its working process is as follows:

- Neighbor discovery → Bidirectional communication
  An OSPF neighbor relationship is set up between adjacent routers, and bidirectional communication is maintained.
- Database synchronization → Full adjacency
  A router uses LSAs to advertise all its link states. LSAs are exchanged between neighbors and the link state database (LSDB) is synchronized to achieve full adjacency.
- Shortest Path Tree (SPT) computation → Formation of a routing table
  The router computes the shortest path to each destination network based on the LSDB and forms an OSPF routing table.

Working Principle

Neighbor Discovery → Bidirectional Communication

Routers send Hello packets through all OSPF-enabled interfaces (or virtual links). If Hello packets can be exchanged between two routers, and parameters carried in the Hello packets can be successfully negotiated, the two routers become neighbors. Routers that are mutually neighbors find their own router IDs from Hello packets sent from neighbors, and bidirectional communication is set up.

A Hello packet includes, but is not limited to, the following information:

- Router ID of the originating router
- Area ID of the originating router interface (or virtual link)
- Subnet mask of the originating router interface (or virtual link)
- Authentication information of the originating router interface (or virtual link)
- Hello interval of the originating router interface (or virtual link)
- Neighbor dead interval of the originating router interface (or virtual link)
- Priority of the originating router interface (used for DR/BDR election)
- IP addresses of the DR and Backup Designated Router (BDR)
- Router ID of the neighbor of the originating router

Database Synchronization → Full Adjacency

After bidirectional communication is set up between neighbor routers, the DD, LSR, LSU, and LSAck packets are used to exchange LSAs and set up the adjacency. The brief process is as follows:

- A router generates an LSA to describe all link states on the router.
- The LSA is exchanged between neighbors. When a router receives the LSA from its neighbor, it copies the LSA and saves the copy in the local LSDB, and then advertises the LSA to other neighbors.
- When the router and its neighbors obtain the same LSDB, full adjacency is achieved.

OSPF will be very quiet without changes in link costs or network addition or deletion. If any change takes place, the changed link states are advertised to quickly synchronize the LSDB.

SPT Computation → Formation of a Routing Table

After the complete LSDB is obtained from the router, the Dijkstra algorithm is run to generate an SPT from the local router to each destination network. The SPT records the destination networks, next-hop addresses, and costs. OSPF generates a routing table based on the SPT.

If changes in link costs or network addition or deletion take place, the LSDB will be updated. The router again runs the Dijkstra algorithm, generates a new SPT, and updates the routing table.

The Dijkstra algorithm is used to find a shortest path from a vertex to other vertices in a weighted directed graph.

OSPF Network Types

A router does not necessarily need to exchange LSAs with every neighbor and set upan adjacency with every neighbor. To improve efficiency, OSPF classifies networks that use various link layer protocols into five types so that LSAs are exchanged in different ways to set upan adjacency:

- Broadcast
  Neighbors are discovered, and the DR and BDR are elected.
  The DR (or BDR) exchanges LSAs with all other routers to set up an adjacency. Except the DR and BDR, all other routers do not

exchange LSAs with each other, and the adjacency is not set up.

Ethernet and fiber distributed data interface (FDDI) belong to the broadcast network type by default.

- Non-broadcast multiple access (NBMA)

  Neighbors are manually configured, and the DR and BDR are elected.

  The DR (or BDR) exchanges LSAs with all other routers to set up an adjacency. Except the DR and BDR, all other routers do not exchange LSAs with each other, and the adjacency is not set up.

  X.25, frame relay, and ATM belong to NBMA networks by default.

- Point-to-point (P2P)

  Neighbors are automatically discovered, and the DR or BDR is not elected.

  LSAs are exchanged between routers at both ends of the link, and the adjacency is set up.

  PPP,HDLC, and LAPB belongs to the P2P network type by default.

- Point-to-multipoint (P2MP)

  Neighbors are automatically discovered, and the DR or BDR is not elected.

  LSAs are exchanged between any two routers, and the adjacency is set up.

  Networks without any link layer protocol belong to the P2MP network type by default. P2MP broadcast

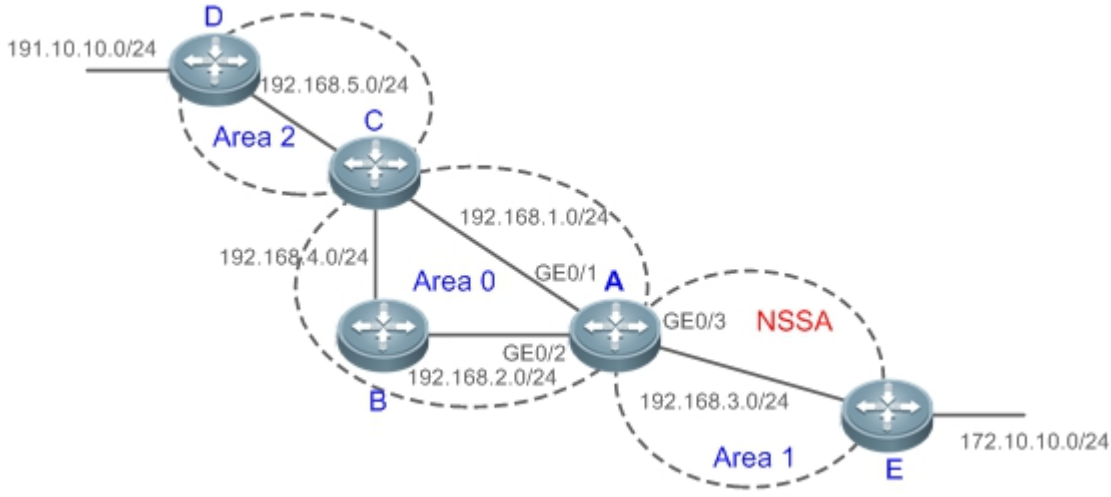
  Neighbors are manually configured, and the DR or BDR is not elected.

  LSAs are exchanged between any two routers, and the adjacency is set up.

  Networks without any link layer protocol belong to the P2MP network type by default.

OSPF Route Types

Figure 2- 5



Display the OSPF routes (marked in red) in the routing table of Router A.

```
A#show ip route


Codes:    C - connected, S - static, R - RIP, B - BGP

O - OSPF, IA - OSPF inter area

N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2

E1 - OSPF external type 1, E2 - OSPF external type 2
```

```
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
O N2 172.10.10.0/24 [110/20] via 192.168.3.2, 00:01:00,GigabitEthernet 0/3
O E2 191.10.10.0/24 [110/20] via 192.168.1.2, 01:11:26,GigabitEthernet 0/1
C       192.168.1.0/24 is directly connected,GigabitEthernet 0/1
C       192.168.1.1/32 is local host.
C       192.168.2.0/24 is directly connected,GigabitEthernet 0/2
C       192.168.2.1/32 is local host.
C       192.168.3.0/24 is directly connected,GigabitEthernet 0/3
C       192.168.3.1/32 is local host.
O       192.168.4.0/24 [110/2] via 192.168.2.2, 00:00:02,GigabitEthernet 0/2
O IA 192.168.5.0/24 [110/3] via 192.168.1.2, 00:01:02,GigabitEthernet 0/1
```

A mark is displayed in front of each OSPF route to indicate the type of the route. There are six types of OSPF routes:

- O: Intra-area route
  This type of route describes how to arrive ata destination network in the local area. The cost of this type of route is equal to the cost of the route from the local router to the destination network.
- IA: Inter-area route
  This type of route describes how to arrive at a destination network in another area. The cost of this type of route is equal to the cost of the route from the local router to the destination network.
- E1: Type 1 external route
  This type of route describes how to arrive at a destination network outside the AS. The cost of this type of route is equal to the cost of the route from the local router to the ASBR plus the cost of the route from the ASBR to the destination network. This type of route does not exist on routers in the stub or NSSA area.
- E2: Type 2 external route
  This type of route describes how to arrive at a destination network outside the AS. The cost of this type of route is equal to the cost of the route from the ASBR to the destination network. This type of route does not exist on routers in the stub or NSSA area.
- N1: Type 1 external route of the NSSA area
  This type of route describes how to arrive at a destination network outside the AS through the ASBR in the NSSA area. The cost of this type of route is equal to the cost of the route from the local router to the ASBR plus the cost of the route from the ASBR to the destination network. This type of route exists only on routers in the NSSA area.
- N2: Type 2 external route of the NSSA area
  This type of route describes how to arrive at a destination network outside the AS through the ASBR in the NSSA area. The cost of this type of route is equal to the cost of the route from the ASBR to the destination network. This type of route exists only on routers in the NSSA area.

Reliability of E2 and N2 routes is poor. OSPF believes that the cost of the route from the ASBR to a destination outside an AS is far greater than the cost of the route to the ASBR within the AS. Therefore, when the route cost is computed, only the cost of the route from the ASBR to a destination outside an AS is considered.

Related Configuration

Enabling OSPF

OSPF is disabled by default.

Run the **router ospf** 1 command to create an OSPF process on the router.

Run the **network area** command to enable OSPF on the interface and specify the area ID.

Run the **area virtual-link** command to create a virtual link on the router. The virtual link can be treated as a logical interface.

Router ID

By default, the OSPF process elects the largest IP address among the IP addresses of all the loopback interfaces as the router ID. If the loopback interfaces configured with IP addresses are not available, the OSPF process elects the largest IP address among the IP addresses of all the loopback interfaces as the router ID.

Alternatively, you can run the **router-id** command to manually specify the router ID.

Protocol Control Parameters

Run the **ip ospf hello-interval** command to modify the Hello interval on the interface. The default value is 10s (or 30s for NBMA networks).

Run the **ip ospf dead-interval** command to modify the neighbor dead interval on the interface. The default value is four times the Hello interval.

Use the **poll-interval** parameter in the **neighbor** command to modify the neighbor polling interval on the NBMA interface. The default value is 120s.

Run the **ip ospf transmit-delay** command to modify the LSU packet transmission delay on the interface. The default value is 1s.

Run the **ip ospf retransmit-interval** command to modify the LSU packet retransmission interval on the interface. The default value is 5s.

Use the **hello-interval** parameter in the **area virtual-link**command to modify the Hello interval on the virtual link. The default value is 10s.

Use the **dead-interval** parameter in the **area virtual-link**command to modify the neighbor dead interval on the virtual link. The default value is four times the Hello interval.

Use the **transmit-delay** parameter in the **area virtual-link**command to modify the LSU packet transmission delay on the virtual link. The default value is 1s.

Use the **retransmit-interval** parameter in the **area virtual-link**command to modify the LSU packet retransmission interval on the virtual link. The default value is 5s.

Run the **timers throttle lsa all** command to modify parameters of the exponential backoff algorithm that generates LSAs. The default values of these parameters are 0 ms, 5000 ms, and 5000 ms.

Run the **timerspacinglsa-group** command to modify the LSA group update interval. The default value is 30s.

Run the **timers pacing lsa-transmit** command to modify the LS-UPD packet sending interval and the number of sent LS-UPD packets. The default values are 40 ms and 1.

Run the **timers lsa arrival** command to modify the delay after which the same LSA is received. The default value is 1000 ms.

Run the **timers throttle spf** command to modify the SPT computation delay, minimum interval between two SPT computations, and maximum interval between two SPT computations. The default values are 1000 ms, 5000 ms, and 10000 ms.

OSPF Network Types

By default, Ethernet and FDDI belong to the broadcast type, X.25, frame relay, and ATM belong to the NBMA type, and PPP, HDLC, and LAPB belong to the P2P type.

Run the **ip ospf network** command to manually specify the network type of an interface.

Run the **neighbor** command to manually specify a neighbor. For the NBMA and P2MP non-broadcast types, you must manually specify neighbors.

Run the **ip ospf priority** command to adjust the priorities of interfaces, which are used for DR/BDR election. The DR/BDR election is required for the broadcast and NBMA types. The router with the highest priority wins in the election, and the router with the priority of 0 does not participate in the election. The default value is 1.

2.3.2 OSPF Route Management

Plan or optimize OSPF routes through manual configuration to implement management of OSPF routes.

Working Principle

(Totally) Stub Area and (Totally)NSSA Area

The (totally) stub and (totally)NSSA areas help reduce the protocol interaction load and the size of the routing table.

- If an appropriate area is configured as a (totally) stub or NSSA area, advertisement of a large number of Type 5 and Type 3 LSAs can be avoided within the area.

Area	Type1 and Type2 LSAs	Type 3 LSA	Type 4 LSA	Type 5 LSA	Type 7 LSA
Non (totally) stub area and NSSA area	Allowed	Allowed	Allowed	Allowed	Not allowed
Stub area	Allowed	Allowed (containing one default route)	Not allowed	Not allowed	Not allowed
Totally stub area	Allowed	Only one default route is allowed.	Not allowed	Not allowed	Not allowed
NSSA area	Allowed	Allowed (containing one default route)	Allowed	Not allowed	Allowed
Totally NSSA area	Allowed	Only one default route is allowed.	Allowed	Not allowed	Allowed

The ABR uses Type 3LSAs to advertise a default route to the (totally) stub or NSSA area.

The ABR converts Type 7 LSAs in the totally NSSA area to Type 5LSAs, and advertise Type5LSAs to the backbone area.

- If an area is appropriately configured as a (totally) stub area or an NSSA area, a large number of E1, E2, and IA routes will not be added to the routing table of a router in the area.

Area	Routes Available in the Routing Table of a Router Inside the Area
Non (totally) stub area and NSSA	O: a route to a destination network in the local area

area	IA: a route to a destination network in another area E1 or E2: a route or default route to a destination network segment outside the AS (via any ASBR in the AS)
Stub area	O: a route to a destination network in the local area IA: a route or a default route to a destination network in another area
Totally stub area	O: a route to a destination network in the local area IA: a default route
NSSA area	O: a route to a destination network in the local area IA: a route or a default route to a destination network in another area N1 or N2: a route or default route to a destination network segment outside the AS (via any ASBR in the local area)
Totally NSSA area	O: a route to a destination network in the local area IA: a default route N1 or N2: a route or default route to a destination network segment outside the AS (via any ASBR in the local area)

Route Redistribution

Route redistribution refers to the process of introducing routes of other routing protocols, routes of other OSPF processes, static routes, and direct routes that exist on the device to an OSPF process so that these routes can be advertised to neighbors using Type 5 and Type 7 LSAs. A default route cannot be introduced during route redistribution.

Route redistribution is often used for interworking between ASs. You can configure route redistribution on an ASBR to advertise routes outside an AS to the interior of the AS, or routes inside an AS to the exterior of the AS.

Default Route Introduction

By configuring a command on an ASBR, you can introduce a default route to an OSPF process so that the route can be advertised to neighbors using Type 5 and Type 7 LSAs.

Default route introduction is often used for interworking between ASs. One default route is used to replace all the routes outside an AS.

Route Summarization

Route summarization is a process of summarizing routing information with the same prefix into one route, and advertising the summarized route (replacing a large number of individual routes) to neighbors. Route summarization helps reduce the protocol interaction load and the size of the routing table.

By default, the ABR advertises inter-area routing information by using Type3 LSAs within a network segment, and advertises redistributed routing information by using Type 5 and Type 7 LSAs.If continuous network segments exist, it is recommended that you configure route summarization.

When configuring route summarization, the summarization range may exceed the actual network scope of routes. If data is sent to a network beyond the summarization range, a routing loop may be formed and the router processing load may increase.To prevent these problems, the ABR or ASBR automatically adds a discard route to the routing table. This route will not be advertised.

Route Filtering

OSPF supports route filtering to ensure security and facilitate control when the routing information is being learned, exchanged, or used.

Using configuration commands, you can configure route filtering for the following items:

- Interface: The interface is prevented from sending routing information (any LSAs) or exchanging routing information (any LSAs) with neighbors.
- Routing information advertised between areas: Only the routing information that meets the filtering conditions can be advertised to another area (Type 3 LSAs).
- Routing information outside an AS: Only the routing information that meets the filtering conditions can be redistributed to the OSPF process(Type 5 and Type 7 LSAs).
- LSAs received by a router: In the OSPF routing table, only the routes that are computed based on the LSAs meeting the filtering conditions can be advertised.

Route Cost

If redundancy links or devices exist on the network, multiple paths may exist from the local device to the destination network. OSPF selects the path with the minimum total cost to form an OSPF route. The total cost of a path is equal to the sum of the costs of individual links along the path.The total cost of a path can be minimized by modifying the costs of individual links along the path. In this way, OSPF selects this path to form a route.

Using configuration commands, you can modify the link costs:

- Cost from an interface to a directly connected network segment and cost from the interface to a neighbor
- Costfrom an ABR to the inter-area summarization network segment and cost from the ABR to the default network segment
- Costfrom an ASBR to an external network segment and cost from the ASBR to the default network segment

Both the cost and the metric indicate the cost and are not differentiated from each other.

OSPF Administrative Distance

The administrative distance (AD) evaluates reliability of a route, and the value is an integer ranging from 0 to 255. A smaller AD value indicates that the route is more trustworthy. If multiples exist to the same destination, the route preferentially selects a route with a smaller AD value. The route with a greater AD value becomes a floating route, that is, a standby route of the optimum route.

By default, the route coming from one source corresponds to an AD value. The AD value is a local concept. Modifying the AD value affects route selection only on the current router.

Route Source	Directly-Connected Network	Static Route	OSPF Route	RIP Route	Unreachable Route
Default AD	0	1	110	120	255

Related Configuration

Stub Area and NSSA Area

No stub or NSSA area is configured by default.

Run the **area stub** command to configure a specified area as a stub area.

Run the **area nssa** command to configure a specified area as an NSSA area.

The backbone area cannot be configured as a stub or an NSSA area.

A transit area (with virtual links going through) cannot be configured as a stub or an NSSA area.

An area containing an ASBR cannot be configured as a stub area.

Route Redistribution and Default Route Introduction

By default, routes are not redistributed and the default route is not introduced.

Run the **redistribute** command to configure route redistribution.

Run the **default-information originate** command to introduce the default route.

After configuring route redistribution and default route introduction, the route automatically becomes an ASBR.

Route Summarization

By default, routes are not summarized. If route summarization is configured, a discard route will be automatically added.

Run the **arearange** command to summarize routes distributed between areas (Type 3 LSA) on the ABR.

Run the **summary-address** command to summarize redistributed routes (Type 5 and Type 7 LSAs) on the ASBR.

Run the **discard-route** command to add a discard route to the routing table.

Route Filtering

By default, routes are not filtered.

Run the **passive-interface** command to configure a passive interface. Routing information (any LSAs) cannot be exchanged on a passive interface.

Run the **ip ospfdatabase-filter all out** command to prohibit an interface from sending routing information (any LSAs).

Run the **area filter-list** command to filter routing information advertised between areas on the ABR. Only the routing information that meets the filtering conditions can be advertised to another area (Type 3 LSAs).

Use the **route-map** parameter in the **redistribute** command, or use the **distribute-list out** command to filter the external routing information of the AS on the ASBR. Only the routing information that meets the filtering conditions can be redistributed to the OSPF process (Type 5 and Type 7 LSAs).

Run the **distribute-list in** command to filter LSAs received by the router. In the OSPF routing table, only the routes that are computed based on the LSAs meeting the filtering conditions can be advertised.

Route Cost

- Cost from the interface to the directly-connected network segment (cost on the interface)
  The default value is the auto cost. Auto cost = Reference bandwidth/Interface bandwidth
  Run the **auto-costreference-bandwidth** command to set the reference bandwidth of auto cost. The default value is 100 Mbps.
  Run the **ip ospf cost** command to manually set the cost of the interface. The configuration priority of this item is higher than that of the auto cost.
- Cost from the interface to a specified neighbor (that is, cost from the local device to a specified neighbor)
  The default value is the auto cost.
  Use the **cost** parameter in the **neighbor** command to modify the cost from the interface to a specified neighbor. The configuration priority of this item is higher than that of the cost of the interface.
  This configuration item is applicable only to P2MP-type interfaces.
- Cost from the ABR to the inter-area summarization network segment (that is, the cost of the summarized inter-area route)
  If OSPF routing is compatible with RFC1583, the default value is the minimum cost among all costs of the summarized links; otherwise, the default value is the maximum cost among all costs of the summarized links.
  Run the **compatible rfc1583** command to make OSPF routing compatible with RFC1583. By default, OSPF routing is compatible

with RFC1583.

Use the **cost** parameter in the **area range** command to modify the cost of inter-area route summarization.

- Cost from the ABR to the default network segment (that is, the cost of the default route that is automatically advertised by the ABR to the stub or NSSA areas)

  The default value is 1.

  Run the **area default-cost** command to modify the cost of the default route that the ABR automatically advertise to the stub or NSSA areas.

- Cost from the ASBR to an external network segment (that is, the metric of an external route)

  By default, the metric of other types of redistributed routes is 20, and the route type is Type 2 External.

  Run the **default-metric** command to modify the default metric of the external route.

  Use the **metric**,**metric-type** and **route-map** parameters in the **redistribute** command to modify the metric and route type of the external route.

- Cost from the ASBR to the default network segment (that is, the metric of the default route that is manually introduced)

  By default, the metric is 1, and the route type is Type 2 External.

  Use the **metric**,**metric-type** and **route-map** parameters in the **default-information originate** command to modify the metric and route type of the default route that is manually introduced.

  Use the **metric** and **metric-type** parametersof**default-information originate**in the **area nssa** command to modify the metric and type of the default route that is manually introduced to the NSSA area.

- Run the **max-metric router-lsa** command to set metrics of all routes advertised on the router to the maximum value. In this way, the total cost of any path that passes through this router will become very large, and the path can hardly become the shortest path.

OSPF Administrative Distance

By default, the OSPF AD is 110.

Run the **distance** command to set the AD of an OSPF route.

2.3.3 Enhanced Security and Reliability

Use functions such as authentication to enhance security, stability, and reliability of OSPF.

Working Principle

Authentication

Authentication prevents routers that illegally access the network and hosts that forge OSPF packet from participating in the OSPF process. OSPF packets received on the OSPF interface (or at both ends of the virtual link) are authenticated. If authentication fails, the packets are discarded and the adjacency cannot be set up.

Enabling authentication can avoid learning unauthenticated or invalid routes, thus preventing advertising valid routes to unauthenticated devices. In the broadcast-type network, authentication also prevents unauthenticated devices from becoming designated devices, ensuring stability of the routing system and protecting the routing system against intrusions.

MTU Verification

On receiving a DD packet, OSPF checks whether the MTU of the neighbor interface is the same as the MTU of the local interface. If the MTU of the interface specified in the received DD packet is greater than the MTU of the interface that receives the packet, the adjacency cannot be set up. Disabling MTU verification can avoid this problem.

Source Address Verification

Generally, the source address of a packet received by OSPF is in the same network segment as the receiving interface. The addresses at both ends of a P2P link are configured separately and are not necessarily in the same network segment. In this scenario, as the peer address information will be notified during the P2P link negotiation process, OSPF checks whether the source address of the packet is the address advertised by the peer during negotiation. If not, OSPF determines that the packet is invalid and discards this packet. In particular, OSPF does not verify the address of an unnumbered interface.

In some scenarios, the source address of a packet received by OSPF maynot be in the same network segment as the receiving interface, and therefore OSPF address verification fails. For example, the negotiated peer address cannot be obtained on a P2P link. In this scenario, source address verification must be disabled to ensure that the OSPF adjacency can be properly set up.

Two-Way Maintenance

OSPF routers periodically send Hello packets to each other to maintain the adjacency. On a large network, a lot of packets may be sent or received, occupying too much CPU and memory. As a result, some packets are delayed or discarded. If the processing time of Hello packets exceeds the dead interval, the adjacency will be destroyed.

If the two-way maintenance function is enabled, in addition to the Hello packets, the DD, LSU, LSR, and LSAck packets can also be used to maintain the bidirectional communication between neighbors, which makes the adjacency more stable.

Concurrent Neighbor Interaction Restriction

When a router simultaneously exchanges data with multiple neighbors, its performance may be affected. If the maximum number of neighbors that concurrently initiate or accept interaction with the OSPF process, the router can interact with neighbors by batches, which ensures data forwarding and other key services.

Overflow

OSPF requires that routers in the same area store the same LSDB. The number of routers keeps increasing on the network. Some routers, however, cannot store so much routing information due to the limited system resources. The large amount of routing information may exhaust the system resources of routers, causing failures of the routers.

The overflow function limit the number of external routes in the LSDB to control the size of the LSDB.

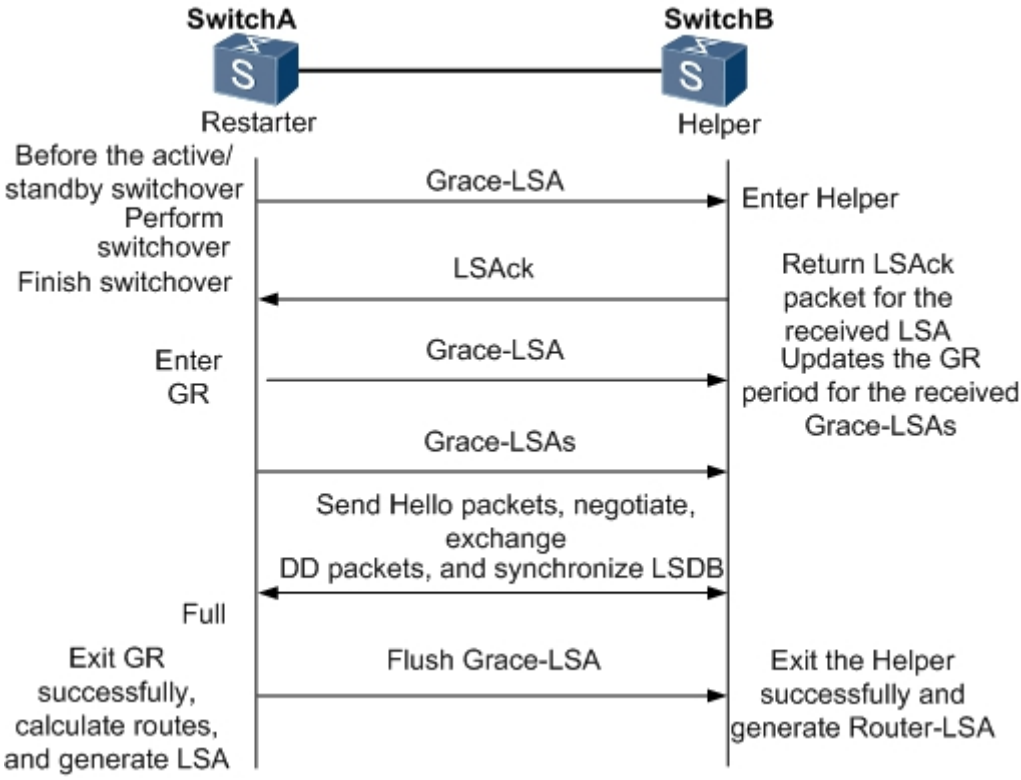
When the number of external routes on a router exceeds the upper limit, the router enters the overflow state. The router deletes the external routes generated by itself from the LSDB, and does not generate new external routes. In addition, the router discards the newly received external routes. After the overflow state timer (5s) expires, if the number of external routes is lower than the upper limit, the normal state is restored.

GR

The control and forwarding separated technology is widely used among routers. On a relatively stable network topology, when a GR-enabled router is restarted on the control plane, data forwarding can continue on the forwarding plane. In addition, actions (such as adjacency re-forming and route computation) performed on the control plane do not affect functions of the forwarding plane. In this way, service interruption caused by route flapping can be avoided, thus enhancing reliability of the entire network.

Currently, the GR function is used only during active/standby switchover and system upgrade.

Figure 2- 6 Normal OSPF GR Process



- The GR process requires collaboration between the restarter and the helper. The restarter is the router where GR occurs. The helper is a neighbor of the restarter.
- When entering or exiting the GR process, the restarter sends a Grace-LSA to the neighbor, notifying the neighbor to enter or exit the helper state.
- When the adjacency between the restarter and the helper reaches the Full state, the router can exit the GR process successfully.

Fast Hello

After a link fault occurs, OSPF senses the death of the neighbor only after a period of time (about 40s). Then, OSPF advertises the information and re-computes the SPT. During this period, traffic is interrupted.

- After the fast Hello function is enabled (that is, the neighbor dead interval is set to 1s), OSPF can sense the death of a neighbor within 1s once a link is faulty. This greatly accelerates route convergence and prevents traffic interruption.

Related Configuration

OSPF Packet Authentication

By default, authentication is disabled.

- Run the **areaauthentication** command to enable the authentication function in the entire area so that the function takes effect on all interfaces in this area. If authentication is enabled in area 0, the function takes effect on the virtual link.
- Run the **ip ospf authentication** command to enable authentication on an interface. This configuration takes precedence over the area-based configuration.
- Run the **ip ospf authentication-key** command to set the text authentication key on an interface.

- Run the **ip ospfmessage-digest-key** command to set the message digest 5 (MD5) authentication key on an interface.
- Use the **authentication** parameter in the **area virtual-link** command to enable authentication at both ends of a virtual link. This configuration takes precedence over the area-based configuration.
- Use the **authentication-key** parameter in the **area virtual-link** command to set the text authentication key at both ends of a virtual link.
- Use the **message-digest-key** parameter in the **area virtual-link** command to set the MD5 authentication key at both ends of a virtual link.

MTU Verification

By default, MTU verification is disabled.

Run the **ip ospf mtu-ignore** command to disable MTU verification on an interface.

Source address verification

By default, source address verification is enabled on a P2P interface.

Run the **ip ospf source-check-ignore** command to disable source address verification on an interface.

Two-Way Maintenance

By default, bidirectional maintenance is enabled.

Run the **two-way-maintain** command to enable two-way maintenance.

Concurrent neighbor Interaction Restriction

Run the **max-concurrent-dd** command to modify the maximum number of neighbors that are concurrently interacting with the current OSPF process. The default value is 5.

Run the **ip router ospf max-concurrent-dd** command to modify the maximum number of neighbors that are concurrently interacting with all OSPF processes on the router. The default value is 10.

Overflow

Run the **overflow memory-lack** command to allow the router to enter the overflow state when the memory is insufficient. By default, the router is allowed to enter the overflow state when the memory is insufficient.

Run the **overflow database** command to allow the router to enter the overflow state when the number of LSAs is too large. By default, the router is not allowed to enter the overflow state when the number of LSAs is too large.

Run the **overflow database external** command to allow the router to enter the overflow state when the number of externalLSAs is too large. By default, the router is not allowed to enter the overflow state when the number of external-LSAs is too large.

GR

By default, the restarter function is disable, and the helper function is enabled.

Run the **graceful-restart** command to configure the restarter function.

Run the **graceful-restart helper** command to configure the helper function.

Fast Hello

By default, the neighbor dead interval on the interface is 40s.

Run the **ip ospf dead-intervalminimal hello-multiplier** command to enable the Fast Hello function on an interface, that is, the neighbor dead interval is 1s.

2.3.4 Network Management

Use functions such as the MIB and Syslog to facilitate OSPF management.

Working Principle

MIB

MIB is the device status information set maintained by a device. You can use the management program to view and set the MIB node.

Multiple OSPF processes can be simultaneously started on a router, but the OSPF MIB can be bound with only one OSPF process.

Trap

A Trap message is a notification generated when the system detects a fault. This message contains the related fault information.

If the Trap function is enabled, the router can proactively send the Trap messages to the network management device.

Syslog

The Syslog records the operations (such as command configuration) performed by users on routers and specific events (such as network connection failures).

If the Syslog is allowed to record the adjacency changes, the network administrator can view the logs to learn the entire process that the OSPF adjacency is set up and maintained.

Related Configuration

MIB

By default, the MIB is bound with the OSPF process with the smallest process ID.

Run the **enable mib-binding** command to bind the MIB with the current OSPF process.

Trap

By default, all traps are disabled, and the device is not allowed to send OSPF traps.

Run the **enable traps** command to enable a specified trap for an OSPF process.

Run the **snmp-server enable traps ospf** command to allow the device to send OSPF traps.

SYSLOG

By default, the Syslog is allowed to record the adjacency changes.

Run the **log-adj-changes** command to allow the Syslog to record the adjacency changes.

2.4 Configuration

Configuration	Description and Command	
Configuring OSPF Basic Functions	(Mandatory) It is used to build an OSPF routing domain.	
	routerospf	Creates an OSPF process.

Configuration	Description and Command	
	router-id	Configures a router ID.
	network area	Enables OSPF on an interface and specifies an area ID.
	area virtual-link	Creates a virtual link.
Setting the Network Type	(Optional) The configurations are mandatory if the physical network is the X.25, frame relay, or ATM network.	
	ip ospf network	Defines the network type.
	neighbor	Specifies a neighbor.
	ip ospf priority	Configures the DR priority.
Configuring Route Redistribution and Default Route	(Optional) The configurations are recommended if the OSPF routing domain is connected with an external network.	
	redistribute	Configures route redistribution.
	default-information originate	Introduces a default route.
Configuring Stub Area and NSSA Area	(Optional) It is used to reduce interaction of routing information and the size of routing table, and enhance stability of routes.	
	areastub	Configures a stub area.
	areanssa	Configures an NSSA area.
Configuring Route Summarization	(Optional) It is used to reduce interaction of routing information and the size of routing table, and enhance stability of routes.	
	arearange	Summarizes routes that are advertised between areas.
	summary-address	Summarizes routes that are introduced through redistribution.
	discard-route	Adds a discard route to the routing table.
Configuring Route Summarization	(Optional) It is used to manually control interaction of routing information and filter available OSPF routes.	
	passive-interface	Configures a passive interface.
	ip ospfdatabase-filter all out	Prohibits an interface from sending LSAs.
	area filter-list	Filters routes that are advertised between areas..
	distribute-list out	Filters routes that are introduced through redistribution.
	distribute-listin	Filters routes that are calculated based on the received LSAs.
Configuring Route Filtering	(Optional) It is used to manually control the shortest route computed by OSPF and determine whether to select an OSPF route preferentially.	
	auto-costreference-bandwidth	Modifies the reference bandwidth of the auto cost.

Configuration	Description and Command	
	ip ospf cost	Modifies the cost in the outbound direction of an interface.
	areadefault-cost	Modifies the cost of the default route in a stub or an NSSA area.
	default-metric	Modifies the default metric of a redistributed route.
	max-metric router-lsa	Configures the maximum metric.
	compatible rfc1583	Enables the routing rules to be compatible with RFC1583.
	distance	Modifies the OSPF AD.
Modifying Route Cost and AD	(Optional) It is used to prevent routers that illegally access the network and hosts that forge OSPF packets from participating in the OSPF protocol process.	
	areaauthentication	Enables authentication and sets the authentication mode in an area.
	ip ospf authentication	Enables authentication and sets the authentication mode on an interface.
	ip ospf authentication-key	Sets the text authentication key on an interface.
	ip ospfmessage-digest-keymd5	Sets the MD5 authentication key on an interface.
Enabling Authentication	(Optional) It is used to prevent the problem that OSPF processes stop running due to over-consumption of the memory.	
	overflow memory-lack	Allows the router to enter the overflow state when the memory is insufficient.
	overflow database	Allows the router to enter the overflow state when the number of LSAs exceeds the preset limit.
	overflow database external	Allows the router to enter the overflow state when the number of external LSAs exceeds the preset limit.
Enabling Overflow	(Optional) It is used to prevent the problem of performance deterioration caused by over-consumption of the CPU.	
	max-concurrent-dd	Modifies the maximum number of con current neighbors on the current OSPF process.
	router ospf max-concurrent-dd	Modifies the maximum number of con current neighbors on all OSPF processes.
Modifying the Maximum Number of Concurrent Neighbors	(Optional) It is used to prevent the problem that the adjacency cannot be set up due to the failure to obtain the peer address.	
	ip ospf source-check-ignore	Disables source address verification on an interface.

Configuration	Description and Command	
Disabling Source Address Verification	(Optional) It is used to prevent the problem that the adjacency cannot be set up due to MTU inconsistency on the neighbor interface.	
	ip ospf mtu-ignore	Disables MTU verification on an interface.
Disabling MTU Verification	(Optional) It is used to prevent termination of the adjacencydue to the delay or loss of Hello packets.	
	two-way-maintain	Enables two-way maintenance.
Enabling Two-Way Maintenance	(Optional) It is used to retain OSPF routing forwarding during restart or active/standby switchover of the OSPF processes to prevent traffic interruption.	
	graceful-restart	Configures the restarter function.
	graceful-restart helper	Configures the helper function.
Enabling GR	(Optional) It is used to retain OSPF routing forwarding during active/standby switchover of the OSPF processes to prevent traffic interruption.	
	nsr	Enables NSR.
Enabling Fast Hello	(Optional) It is used to quickly discover the death of a neighbor to prevent traffic interruption when a link is faulty.	
	ip ospf dead-intervalminimal hello-multiplier	Enabling Fast Hello.
Configuring the Network Management Function	(Optional) The configurations enable users to use the SNMP network management software to manage OSPF.	
	enable mib-binding	Binds the MIB with the current OSPF process.
	enable traps	Enables a specified trap for an OSPF process.
	snmp-server enable traps ospf	Allows the device to send OSPF traps.
	log-adj-changes	Allows the Syslog to record the adjacency changes.
Modifying Protocol Control Parameters	(Optional) You are advised not to modify protocol control parameters unless necessary.	
	ip ospf hello-interval	Modifies the Hello interval.
	ip ospf dead-interval	Modifies the neighbor death interval.
	timers throttle lsa all	Modifies parameters of the exponential backoff algorithm that generates LSAs.
	timers throttle route inter-area	Modifies the inter-area route computation delay.
	timers throttle route ase	Modifies the external route computation delay.
	timerspacinglsa-group	Modifies the LSA group update interval.
	timers pacing lsa-transmit	Modifies the LS-UPD packet sending interval.
	ip ospf transmit-delay	Modifies the LSU packet transmission delay.
	ip ospf retransmit-interval	Modifies the LSU packet retransmission interval.
	timers lsa arrival	Modifies the delay after which the same LSA is received.
	timers throttlespf	Modifies the SPT computation timer.

2.4.1 Configuring OSPF Basic Functions

Configuration Effect

- Set up an OSPF routing domain on the network to provide IPv4 unicast routing service for users on the network.

Notes

- Ensure that the IP unitcast routing function is enabled, that is, **ip routing** is not disabled; otherwise, OSPF cannot be enabled.
- It is strongly recommended that you manually configure the router ID.
- After **ip ospf disable all** is configured, the interface neither sends or receives any OSPF packet, nor participates in OSPF computation even if the interface belongs to the network.

Configuration Steps

Creating an OSPF Process

- Mandatory.
- The configuration is mandatory for every router.

Configuring a Router ID

- (Optional) It is strongly recommended that you manually configure the router ID.
- If the router ID is not configured, OSPF selects an interface IP address. If the IP address is not configured for any interface, or the configured IP addresses have been used by other OSPF instances, you must manually configure the router ID.

Enabling OSPF on an Interface and Specifying an Area ID

- Mandatory.
- The configuration is mandatory for every router.

Verification

- Run the **show ip route ospf** command to verify that the entries of the OSPF routing table are correctly loaded.
- Run the **ping** command to verify that the IPv4 unicast service is correctly configured.

Related Commands

Creating an OSPF Process

Command	**router ospf** *process-id*
Parameter Description	*process-id:* Indicates the OSPF process ID. If the process ID is not specified, the process ID is 1.
Command Mode	Global configuration mode
Usage Guide	Different OSPF processes are independent of each other, and can be treated as different routing protocols that run independently.

Configuring a Router ID

Command	**router-id***router-id*
Parameter Description	*router-id:* Indicates the router ID to be configured. It is expressed in the IP address.
Command Mode	OSPF routing process configuration mode
Usage Guide	Different OSPF processes are independent of each other, and can be treated as different routing protocols that run independently. Each OSPF process uses a unique router ID.

Enabling OSPF on an Interface and Specifying an Area ID

Command	**network***ip-addresswildcard***area***area-id*
Parameter Description	*ip-address*: Indicates the IP address of the interface. *wildcard:* Indicates the IP address comparison mode. **0** indicates accurate matching, and **1** indicates that no comparison is performed. *area-id:* Indicates the ID of an OSPF area. An OSPF area is always associated with an address range. To facilitate management, you can use a subnet as the ID of an OSPF area.
Command Mode	OSPF routing process configuration mode
Usage Guide	By defining *ip-address* and *wildcard*, you can use one command to associate multiple interfaces with one OSPF area. To run OSPF on one interface, you must include the primary IP address of the interface in the IP address range defined by **network area**. If the IP address range defined by **network area** contains only the secondary IP address of the interface, OSPF does not run on this interface. If the interface address matches the IP address ranges defined in the **network** commands of multiple OSPF processes, the OSPF process that the interface is associated with is determined based on the best match method.

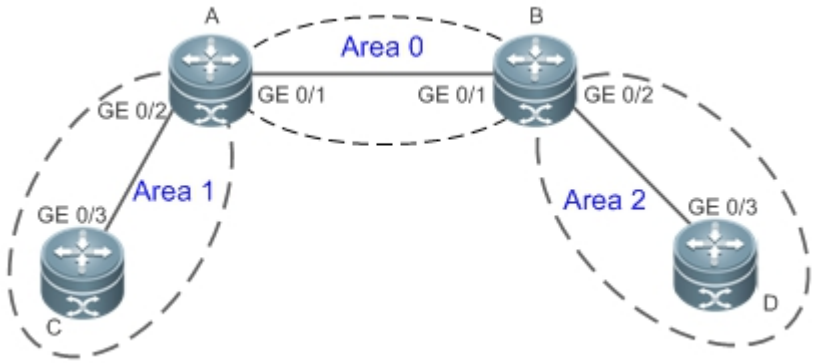
Creating a Virtual Link

Command	**area** *area-id***virtual-link** *router-id* [**authentication** [**message-digest** | **null**]] [**dead-interval**{ *seconds* |**minimal hello-multiplier** *multiplier*}] [**hello-interval***seconds*] [**retransmit-interval***seconds*] [**transmit-delay***seconds*] [[**authentication-key**[**0** |**7**]*key*] | [**message-digest-key***key-id* **md5**[**0** |**7**]*key*]]
Parameter Description	*area-id:* Indicates the ID of the OSPF transit area. The area ID can be a decimal integer or an IP address. *router-id:* Indicates the ID of a neighborrouter on the virtual link. **dead-interval***seconds:* Indicates the time that the neighbor is declared lost. The unit is second. The value ranges from 0 to 2,147,483,647. The setting of this parameter must be consistent with that on a neighbor. **minimal:** Indicates that the Fast Hello function is enabled to set the dead interval to 1s. **hello-multiplier:** Indicates the result of the dead interval multiple by the Hello interval in the Fast Hello function. *multiplier:* Indicates the number of Hello packets sent per second in the Fast Hello function. The value ranges from 3 to 20. **hello-interval** *seconds:* Indicates the interval at which OSPF sends the Hello packet to the virtual link. The unit is second. The value ranges from 1 to 65,535. The setting of this parameter must be consistent with that on a neighbor. **retransmit-interval** *seconds:* Indicates the OSPF LSA retransmission time. The unit is second. The value ranges from 1 to 65,535. **transmit-delay** *seconds:* Indicates the delay after which OSPF sends the LSA. The unit is second. The value ranges from 1 to 65,535. **authentication-key** [**0** | **7**]*key:* Defines the key for OSPF plain text authentication.

	message-digest-key *key-id***md5** [**0** | **7**]key: Defines the key ID and key for OSPF MD5 authentication. **authentication:** Sets the authentication type to plain text authentication. **message-digest:** Sets the authentication type to MD5 authentication. **null:** Indicates that authentication is disabled.
Command Mode	OSPF routing process configuration mode
Usage Guide	In the OSPF routing domain, all areas must be connected to the backbone area. If the backbone area is disconnected, a virtual link must be configured to connect to the backbone area; otherwise, network communication problems will occur. A virtual link must be created between two ABRs, and the area to which both ABRs belong is the transit area. A stub area or an NSSA area cannot be used as a transit area. A virtual link can also be used to connect other non-backbone areas. **router-id** is the ID of an OSPF neighbor router. If you are sure about the value of **router-id**, run the **show ip ospf neighbor** command to confirm the value. You can configure the loopback address as the router ID. The **area virtual-link** command defines only the authentication key of the virtual link. To enable OSPF packet authentication in the areas connected to the virtual link, you must run the **area authentication** command. OSPF supports the Fast Hello function. After the OSPF Fast Hello function is enabled, OSPF finds neighbors and detects neighbor failures faster. You can enable the OSPF Fast Hello function by specifying the **minimal** and **hello-multiplier** keywords and the **multiplier** parameter. The **minimal** keyword indicates that the death interval is set to 1s, and **hello-multiplier** indicates the number of Hello packets sent per second. In this way, the interval at which the Hello packet is sent decreases to less than 1s. If the Fast Hello function is configured for a virtual link, the Hello interval field of the Hello packet advertised on the virtual link is set to 0, and the Hello interval field of the Hello packet received on this virtual link is ignored. No matter whether the Fast Hello function is enabled, the death interval must be consistent and the **hello-multiplier** values can be inconsistent on routers at both ends of the virtual link. Ensure that at least one Hello packet can be received within the death interval. Run the **show ip ospf virtual-links** command to monitor the death interval and Fast Hello interval configured for the virtual link. The **dead-interval minimal hello-multiplier** and **hello-interval** parameters introduced for the Fast Hello function cannot be configured simultaneously.

Configuration Example

Scenario Figure 2-7



Remarks	The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 B: GE 0/1 192.168.1.2 GE 0/2 192.168.3.1 C: GE 0/3 192.168.2.2 D: GE 0/3 192.168.3.2

Configuration Steps

- Configure the interface IP addresses on all routers.
- Enable the IPv4 unicast routing function on all routers. (This function is enabled by default.)
- Configure the OSPF instances and router IDs on all routers.
- Enable OSPF on the interfaces configured on all routers.

A

```
A#configure terminal
A(config)#interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#ip address 192.168.1.1 255.255.255.0
A(config-if-GigabitEthernet 0/1)#exit
A(config)#interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)#ip address 192.168.2.1 255.255.255.0
A(config-if-GigabitEthernet 0/2)#exit
A(config)#router ospf 1
A(config-router)#router-id192.168.1.1
A(config-router)#network 192.168.1.0 0.0.0.255 area 0
A(config-router)#network 192.168.2.0 0.0.0.255 area 1
```

B

```
B#configure terminal
B(config)#interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)#ip address 192.168.1.2 255.255.255.0
B(config-if-GigabitEthernet 0/1)#exit
B(config)#interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/2)#ip address 192.168.3.1 255.255.255.0
B(config-if-GigabitEthernet 0/2)#exit
```

	B(config)#router ospf 1 B(config-router)#router-id192.168.1.2 B(config-router)#network 192.168.1.0 0.0.0.255 area 0 B(config-router)#network 192.168.3.0 0.0.0.255 area 2
C	C#configure terminal C(config)#interface GigabitEthernet 0/3 C(config-if-GigabitEthernet 0/3)#ip address 192.168.2.2 255.255.255.0 C(config-if-GigabitEthernet 0/3)#exit C(config)#router ospf 1 C(config-router)#router-id192.168.2.2 C(config-router)#network 192.168.2.0 0.0.0.255 area 1
D	D#configure terminal D(config)#interface GigabitEthernet 0/3 D(config-if-GigabitEthernet 0/3)#ip address 192.168.3.2 255.255.255.0 D(config-if-GigabitEthernet 0/3)#exit D(config)#router ospf 1 D(config-router)#router-id192.168.3.2 D(config-router)#network 192.168.3.0 0.0.0.255 area 2
Verification	● Verify that the OSPF neighbors are correct on all routers. ● Verify that the routing table is correctly loaded on all routers. ● On Router D, verify that the IP address 192.168.2.2 can be pinged successfully.
A	A# show ip ospf neighbor OSPF process 1, 2 Neighbors, 2 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.2 1 Full/DR 00:00:40192.168.1.2 GigabitEthernet 0/1 192.168.2.2 1 Full/BDR00:00:34 192.168.2.2 GigabitEthernet 0/2 A# show ip route ospf O IA 192.168.3.0/24 [110/2] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
B	B# show ip ospf neighbor OSPF process 1, 2 Neighbors, 2 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.1 1 Full/BDR 00:00:32 192.168.1.1 GigabitEthernet 0/1

	192.168.3.2 1 Full/BDR00:00:30 192.168.3.2 GigabitEthernet 0/2 B# show ip route ospf O IA 192.168.2.0/24 [110/2] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
C	C# show ip ospf neighbor OSPF process 1,1 Neighbors,1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.1 1 Full/BDR 00:00:32 192.168.2.1 GigabitEthernet 0/3 C# show ip route ospf O IA 192.168.1.0/24 [110/2] via 192.168.2.1, 00:19:05, GigabitEthernet 0/3 O IA 192.168.3.0/24 [110/3] via 192.168.2.1, 00:19:05, GigabitEthernet 0/3
D	D# show ip ospf neighbor OSPF process 1,1 Neighbors,1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.21 Full/BDR00:00:30 192.168.3.1 GigabitEthernet 0/3 D# show ip route ospf O IA 192.168.1.0/24 [110/2] via 192.168.3.1, 00:19:05, GigabitEthernet 0/3 O IA 192.168.2.0/24 [110/3] via 192.168.3.1, 00:19:05, GigabitEthernet 0/3 D# ping 192.168.2.2 Sending 5, 100-byte ICMP Echoes to 192.168.2.2, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms.

Common Errors

- OSPF cannot be enabled because the IP unicast routing function is disabled.
- The network segment configured by the **network** command does not include the interface IP addresses.
- The area IDs enabled on adjacent interfaces are inconsistent.
- The same router ID is configured on multiple routers, resulting in a router ID conflict.
- The same interface IP address is configured on multiple routers, resulting in a running error of the OSPF network.

2.4.2 Setting the Network Type

Configuration Effect

- Run OSPF to provide the IPv4 unicast routing serviceif the physical network is X.25, frame relay, or ATM.

Notes

- The OSPF basic functions must be configured.
- The broadcast network sends OSPF packets in multicast mode. Neighbors are automatically discovered, and the DR/BDR election is required.
- The P2P network sends OSPF packets in multicast mode. Neighbors are automatically discovered.
- The NBMA network sends OSPF packets in unicast mode. Neighbors must be manually specified, and the DR/BDR election is required.
- The P2MP network (without the **non-broadcast** parameter) sends OSPF packets in multicast mode. Neighbors are automatically discovered.
- The P2MP network (with the **non-broadcast** parameter) sends OSPF packets in unicast mode. Neighbors must be manually specified.

Configuration Steps

Configuring the Interface Network Type

- Optional.
- The configuration is required on routers at both ends of the link.

Configuring Neighbors

- (Optional) If the interface network type is set to NBMA or P2MP (with the **non-broadcast** parameter), neighbors must be configured.
- Neighbors are configured on routers at both ends of the NBMA or P2MP (with the **non-broadcast** parameter) network.

Configuring the Interface Priority

- (Optional) You must configure the interface priority if a router must be specified as a DR, or a router cannot be specified as a DR.
- Configure the interface priority on a router that must be specified as a DR, or cannot be specified as a DR.

Verification

- Run the **show ip ospf interface** command to verify that the network type of each interface is correct.

Related Commands

Configuring the Interface Network Type

Command	**ip ospf network** { **broadcast**| **non-broadcast**| **point-to-multipoint**[**non-broadcast**] | **point-to-point**}
Parameter Description	**broadcast**: Sets the interface network type to broadcast. **non-broadcast:** Sets the interface network type to non-broadcast. **point-to-multipoint** [**non-broadcast**]: Sets the interface network type to P2MP. If the interface does not have the

	broadcast capability, the **non-broadcast** parameter must be available. **point-to-point:** Sets the interface network type to P2P.
Command Mode	Interface configuration mode
Usage Guide	The broadcast type requires that the interface must have the broadcast capability. The P2P type requires that the interfaces are interconnected in one-to-one manner. The NBMA type requires full-meshed connections, and all interconnected routers can directly communicate with each other. The P2MP type does not raise any requirement.

Configuring Neighbors

Command	**neighbor** *ip-address* [**poll-interval***seconds*] [**priority***priority*] [**cost** *cost*]
Parameter Description	*ip-address*: Indicates the IP address of the neighbor interface. **poll-interval***seconds***:** Indicates the neighbor polling interval. The unit is second. The value ranges from 0 to 2,147,483,647. This parameter is applicable only to the NBMA interface. **priority***priority*: Indicates the neighbor priority. The value ranges from 0 to 255. This parameter is applicable only to the NBMA interface. **cost***cost***:** Indicates the cost required to reach each neighbor. There is no default value. The value ranges from 0 to 65,535. This parameter is applicable only to the P2MP interface.
Command Mode	OSPF routing process configuration mode
Usage Guide	Neighbors must be specified for the NBMA or P2MP (non-broadcast) interfaces. The neighbor IP address must be the primary IP address of this neighbor interface. If a neighbor router becomes inactive on the NBMA network, OSPF still sends Hello packets to this neighbor even if no Hello packet is received within the router death time. The interval at which the Hello packet is sent is called polling interval. When running for the first time, OSPF sends Hello packets only to neighbors whose priorities are not 0. In this way, neighbors with priorities set to 0 do not participate in the DR/BDR election. After a DR/BDR is elected, the DR/BDR sends the Hello packets to all neighbors to set up the adjacency. The P2MP (non-broadcast) network cannot dynamically discover neighbors because it does not have the broadcast capability. Therefore, you must use this command to manually configure neighbors for the P2MP (non-broadcast) network. In addition, you can use the **cost** parameter to specify the cost to reach each neighbor on the P2MP network.

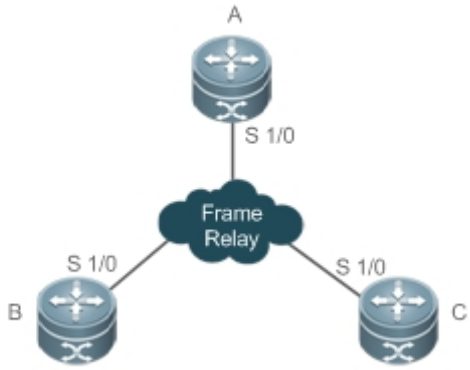
Configuring the Interface Priority

Command	**ip ospf priority** *priority*
Parameter Description	*priority*: Indicates the OSPF priority of an interface. The value ranges from 0 to 255.
Command Mode	Interface configuration mode
Usage Guide	The OSPF interface priority is contained in the Hello packet. When the DR/BDR election occurs on the OSPF broadcast network, the router with the highest priority becomes the DR or BDR. If the priorities are the same, the router with the largest router ID becomes the DR or BDR. A router with the priority set to 0 does not participate in the DR/BDR election. This command is applicable only to the OSPF broadcast and NBMA interfaces.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Setting the Interface Network Type to P2MP

Scenario Figure 2- 8	 **Remarks**: The interface IP addresses are as follows: A: S1/0 192.168.1.2 B: S1/0 192.168.1.3 C: S1/0 192.168.1.4
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Set the interface network type to P2MP on all routers.
A	A#configure terminal A(config)# interface Serial1/0 A(config-Serial1/0)# encapsulation frame-relay A(config-Serial1/0)# ip ospf network point-to-multipoint
B	B#configure terminal B(config)# interface Serial1/0 B(config-Serial1/0)# encapsulation frame-relay B(config-Serial1/0)# ip ospf network point-to-multipoint
C	C#configure terminal C(config)# interface Serial1/0 C(config-Serial1/0)# encapsulation frame-relay C(config-Serial1/0)# ip ospf network point-to-multipoint

Verification	Verify that the interface network type is P2MP. A# show ip ospf interface Serial1/0 Serial1/0 is up, line protocol is up Internet Address 192.168.1.2/24, Ifindex 2, Area 0.0.0.1, MTU 1500 Matching network config: 192.168.1.0/24 Process ID 1, Router ID 192.168.1.2, Network Type POINTOMULTIPOINT, Cost: 1 Transmit Delay is 1 sec, State Point-To-Point Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5 Hello due in 00:00:02 Neighbor Count is 1, Adjacent neighbor count is 0 Crypt Sequence Number is 4787 Hello received 465 sent 466, DD received 8 sent 8 LS-Req received 2 sent 2, LS-Upd received 8 sent 21 LS-Ack received 14 sent 7, Discarded 3

Common Errors

- The network types configured on interfaces at two ends are inconsistent, causing abnormal route learning.
- The network type is set to NBMA or P2MP (with the **non-broadcast** parameter), but neighbors are not specified.

2.4.3 Configuring Route Redistribution and Default Route

Configuration Effect

- In the OSPF domain, introduce a unicast route to other AS domains so that the unicast routing service to other AS domainscan be provided for users in the OSPF domain.
- In the OSPF domain, inject a default route to other AS domains so that the unicast routing service to other AS domains can be provided for users in the OSPF domain.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Configuring External Route Redistribution

- (Optional) This configuration is required if external routes of the OSPF domain should be introduced to an ASBR.
- This configuration is performed on an ASBR.

Generating a Default Route

- (Optional) This configuration is required if the default route should be introduced to an ASBR so that other routers in the OSPF domain access other AS domains through this ASBR by default.

- This configuration is performed on an ASBR.

Verification

- On a router inside the OSPF domain, run the **show ip route** command to verify that the unicast routes to other AS domains are loaded.
- On a router inside the OSPF domain, run the **show ip route** command to verify that the default route to the ASBR is loaded.
- Run the **ping** command to verify that the IPv4 unicast service to other AS domains is correct.

Related Commands

Configuring External Route Redistribution

Command	**redistribute** { **connected** | **ospf***process-id* | **rip** | **static** }[**match** { **internal** | **external** [**1** | **2**] | **nssa-external** [**1** | **2**] }] [**metric** *metric-value*] [**metric-type** { **1** | **2** }] [**route-map** *route-map-name*] [**subnets**] [**tag***tag-value*]
Parameter Description	**connected:** Indicates redistribution from direct routes. **ospf** *process-id:* Indicates redistribution from OSPF.*process-id* specifies an OSPF process. The value ranges from 1 to 65,535. **rip:** Indicates redistribution from RIP. **static:** Indicates redistribution from static routes. **match:** Used only when OSPF routes are redistributed. Only the routes meeting the filtering conditions are redistributed. By default, all OSPF routes can be redistributed. **metric** *metric-value:* Specifies the metric of the OSPF external LSA. *metric-value* specifies the size of the metric. The value ranges from 0 to 16,777,214. **metric-type** { **1** | **2** }: Setsthe external route type, which can be E-1 or E-2. **route-map** *route-map-name:* Setsthe redistribution filtering rules. **subnets:** Specifiesthe non-standard networks for redistribution. **tag** *tag-value:* Specifies the tag value of the route that is redistributed into the OSPF routing domain. The value ranges from 0 to 4,294,967,295.
Command Mode	OSPF routing process configuration mode
Usage Guide	After this command is configured, the router becomes an ASBR, imports related routing information to the OSPF domain, and advertises the routing information as Type 5 LSAs to other OSPF routers in the domain. If you configure redistribution of OSPF routes without specifying the **match** parameter, OSPF routes of all sub-types can be distributed by default. The latest setting of the **match** parameter is used as the initial **match** parameter. Only routes that match the sub-types can be redistributed. You can use the **no** form of the command to restore the default value of **match**. For details, see the configuration example. If **route-map** is specified, the filtering rules specified in **route-map** are applicable to original parameters of redistribution. For redistribution of OSPF, the routemap is used for filtering only when the redistributed routes meet criteria specified by **match**. The **set metric** value associated with **route-map** should fall into the range of 0 to 16,777,214. If the value exceeds this range, routes cannot be introduced. The configuration rules for the **no** form of the **redistribute** command are as follows: 1. If some parameters are specified in the **no** form of the command, default values of these parameters will be restored. 2. If no parameter is specified in the **no** form of the command, the entire command will be deleted.

Introducing a Default Route

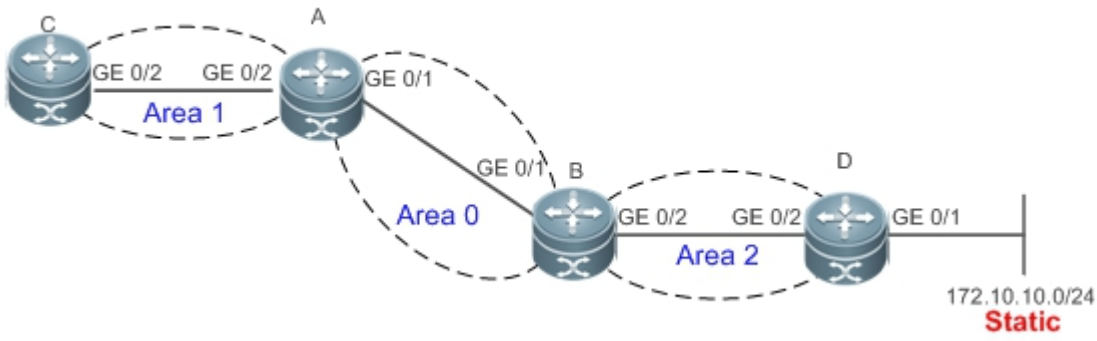
Command	**default-information originate** [**always**] [**metric** *metric*] [**metric-type** *type*] [**route-map***map-name*]
Parameter Description	**always:** Enables OSPF to generate a default route regardless of whether the local router has a default route. **metric** *metric:* Indicates the initial metric of the default route. The value ranges from 0 to 16,777,214. **metric-type***type:* Indicates the type of the default route. OSPF external routes are classified into two types: Type 1: The metric varies with routers; Type 2: The metric is the same for all routers. Type 1 external routes are more trustworthy than Type 2 external routes. **route-map** *map-name:* Indicates the associated route-map name. By default, no route-map is associated.
Command Mode	OSPF routing process configuration mode
Usage Guide	When the **redistribute** or **default-information** command is executed, the OSPF router automatically becomes an ASBR. The ASBR, however, does not automatically generate or advertise a default route to all routers in the OSPF routing domain. To have the ASBR generates a default route, configure the **default-information originate** command. If **always** is specified, the OSPF routing process advertises an external default route to neighbors regardless of whether a default route exists. This default route, however, is not displayed on the local router. To confirm whether the default route is generated, run the **show ip ospf database** command to display the OSPF link status database. The external link with the ID 0.0.0.0 describes the default route. On an OSPF neighbor, you can run the **show ip route** command to see the default route. The metric of the external default route can only be defined in the **default-information originate** command, instead of the **default-metric** command. OSPF has two types of external routes. The metric of the Type 1 external route changes, but the metric of the Type 2 external route is fixed. If two parallel paths to the same destination have the same route metric, the priority of the Type 1 route is higher than that of the Type 2 route. Therefore, the **show ip route** command displays only the Type 1 route. A router in the stub area cannot generate an external default route. The **set metric** value associated with **route-map** should fall into the range of 0 to 16,777,214. If the value exceeds this range, routes cannot be introduced.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Configuring Static Route Redistribution

Scenario Figure 2-9



Remarks	The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 B: GE 0/1 192.168.1.2 GE 0/2 192.168.3.1 C: GE 0/2 192.168.2.2 D: GE 0/1 192.168.6.2 GE 0/2 192.168.3.2

Configuration Steps

- Configure the interface IP addresses on all routers. (Omitted)
- Configure the OSPF basic functions on all routers. (Omitted)
- Introduce an external static route to Router D.

D

```
D# configure terminal
D(config)# ip route 172.10.10.0 255.255.255.0 192.168.6.3
D(config)#router ospf 1
D(config-router)# redistribute staticsubnets
```

Verification

- On Router D, run the **show ip ospf database external brief** command to verify that an LSA corresponding to an external route is generated.
- On Router C, run the **show ip route ospf** command to verify that the external static route has been introduced.

D

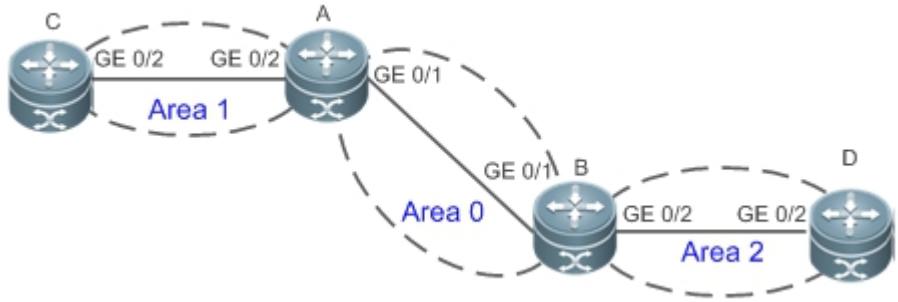
```
D# show ip ospf database external brief
                OSPF Router with ID (192.168.22.30) (Process ID 1)
                    AS External Link States

Link ID         ADV Router      Age  Seq#       CkSum  Route                Tag
172.10.10.0     192.168.22.30   11   0x80000001 0xa4bb E2 172.10.10.0/24     0
```

C

```
C# show ip route ospf
O E2 172.10.10.0/24 [110/20] via 192.168.2.1, 00:18:03, GigabitEthernet 0/2
```

Configuring the Default Route

Scenario Figure 2- 10



Remarks	The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 B: GE 0/1 192.168.1.2 GE 0/2 192.168.3.1 C: GE 0/2 192.168.2.2 D: GE 0/2 192.168.3.2

Configuration Steps

- Configure the interface IP addresses on all routers. (Omitted)
- Configure the OSPF basic functions on all routers. (Omitted)
- Configure the default route on Router D.

D

```
D# configure terminal
D(config)#router ospf 1
D(config-router)#default-information originate always
```

Verification

- On Router D, run the **show ip ospf database external brief** command to verify that an LSA corresponding to the default route is generated.
- On Router C, run the **show ip route ospf** command to verify that the OSPF default route exists.

D

```
D#show ip ospf database external brief
                OSPF Router with ID (192.168.22.30) (Process ID 1)
                    AS External Link States

Link ID         ADV Router      Age  Seq#       CkSum  Route          Tag
0.0.0.0         192.168.22.30   565  0x80000002 0xa190 E2 0.0.0.0/0    1
```

C

```
C# show ip route ospf
O E20.0.0.0/0 [110/20] via 192.168.2.1, 00:18:03, GigabitEthernet 0/2
```

Common Errors

- The subnet route is not introduced because the **subnets** parameter in the **redistribute** command is not configured.
- A routing loop is formed because the **default-information originate always** command is configured on multiple routers.
- Routes cannot be introduced because route redistribution is configured on a router in the stub area.

2.4.4 Configuring Stub Area and NSSA Area

Configuration Effect

- Configure an area located on the stub as a stub area to reduce interaction of routing information and the size of routing table, and enhance stability of routes.

Notes

- The OSPF basic functions must be configured.
- A backbone or transit area cannot be configured as a stub or an NSSA area.
- A router in the stub area cannot introduce external routes, but a router in the NSSA area can introduce external routes.

Configuration Steps

Configuring a Stub Area

- (Optional) This configuration is required if you wish to reduce the size of the routing table on routers in the area.
- The area must be configured as a stub area on all routers in this area.

Configuring an NSSA Area

- (Optional) This configuration is required if you wish to reduce the size of the routing table on routers in the area and introduce OSPF external routes to the area.
- The area must be configured as an NSSA area on all routers in this area.

Verification

Verifying the Stub Area

- On a router in the stub area, run the **show ip route** command to verify that the router is not loaded with any external routes.

Verifying the NSSA Area

- On a router in the NSSA area, run the **show ip ospf database** command to verify that the introduced external route generates Type 7 LSAs.
- On a router in the backbone area, run the **show ip route** command to verify that the router is loaded with external routes introduced from the NSSA area.

Related Commands

Configuring a Stub Area

Command	**area** *area-id* **stub** [**no-summary**]
Parameter Description	*area-id:* Indicates the ID of the stub area. **no-summary:** Prohibits the ABR from sending network summary LSAs. At this time, the stub can be called totally stub area. This parameter is configured only when the router is an ABR.
Command Mode	OSPF routing process configuration mode
Usage Guide	You must run the **area stub** command on all routers in the OSPF stub area. The ABR sends only three types of LSAs to the

	stub area: (1) Type 1: Router LSA; (2) Type 2: Network LSA; (3) Type 3: Network Summary LSA. From the routing table point of view, a router in the stub area can learn only the internal routes of the OSPF routing domain, including the internal default route generated by an ABR. A router in the stub area cannot learn external routes of the OSPF routing domain. To configure a totally stub area, add the **no-summary** keyword when running the **area stub** command on the ABR. A router in the totally stub area can learn only the internal routes of the local area, including the internal default route generated by an ABR. You can run either the **area stub** or **area default-cost** command to configure an OSPF area as a stub area. If **area stub** is used, you must configure this command on all routers connected to the stub area. If **area default-cost** is used, run this command only on the ABR in the stub area. The **area default-cost** command defines the initial cost (metric) of the internal default route.

Configuring an NSSA Area

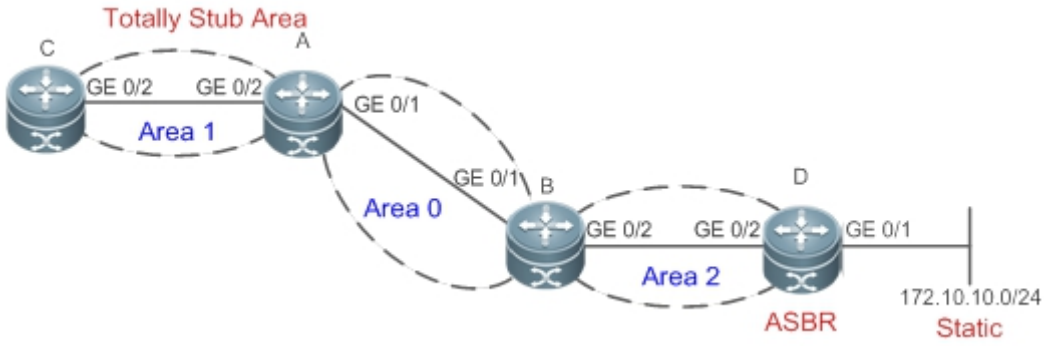
Command	**area** *area-id* **nssa** [**no-redistribution**] [**default-information-originate**[**metric***value*] [**metric-type***type*]] [**no-summary**] [**translator** [**stability-interval***seconds* | **always**] **]**
Parameter Description	*area-id:* Indicates the ID of the NSSA area. **no-redistribution:** Select this option if the router is an NSSA ABR and you want to use only the **redistribute** command to introduce the routing information into a common area instead of an NSSA area. **default-information-originate:** Indicates that a default Type 7 LSA is generated and introduced to the NSSA area. This option takes effect only on an NSSA ABR or ASBR. **metric***value:* Specifies the metric of the generated default LSA. The value ranges from 0 to 16,777,214. The default value is 1. **metric-type***type:* Specifies the route type of the generated default LSA. The values include 1 and 2. **1** represents N-1, and **2** represents N-2. The default value is 2. **no-summary:** Prohibits the ABR in the NSSA area from sending summary LSAs (Type-3 LSA). **translator:** Indicatesthat the NSSA ABR is a translator. **stability-interval***seconds:* Indicates the stability interval after the NSSA ABR is changed from a translator to a non-translator. The unit is second. The default value is 40. The value ranges from 0 to 2,147,483,647. **always:** Indicates that the current NSSA ABR always acts as a translator. The default value is the standby translator.
Command Mode	OSPF routing process configuration mode
Usage Guide	The **default-information-originate** parameter is used to generate a default Type 7 LSA. This parameter has different functions on the ABR and the ASBR in the NSSA area. On the ABR, a Type 7 LSA default route is generated regardless of whether the default route exists in the routing table. On the ASBR (not an ABR), a Type 7 LSA default route is generated only when the default route exists in the routing table. If the **no-redistribution** parameter is configured on the ASBR, other external routes introduced by OSPF through the **redistribute** command cannot be advertised to the NSSA area. This parameter is generally used when a router in the NSSA area acts both as the ASBR and the ABR. It prevents external routing information from entering the NSSA area. To further reduce the number of LSAs sent to the NSSA area, you can configure the **no-summary** parameter on the ABR to prevent the ABR from sending the summary LSAs (Type 3 LSA) to the NSSA area. **area default-cost** is used on an ABR or ASBR connected to the NSSA area. This command configures the cost of the default route sent from the ABR/ASBR to the NSSA area. By default, the cost of the default route sent to the NSSA area is 1. If an NSSA area has two or more ABRs, the ABR with the largest router ID is elected by default as the translator for

	converting Type 7 LSAs into Type 5 LSAs. If the current device is always the translator ABR for converting Type 7 LSAs into Type 5 LSAs, use the **translator always** parameter. If the translator role of the current device is replaced by another ABR, the conversion capability is retained during the time specified by **stability-interval**. If the router does not become a translator again during **stability-interval**, LSAs that are converted from Type 7 to Type 5 will be deleted from the AS after **stability-interval** expires. To prevent a routing loop, LSAs that are converted from Type 7 to Type 5 will be deleted from the AS immediately after the current device loses the translator role even if **stability-interval** does not expire. In the same NSSA area, it is recommended that **translator always** be configured on only one ABR.

Configuration Example

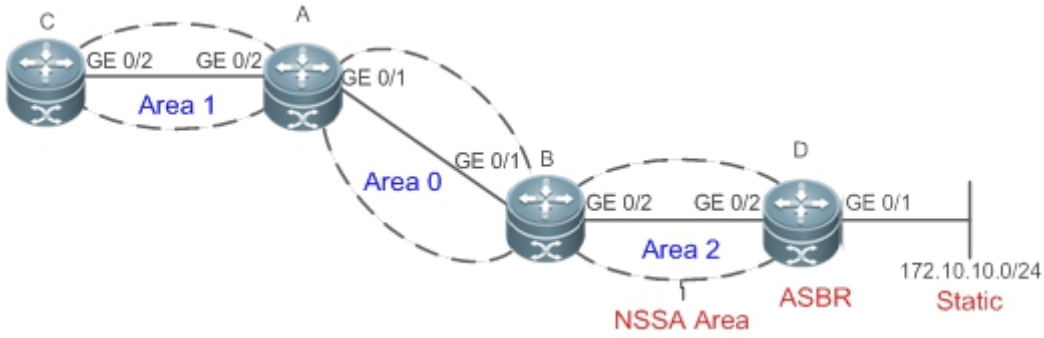
The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Configuring a Stub Area

Scenario Figure 2- 11	 **Remarks**: The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 B: GE 0/1 192.168.1.2 GE 0/2 192.168.3.1 C: GE 0/2 192.168.2.2 D: GE 0/1 192.168.6.2 GE 0/2 192.168.3.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Introduce an external static route to Router D. ● Configure area 1 as the stub area on Router A and Router C.
D	D# configure terminal D(config)#router ospf 1 D(config-router)# redistribute staticsubnets
A	A# configure terminal A(config)#router ospf 1 A(config-router)#area 1 stubno-summary

C	C# configure terminal C(config)#router ospf 1 C(config-router)#area 1 stub
Verification	On Router C, run the **show ip route ospf** command to display the routing table. Verify that there is only one default inter-area route, and no external static route is introduced from Router D.
	C#show ip route ospf O*IA 0.0.0.0/0 [110/2] via 192.168.2.1, 00:30:53, GigabitEthernet 0/2

Configuring an NSSA Area

Scenario Figure 2- 12	 **Remarks**: The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 B: GE 0/1 192.168.1.2 GE 0/2 192.168.3.1 C: GE 0/2 192.168.2.2 D: GE 0/1 192.168.6.2 GE 0/2 192.168.3.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Introduce an external static route to Router D. ● Configure area 2 as the NSSA area on Router B and Router D.
B	B# configure terminal B(config)#router ospf 1 B(config-router)#area 2 nssa
D	D# configure terminal D(config)#ip route 172.10.10.0 255.255.255.0 192.168.6.2 D(config)#router ospf 1 D(config-router)#redistribute static subnets D(config-router)#area 2 nssa

Verification	● On Router D, verify that the Type 7 LSA, 172.10.10.0/24, is generated. ● On Router B, verify that Type 5 and Type 7 LSAs coexist on 172.10.10.0/24. ● On Router B, verify that the N-2 route of 172.10.10.0/24 is generated.
D	D# show ip ospf database nssa-external OSPF Router with ID (192.168.6.2) (Process ID 1) NSSA-external Link States (Area 0.0.0.1 [NSSA]) LS age: 61 Options: 0x8 (-|-|-|-|N/P|-|-|-) LS Type: AS-NSSA-LSA Link State ID: 172.10.10.0 (External Network Number For NSSA) Advertising Router: 192.168.6.2 LS Seq Number: 80000001 Checksum: 0xc8f8 Length: 36 Network Mask: /24 Metric Type: 2 (Larger than any link state path) TOS: 0 Metric: 20 NSSA: Forward Address: 192.168.6.2 External Route Tag: 0
B	B# show ip ospf database nssa-external OSPF Router with ID (192.168.3.1) (Process ID 1) NSSA-external Link States (Area 0.0.0.1 [NSSA]) LS age: 314 Options: 0x8 (-|-|-|-|N/P|-|-|-) LS Type: AS-NSSA-LSA Link State ID: 172.10.10.0 (External Network Number For NSSA) Advertising Router: 192.168.6.2 LS Seq Number: 80000001 Checksum: 0xc8f8 Length: 36 Network Mask: /24 Metric Type: 2 (Larger than any link state path) TOS: 0

```
            Metric: 20
            NSSA: Forward Address: 192.168.6.2
            External Route Tag: 0
B# show ip ospf database external
                OSPF Router with ID (192.168.3.1) (Process ID 1)
                    AS External Link States
  LS age: 875
  Options: 0x2 (-|-|-|-|-|-|E|-)
  LS Type: AS-external-LSA
  Link State ID: 172.10.10.0 (External Network Number)
  Advertising Router: 192.168.3.1
  LS Seq Number: 80000001
  Checksum: 0xd0d3
  Length: 36
  Network Mask: /24
            Metric Type: 2 (Larger than any link state path)
            TOS: 0
            Metric: 20
            Forward Address: 192.168.6.2
            External Route Tag: 0

B# show ip route ospf
O N2    172.10.10.0/24 [110/20] via 192.168.3.2, 00:06:53, GigabitEthernet 0/2
```

Common Errors

- Configurations of the area type are inconsistent on routers in the same area.
- External routes cannot be introduced because route redistribution is configured on a router in the stub area.

2.4.5 Configuring Route Summarization

Configuration Effect

- Summarize routes to reduce interaction of routing information and the size of routing table, and enhance stability of routes.
- Shield or filter routes.

Notes

- The OSPF basic functions must be configured.

- The address range of summarized routes may exceed the actual network range in the routing table. If data is sent to a network beyond the summarization range, a routing loop may be formed and the router processing load may increase. To prevent these problems, a discard route must be added to the routing table or shield or filter routes.

Configuration Steps

Configuring Inter-Area Route Summarization

- (Optional) This configuration is required when routes of the OSPF area need to be summarized.
- Unless otherwise required, this configuration should be performed on an ABR in the area where routes to be summarized are located.

Configuring External Route Summarization

- (Optional) This configuration is required when routes external to the OSPF domain need to be summarized.
- Unless otherwise required, this configuration should be performed on an ASBR to which routes to be summarized are introduced.

Verification

Run the **show ip route ospf** command to verify that individual routes do not exist and only the summarized route exists.

Related Commands

Configuring Inter-Area Route Summarization

Command	**area** *area-id* **range** *ip-address net-mask* [**advertise** | **not-advertise**] [**cost** *cost*]
Parameter Description	*area-id*: Specifies the ID of the OSPF area to which the summarized route should be injected. The area ID can be a decimal integer or an IP address. *ip-address net-mask*: Defines the network segment of the summarized route. **advertise** | **not-advertise**: Specifies whether the summarized route should be advertised. **cost** *cost:* Indicates the metric of the summarized route. The value ranges from 0 to 16777215.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command can be executed only on the ABR. It is used to combine or summarize multiple routes of an area into one route, and advertise the route to other areas. Combination of the routing information occurs only on the boundary of an area. Routers inside the area can learn specific routing information, whereas routers in other areas can learn only one summarized route. In addition, you can set **advertise** or **not-advertise** to determine whether to advertise the summarized route to shield and filter routes. By default, the summarized route is advertised. You can use the **cost** parameter to set the metric of the summarized route. You can configure route summarization commands for multiple areas. This simplifies routes in the entire OSPF routing domain, and improve the network forwarding performance, especially for a large-sized network. When multiple route summarization commands are configured and have the inclusive relationship with each other, the area range to be summarized is determined based on the maximum match principle.

Configuring External Route Summarization

Command	**summary-address** *ip-address net-mask* [**not-advertise** | **tag** *value*]
Parameter Description	*ip-address*: Indicates the IP address of the summarized route. *net-mask*: Indicates the subnet mask of the summarized route.

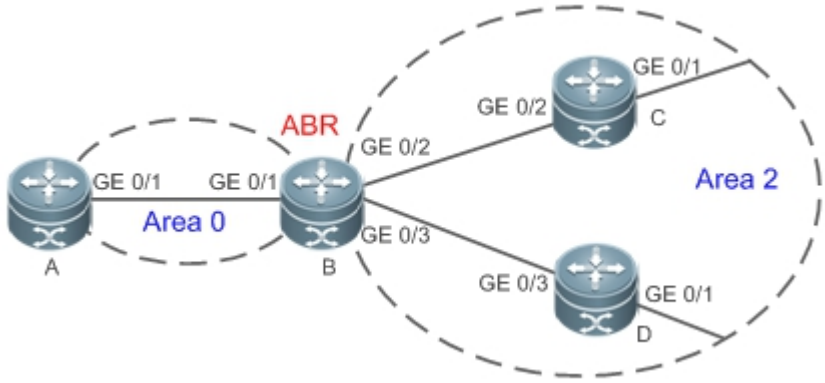
	not-advertise: Indicates that the summarized route is not advertised. If this parameter is not specified, the summarized route is advertised. **tag***value:* Indicates the tag of the summarized route. The value ranges from 0 to 4,294,967,295.
Command Mode	OSPF routing process configuration mode
Usage Guide	When routes are redistributed from other routing processes and injected to the OSPF routing process, each route is advertised to the OSPF routers using an external LSA. If the injected routes are a continuous address space, the ABR can advertised only one summarized route to significantly reduce the size of the routing table. **area range** summarizesthe routes between OSPF routes, whereas **summary-address** summarizes external routes of the OSPF routing domain. When configured on the NSSA ABR translator, **summary-address** summarizes redistributed routes and routes obtained based on the LSAs that are converted from Type 7 to Type 5. When configured on the ASBR (not an NSSA ABR translator), **summary-address** summarizes only redistributed routes.

Configuring a Discard Route

Command	**discard-route** { **internal** | **external** }
Parameter Description	**internal:** Indicates that the discard route generated by the **area range** command can be added. **external:** Indicates that the discard route generated by the **summary-address** command can be added.
Command Mode	OSPF routing process configuration mode
Usage Guide	The address range of summarized routes may exceed the actual network range in the routing table. If data is sent to a network beyond the summarization range, a routing loop may be formed and the router processing load may increase. To prevent these problems, a discard route must be added to the routing table on the ABR or ASBR. This route is automatically generated, and is not advertised.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 13	 **Remarks** The interface IP addresses are as follows: A: GE0/1 192.168.1.1 B: GE0/1 192.168.1.2 GE0/2 172.16.2.1 GE0/3 172.16.3.1 C: GE0/2 172.16.2.2 GE0/1 172.16.4.2 D: GE0/2 172.16.3.2 GE0/1 172.16.5.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Summarize routes of area 2 on Router B.
B	B# configure terminal B(config)#router ospf 1 B(config-router)#area 2 range 172.16.0.0 255.255.0.0
Verification	On Router A, verify that the entry 172.16.0.0/16 is added to the routing table.
A	A#show ip route ospf O IA 172.16.0.0/16 [110/2] via 192.168.1.2, 00:01:04, GigabitEthernet 0/1

Common Errors

- Inter-area route summarization cannot be implemented because the **area range** command is configured on a non-ABR device.

2.4.6 Configuring Route Filtering

Configuration Effect

- Routes that do not meet filtering conditions cannot be loaded to the routing table, or advertised to neighbors. Network users cannot access specified destination network.

Notes

- The OSPF basic functions must be configured.
- Filtering routes by using the **distribute-list in** command affects forwarding of local routes, but does not affect route computation based on LSAs. Therefore, if route filtering is configured on the ABR, Type 3 LSAs will still be generated and advertised to other areas because routes can still be computed based on LSAs. As a result, black-hole routes are generated. In this case, you can run

the **area filter-list** or **area range** (containing the **not-advertise** parameter) command on the ABR to prevent generation of black-hole routes.

Configuration Steps

Configuring Inter-Area Route Filtering

- (Optional) This configuration is recommended if users should be restricted from accessing the network in a certain OSPF area.
- Unless otherwise required, this configuration should be performed on an ABR in the area where filtered routes are located.

Configuring Redistributed Route Filtering

- (Optional) This configuration is required if external routes introduced by the ASBR need to be filtered.
- Unless otherwise required, this configuration should be performed on an ASBR to which filtered routes are introduced.

Configuring Learned Route Filtering

- (Optional) This configuration is required if users should be restricted from accessing a specified destination network.
- Unless otherwise required, this configuration should be performed on a router that requires route filtering.

Verification

- Run the **show ip route** command to verify that the router is not loaded with routes that have been filtered out.
- Run the **ping** command to verify that the specified destination network cannot be accessed.

Related Commands

Configuring a Passive Interface

Command	**passive-interface**{ **default** | *interface-type interface-number* | *interface-type interface-numberip-address*}
Parameter Description	*interface-type interface-number*: Indicates the interface that should be configured as a passive interface. **default:** Indicates that all interface will be configured as passive interfaces. *interface-type interface-numberip-address:* Specifies an address of the interface as the passive address.
Command Mode	OSPF routing process configuration mode
Usage Guide	To prevent other routers on the network from learning the routing information of the local router, you can configure a specified network interface of the local router as the passive interface, or a specified IP address of a network interface as the passive address.

Configuring the LSA Update Packet Filtering

Command	**ip ospf database-filter all out**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Enable this function on an interface to prevent sending the LSA update packet on this interface. After this function is enabled, the local router does not advertise the LSA update packet to neighbors, but still sets up the adjacency with neighbors and receives LSAs from neighbors.

Configuring Inter-Area Route Filtering

Command	**area***area-id***filter-list** {**access***acl-name*| **prefix** *prefix-name*} {**in** | **out**}
Parameter Description	*area-id*: Indicates the area ID. **access** *acl-name*: Indicates the associated ACL. **prefix** *prefix-name*: Indicates the associated prefix list. **in** | **out:** Filters routes that are received by or sent from the area.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command can be configured only on an ABR. Use this command when it is required to configure filtering conditions for inter-area routes on the ABR.

Configuring Redistributed Route Filtering

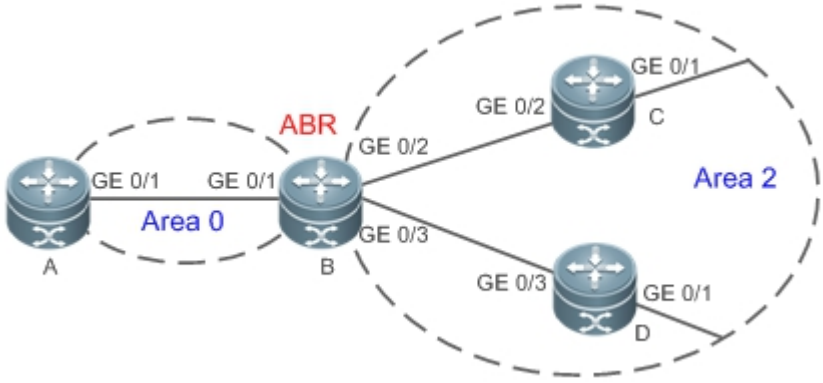
Command	**distribute-list** { [*access-list-number* | *name*] | **prefix** *prefix-list-name* } **out** [**connected | ospf** *process-id* **| rip | static**]
Parameter Description	*access-list-number* | *name*: Uses the ACL for filtering. **prefix** *prefix-list-name*: Uses the prefixlist for filtering. **connected | ospf** *process-id* | **rip** | **static**: Indicates the source of routes to be filtered.
Command Mode	OSPF routing process configuration mode
Usage Guide	**distribute-list out** is similar to **redistribute route-map**, and is used to filter routes that are redistributed from other protocols to OSPF. The **distribute-list out** command itself does not redistribute routes, and is generally used together with the **redistribute** command. The ACL and the prefixlist filtering rules are mutually exclusive in the configuration. That is, if the ACL is used for filtering routes coming from a certain source, the prefixlist cannot be configured to filter the same routes.

Configuring Learned Route Filtering

Command	**distribute-list** {[*access-list-number* | *name*] | **prefix***prefix-list-name* [**gateway** *prefix-list-name*] | **route-map** *route-map-name* } **in** [*interface-typeinterface-number*]
Parameter Description	*access-list-number* | *name*: Uses the ACL for filtering. **gateway***prefix-list-name*: Uses the gateway for filtering. **prefix***prefix-list-name*: Uses the prefixlist for filtering. **route-map** *route-map-name:* Uses the route map for filtering. *interface-type interface-number:* Specifies the interface for which LSA routes are filtered.
Command Mode	OSPF routing process configuration mode
Usage Guide	Filter routes that are computed based on received LSAs. Only routes meeting the filtering conditions can be forwarded. The command does not affect the LSDB or the routing tables of neighbors. The ACL, prefix list, and route map filtering rules are mutually exclusive in the configuration. That is, if the ACL is used for filtering routes of a specified interface, the prefix list or router map cannot be configured for filtering routes of the same interface.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 14	 **Remarks**: The interface IP addresses are as follows: A: GE0/1 192.168.1.1 B: GE0/1 192.168.1.2 GE0/2 172.16.2.1 GE0/3 172.16.3.1 C: GE0/2 172.16.2.2 GE0/3 172.16.4.2 D: GE0/2 172.16.3.2 GE0/3 172.16.5.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, configure route filtering.
A	A# configure terminal A(config)#access-list 3 permit host 172.16.5.0 A(config)#router ospf 1 A(config-router)#distribute-list 3 in GigabitEthernet 0/1
Verification	● On Router A, check the routing table. Verify that only the entry 172.16.5.0/24 is loaded.
A	A# show ip route ospf O 172.16.5.0/24 [110/2] via 192.168.1.2, 10:39:40, GigabitEthernet 0/1

Common Errors

- Filtering routes by using the **distribute-list in** command affects forwarding of local routes, but does not affect route computation based on LSAs. Therefore, if route filtering is configured on the ABR, Type 3 LSAs will still be generated and advertised to other areas because routes can still be computed based on LSAs. As a result, black-hole routes are generated.

2.4.7 Modifying Route Cost and AD

Configuration Effect

- Change the OSPF routes to enable the traffic pass through specified nodes or avoid passing through specified nodes.
- Change the sequence that a router selects routes so as to change the priorities of OSPF routes.

Notes

- The OSPF basic functions must be configured.

- If you run the **ip ospf cost** command to configure the cost of an interface, the configured cost will automatically overwrite the cost that is computed based on the auto cost.

Configuration Steps

Configuring the Reference Bandwidth

- Optional.
- A router is connected with lines with different bandwidths. This configuration is recommended if you wish to preferentially select the line with a larger bandwidth.

Configuring the Cost of an Interface

- Optional.
- A router is connected with multiple lines. This configuration is recommended if you wish to manually specify a preferential line.

Configuring the Default Metric for Redistribution

- Optional.
- This configuration is mandatory if the cost of external routes of the OSPF domain should be specified when external routes are introduced to an ASBR.

Configuring the Maximum Metric

- Optional.
- A router may be unstable during the restart process or a period of time after the router is restarted, and users do not want to forward data through this router. In this case, this configuration is recommended.

Configuring the AD

- Optional.
- This configuration is mandatory if you wish to change the priorities of OSPF routes on a router that runs multiple unicast routing protocols.

Verification

- Run the **show ip ospf interface** command to verify that the costs of interfaces are correct.
- Run the **show ip route** command to verify that the costs of external routes introduced to the ASBR are correct.
- Restart the router. Within a specified period of time, data is not forwarded through the restarted router.

Related Commands

Configuring the Reference Bandwidth

Command	**auto-costreference-bandwidth** *ref-bw*
Parameter Description	*ref-bw*: Indicates the reference bandwidth. The unit is Mbps. The value ranges from 1 to 4,294,967.
Command Mode	OSPF routing process configuration mode
Usage Guide	By default, the cost of an OSPF interface is equal to the reference value of the auto cost divided by the interface

	bandwidth. Run the **auto-cost** command to obtain the reference value of the auto cost. The default value is 100 Mbps. Run the **bandwidth** command to set the interface bandwidth. The costs of OSPF interfaces on several typical lines are as follows: 64Kbps serial line: The cost is 1562. E1 line: The cost is 48. 10M Ethernet: The cost is 10. 100M Ethernet: The cost is 1. If you run the **ip ospf cost** command to configure the cost of an interface, the configured cost will automatically overwrite the cost that is computed based on the auto cost.

Configuring the Cost of an Interface

Command	**ip ospf cost***cost*
Parameter Description	*cost*: Indicates the cost of an OSPF interface. The value ranges from 0 to 65,535.
Command Mode	Interface configuration mode
Usage Guide	By default, the cost of an OSPF interface is equal to the reference value of the auto cost divided by the interface bandwidth. Run the **auto-cost** command to obtain the reference value of the auto cost. The default value is 100 Mbps. Run the **bandwidth** command to set the interface bandwidth. The costs of OSPF interfaces on several typical lines are as follows: 64Kbps serial line: The cost is 1562. E1 line: The cost is 48. 10M Ethernet: The cost is 10. 100M Ethernet: The cost is 1. If you run the **ip ospf cost** command to configure the cost of an interface, the configured cost will automatically overwrite the cost that is computed based on the auto cost.

Configuring the Cost of the Default Route in a Stub or an NSSA Area

Command	**area***area-id* **default-cost***cost*
Parameter Description	*area-id*: Indicates the ID of the stub or NSSA area. *cost*: Indicates the cost of the default summarized route injected to the stub or NSSA area. The value ranges from 0 to 16,777,215.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command takes effect only on an ABR in a stub area or an ABR/ASBR in an NSSA area. An ABR in a stub area or an ABR/ASBR in an NSSA area is allowed to advertise an LSA indicating the default route in the stub or NSSA area. You can run the **area default-cost** command to modify the cost of the advertised LSA.

Configuring the Default Metric for Redistribution

Command	**default-metric** *metric*
Parameter Description	*metric*: Indicates the default metric of the OSPF redistributed route. The value ranges from 1 to 16,777,214.

Command Mode	OSPF routing process configuration mode
Usage Guide	The **default-metric**command must be used together with the **redistribute**command to modify the initial metrics of all redistributed routes. The **default-metric**command does not take effect on external routes that are injected to the OSPF routing domain by the **default-information originate** command.

Configuring the Maximum Metric

Command	**max-metric router-lsa** [**external-lsa** [*max-metric-value*]] [**include-stub**] [**on-startup**[*seconds*]] [**summary-lsa** [*max-metric-value*]]
Parameter Description	**router-lsa**: Sets the metrics of non-stub links in the Router LSA to the maximum value (0xFFFF). **external-lsa**: Allows a router to replace the metrics of external LSAs (including Type 5 and Type 7 LSAs) with the maximum metric. *max-metric-value*: Indicates the maximum metric of the LSA. The default value is 16711680. The value ranges from 1 to 16,777,215. **include-stub:** Sets the metrics of stub links in the Router LSA advertised by the router to the maximum value. **on-startup:** Allows a router to advertises the maximum metric when started. *seconds:* Indicates the interval at which the maximum metric is advertised. The default value is 600s. The value ranges from 5 to 86,400. **summary-lsa:** Allows a router to replace the metrics of summary LSAs (including Type 3 and Type 4 LSAs) with the maximum metric.
Command Mode	OSPF routing process configuration mode
Usage Guide	After the **max-metric router-lsa** command is executed, the metrics of the non-stub links in the Router LSAs generated by the router will be set to the maximum value (0xFFFF). If you cancel this configuration or the timer expires, the normal metrics of the links are restored. By default, if the **max-metric router-lsa** command is executed, the stub links still advertise common metrics, that is, the costs of outbound interfaces. If the **include-stub** parameter is configured, the stub links will advertise the maximum metric. If an ABR does not wish to transfer inter-area traffic, use the **summary-lsa** parameter to set the metric of the Summary LSA to the maximum metric. If an ASBR does not wish to transfer external traffic, use the **external-lsa** parameter to set the metric of the external LSA to the maximum metric. The **max-metric router-lsa** command is generally used in the following scenarios: Restart a device. After the device is restarted, IGP generally converges faster, and other devices attempt to forward traffic through the restarted device. ● Add a device to the network but the device is not used to transfer traffic. The device is added to the network. If a candidate path exists, the current device is not used to transfer traffic. If a candidate path does not exist, the current device is still used to transfer traffic. ● Delete a device gracefully from the network. After the **max-metric router-lsa** command is executed, the current device advertises the maximum metric among all metrics of routes. In this way, other devices on the network can select the standby path for data transmission before the device is shut down. In the earlier OSPF version (RFC1247 or earlier), the links with the maximum metric (0xFFFF) in the LSAs do not participate in the SPF computation, that is, no traffic is sent to routers that generate these LSAs.

Configuring RFC1583Compatibility

Command	**compatible rfc1583**
Parameter Description	N/A
Command Mode	OSPF routing process configuration mode
Usage Guide	When there are multiple paths to an ASBR or the forwarding address of an external route, RFC1583 and RFC2328 define different routing rules. If RFC1583 compatibilityis configured, a path in the backbone area or an inter-area path is preferentially selected. If RFC1583 compatibilityis not configured, a path in a non-backbone area is preferentially selected.

Configuring the AD

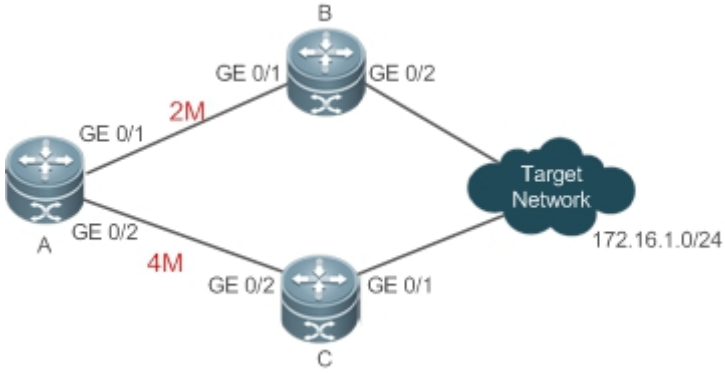
Command	**distance** { *distance* | **ospf** { [**intra-area***distance*] [**inter-area***distance*][**external** *distance*]} }
Parameter Description	*distance*: Indicates the AD of a route. The value ranges from 1 to 255. **intra-area** *distance*: Indicates the AD of an intra-area route. The value ranges from 1 to 255. **inter-area** *distance*: Indicates the AD of an inter-area route. The value ranges from 1 to 255. **external** *distance:* Indicates the AD of an external route. The value ranges from 1 to 255.
Command Mode	OSPF routing process configuration mode
Usage Guide	Use this command to specify different ADs for different types of OSPF routes.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Configuring the Cost of an Interface

Scenario Figure 2- 15



Remarks	The interface IP addresses are as follows: A: GE0/1 192.168.1.1 GE0/2 192.168.2.1 B: GE0/1 192.168.1.2 GE0/2 192.168.3.2 C: GE0/1 192.168.4.2 GE0/2 192.168.2.2

Configuration Steps

- Configure the interface IP addresses on all routers. (Omitted)
- Configure the OSPF basic functions on all routers. (Omitted)
- On Router A, configure the cost of each interface.

A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip ospf cost 10 A(config)# interface GigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)# ip ospf cost 20
Verification	On Router A, check the routing table. The next hop of the optimum path to 172.16.1.0/24 is Router B.
A	A# show ip route ospf O E2172.16.1.0/0 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1

Common Errors

- If the cost of an interface is set to 0 in the **ip ospf cost** command, a route computation error may occur. For example, a routing loop is obtained.

2.4.8 Enabling Authentication

Configuration Effect

- All routers connected to the OSPF network must be authenticated to ensure stability of OSPF and protect OSPF against intrusions.

Notes

- The OSPF basic functions must be configured.
- If authentication is configured for an area, the configuration takes effect on all interfaces that belong to this area.
- If authentication is configured for both an interface and the area to which the interface belongs, the configuration for the interface takes effect preferentially.

Configuration Steps

Configuring the Authentication Type of an Area

- (Optional) This configuration is recommended if the same authentication type should be used on all interfaces in the same area.
- This configuration is required if a router accesses a network that requires authentication.

Configuring the Authentication Type of an Interface

- (Optional) This configuration is recommended if the different authentication types should be used on different interfaces in the same area.
- This configuration is required if a router accesses a network that requires authentication.

Configuring a Plain Text Authentication Key for an Interface

- Optional.
- This configuration is required if a router accesses a network that requires plain text authentication.

Configuring an MD5 Authentication Key for an Interface

- (Optional) MD5 authentication features a high security, and therefore is recommended. You must configure either plain text authentication or MD5 authentication.
- This configuration is required if a router accesses a network that requires MD5 authentication.

Verification

- If routers are configured with different authentication keys, run the **show ip ospf neighbor** command to verify that there is no OSPF neighbor.
- If routers are configured with the same authentication key, run the **show ip ospf neighbor** command to verify that there are OSPF neighbors.

Related Commands

Configuring the Authentication Type of an Area

Command	**area** *area-id***authentication** [**message-digest**]
Parameter Description	*area-id*: Indicatesthe ID of the area where OSPF authentication is enabled. The area ID can be a decimal integer or an IP address. **message-digest**: Enables MD5 authentication.
Command Mode	OSPF routing process configuration mode
Usage Guide	The FSOS supports three authentication types: (1) Type 0: No authentication is required. If this command is not configured to enable OSPF authentication, the authentication type in the OSPF data packet is 0. (2) Type 1: The authentication type is plain text authentication if this command is configured but does not contain the **message-digest** parameter. (3) Type 3: The authentication type is MD5 authentication if this command is configured and contains the **message-digest** parameter. All routers in the same OSPF area must use the same authentication type. If authentication is enabled, the authentication key must be configured on interfaces that are connected to neighbors. You can run the interface configuration command **ip ospf authentication-key** to configure the plain text authentication key, or **ip ospf message-digest-key** to configure the MD5 authentication key.

Configuring the Authentication Type of an Interface

Command	**ip ospfauthentication** [**message-digest** | **null**]
Parameter Description	**message-digest**: Indicates that MD5 authentication is enabled on the current interface. **null**: Indicates that authentication is disabled.
Command Mode	Interface configuration mode
Usage Guide	If the **ip ospfauthentication** command does not contain any option, it indicates that plain text authentication is enabled. If you use the **no** form of the command to restore the default authentication mode, whether authentication is enabled is determined by the authentication type that is configured in the area to which the interface belongs. If the authentication type is set to null, authentication is disabled forcibly.When authentication is configured for both an interface and the area to which the interface belongs, the authentication type configured for the interface is used preferentially.

Configuring a Plain Text Authentication Key for an Interface

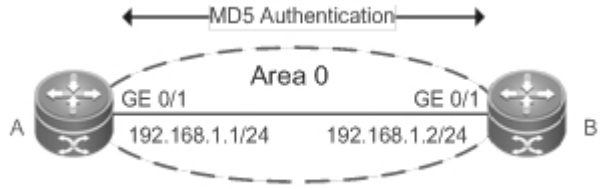
Command	**ip ospf authentication-key**[**0** |**7**]*key*
Parameter Description	**0**: Indicates that the key is displayed in plain text. **7**: Indicates that the key is displayed in cipher text. *key*: Indicates the key. The key is a string of up to eight characters.
Command Mode	Interface configuration mode
Usage Guide	The key configured by the **ip ospf authentication-key** command will be inserted to the headers of all OSPF packets. If the keys are inconsistent, two directly connected devices cannot set up the OSPF adjacency and therefore cannot exchange the routing information. Different keys can be configured for different interface, but all routers connected to the same physical network segment must be configured with the same key. You can enable or disable authentication in an OSPF area by running the **areaauthentication** command in OSPF routing process configuration mode. You can also enable authentication on an individual interface by running the **ip ospf authentication** command in interface configuration mode. When authentication is configured for both an interface and the area to which the interface belongs, the authentication type configured for the interface is used preferentially.

Configuring an MD5 Authentication Key for an Interface

Command	**ip ospf message-digest-key** *key-id* **md5**[**0** |**7**]*key*
Parameter Description	*key-id:* Indicates the key ID. The value ranges from 1 to 255. **0**: Indicates that the key is displayed in plain text. **7**: Indicates that the key is displayed in cipher text. *key:* Indicates the key. The key is a string of up to 16 characters.
Command Mode	Interface configuration mode
Usage Guide	The key configured by the **ip ospf message-digest-key** command will be inserted to the headers of all OSPF packets. If the keys are inconsistent, two directly connected devices cannot set up the OSPF adjacency and therefore cannot exchange the routing information. Different keys can be configured for different interface, but all routers connected to the same physical network segment must be configured with the same key. The same key ID on neighbor routers must correspond to the same key. You can enable or disable authentication in an OSPF area by running the **area authentication** command in OSPF routing process configuration mode. You can also enable authentication on an individual interface by running the **ip ospf authentication** command in interface configuration mode. When authentication is configured for both an interface and the area to which the interface belongs, the authentication type configured for the interface is used preferentially. The FSOS software supports smooth modification of the MD5 authentication key. A new MD5 authentication key must be first added before the old key can be deleted. When an OSPF MD5 authentication key is added to a router, the router determines that other routers do not use the new key yet and therefore uses different keys to send multiple OSPF packets until it confirms that the new key has been configured on neighbors. After configuring the new key all routers, you can delete the old key.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 16	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Configure the authentication type and MD5 authentication key on all routers.
A	A# configure terminal A(config)#router ospf 1 A(config-router)#area 0 authentication message-digest A(config-router)#exit A(config)#interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ip ospf message-digest-key 1 md5 hello
B	B# configure terminal B(config)#router ospf 1 B(config-router)#area 0 authentication message-digest B(config-router)#exit B(config)#interface GigabitEthernet 0/3 B(config-if-GigabitEthernet 0/3)#ip ospf message-digest-key 1 md5 hello
Verification	On Router A and Router B, verify that the OSPF neighbor status is correct.
A	A#show ip ospf neighbor OSPF process 1, 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.2 1 Full/DR 00:00:32 192.168.1.2 GigabitEthernet 0/1
B	A#show ip ospf neighbor OSPF process 1, 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.1 1 Full/DR 00:00:32 192.168.1.1 GigabitEthernet 0/1

Common Errors

- The authentication modes configured on routers are inconsistent.
- The authentication keys configured on routers are inconsistent.

2.4.9 Enabling Overflow

Configuration Effect

- New routes are not loaded to routers when the router memory is insufficient.
- New routes are not loaded to routers when the usage of the database space reaches the upper limit.

Notes

- The OSPF basic functions must be configured.
- After a router enters the overflow state, you can run the **clear ip ospf process** command, or stop and then restart the OSPF to exit the overflow state.

Configuration Steps

Configuring the Memory Overflow Function

- Optional.
- This configuration is recommended if a large number of routes exist in the domain and may cause insufficiency of the router memory.

Configuring the Database Overflow Function

- Optional.
- This configuration is recommended if a large number of routes exist in the domain and may cause insufficiency of the router memory.

Configuring the External LSA Database Overflow Function

- Optional.
- This configuration is recommended if the ASBR introduces a large number of external routes and the router memory may be insufficient.

Verification

- After the memory becomes insufficient, add new routers to the network, and run the **show ip route** command to verify that new routes are not loaded.
- After the usage of the database space reaches the upper limit, add new routers to the network, and run the **show ip route** command to verify that new routes are not loaded.

Related Commands

Configuring the Memory Overflow Function

Command	**overflow memory-lack**
Parameter Description	N/A

Command Mode	OSPF routing process configuration mode
Usage Guide	The OSPF process enters the overflow state to discard newly-learned external routes. This behavior can effectively ensure that the memory usage does not increase. After the overflow function is enabled, the OSPF process enters the overflow state and discards newly-learned external routes, which may cause a routing loop on the entire network. To reduce the occurrence probability of this problem, OSPF generates a default route to the null interface, and this route always exists in the overflow state. You can run the **clear ip ospf process** command to reset the OSPF process so that the OSPF process can exit the overflow state. You can use the **no** form of the command to prevent the OSPF process from entering the overflow state when the memory is insufficient. This, however, may lead to over-consumption of the memory resource, after which the OSPF process will stop and delete all the learned routes.

Configuring the Database Overflow Function

Command	**overflow database***number* [**hard** | **soft**]
Parameter Description	*number*: Indicates the maximum number of LSAs. The value ranges from 1 to 4,294,967,294. **hard**: Indicates that the OSPF process will be stopped if the number of LSAs exceeds the limit. **soft:** Indicates that a warning will be generated if the number of LSAs exceeds the limit.
Command Mode	OSPF routing process configuration mode
Usage Guide	If the number of LSAs exceeds the limit, use the **hard** parameter if the OSPF process should be stopped, and use the **soft** parameter if a warning should be generated without stopping the OSPF process.

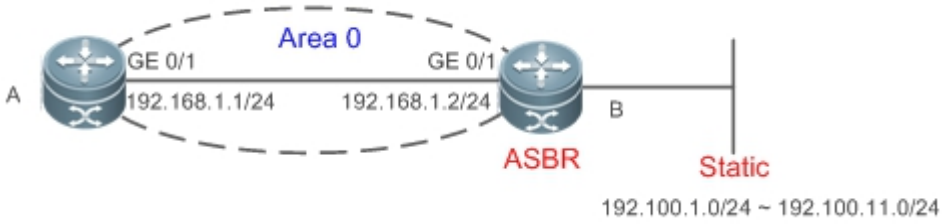
Configuring the External LSA Database Overflow Function

Command	**overflow database external** *max-dbsize wait-time*
Parameter Description	*max-dbsize*: Indicates the maximum number of external LSAs. This value must be the same on all routers in the same AS. The value ranges from 0 to 2,147,483,647. *wait-time*: Indicates the waiting time after a router in overflow state attempts to restore the normal state. The value ranges from 0 to 2,147,483,647.
Command Mode	OSPF routing process configuration mode
Usage Guide	When the number of external LSAs of a router exceeds the configured **max-dbsize**, the router enters the overflow state. In this state, the router no longer loads external LSAs and deletes external LSAs that are generated locally. After *wait-time* elapses, the device restores the normal state, and loads external LSAs again. When using the overflow function, ensure that the same **max-dbsize** is configured on all routers in the OSPF backbone area and common areas; otherwise, the following problems may occur: Inconsistent LSDBs throughout network are inconsistent, and the failure to achieve the full adjacency Incorrect routes, including routing loops Frequent retransmission of AS external LSAs

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Configuring the External LSA Database Overflow Function

Scenario Figure 2- 17	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router B, configure redistribution and introduce external static routes. ● On Router B, configure the maximum number of external LSAs.
B	B# configure terminal B(config)# router ospf 1 B(config-router)# redistribute static subnets
A	A# configure terminal A(config)# router ospf 1 A(config-router)# overflow database external 10 3
Verification	On Router B, configure 11 static routes (192.100.1.0/24 to 192.100.11.0/24). On Router A, verify that only 10 static routes are loaded.
A	A# show ip route ospf O E2 192.100.1.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.2.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.3.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.4.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.5.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.6.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.7.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.8.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.9.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1 O E2 192.100.10.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1

Common Errors

- The OSPF adjacency is abnormal because the maximum number of LSAs is inconsistent on different routers.

2.4.10 Modifying the Maximum Number of Concurrent Neighbors

Configuration Effect

- Control the maximum number of concurrent neighbors on the OSPF process to ease the pressure on the device.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Configuring the Maximum Number of Concurrent Neighbors on the OSPF Process

- (Optional) This configuration is recommended if you wish to set up the OSPF adjacencymore quickly when a router is connected with a lot of other routers.
- This configuration is performed on a core router.

Verification

- Run the **show ip ospf neighbor** command to display the number of neighbors that are concurrently interacting with the OSPF process.

Related Commands

Configuring the Maximum Number of Concurrent Neighbors on the Current Process

Command	**max-concurrent-dd***number*
Parameter Description	*number:* Specifies the maximum number of neighbors that are concurrently interacting with the OSPF process. The value ranges from 1 to 65,535.
Command Mode	OSPF routing process configuration mode
Usage Guide	When the performance of a router is affected because the router exchanges data with multiple neighbors, you can configure this command to restrict the maximum of neighbors with which one OSPF process can concurrently initiates or accepts interaction.

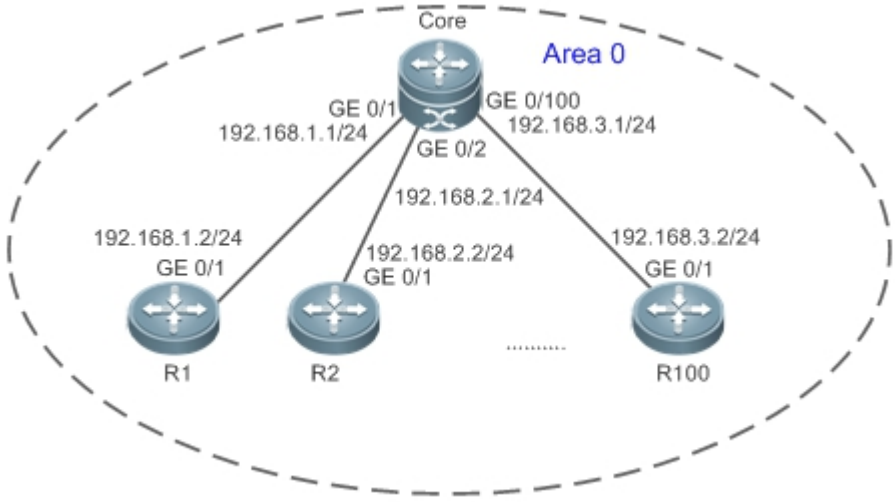
Configuring the Maximum Number of Concurrent Neighbors on All Processes

Command	**router ospf max-concurrent-dd***number*
Parameter Description	*number:* Specifies the maximum number of neighbors that are concurrently interacting with the OSPF process. The value ranges from 1 to 65,535.
Command Mode	Global configuration mode
Usage Guide	When the performance of a router is affected because the router exchanges data with multiple neighbors, you can configure this command to restrict the maximum of neighbors with which all OSPF processes can concurrently initiate or accept interaction.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Configuring the Maximum Number of Concurrent Neighbors on the OSPF Process

Scenario Figure 2- 18	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On the router Core, set the maximum number of concurrent neighbors to 4.
Core	Core# configure terminal Core(config)# router ospf max-concurrent-dd 4
Verification	On therouter Core, check the neighbor status and verify that at most eight neighbors concurrently interact with the OSPF process.

2.4.11 Disabling Source Address Verification

Configuration Effect

- The unicast routing service can be provided even if the interface IP addresses of neighbor routers are not in the same network segment.

Notes

- The OSPF basic functions must be configured.
- Source address verification cannot be disabled on a broadcast or NBMA network.

Configuration Steps

Disabling Source Address Verification

- (Optional) This configuration is mandatory if an adjacency should be set up between routers with interface IP addresses in different network segments.
- This configuration is performed on routers with interface IP addresses in different network segments.

Verification

- An adjacency can be set up between routers in different network segments.

Related Commands

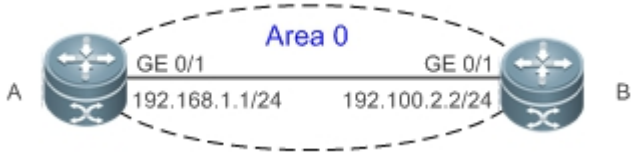
Disabling Source Address Verification

Command	**ip ospf source-check-ignore**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Generally, the source address of a packet received by OSPF is in the same network segment as the receiving interface. The addresses at both ends of a P2P link are configured separately and are not necessarily in the same network segment. In this scenario, as the peer address information will be notified during the P2P link negotiation process, OSPF checks whether the source address of the packet is the address advertised by the peer during negotiation. If not, OSPF determines that the packet is invalid and discards this packet. In particular, OSPF does not verify the address of an unnumbered interface. In some scenarios, the source address may not meet the preceding requirement, and therefore OSPF address verification fails. For example, the negotiated peer address cannot be obtained on a P2P link. In this scenario, source address verification must be disabled to ensure that the OSPF adjacency can be properly set up.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Disabling Source Address Verification

Scenario Figure 2- 19	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Set the network types of interfaces on all routers to P2P. ● Disable source address verification on all routers.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip ospf network point-to-point A(config-if-GigabitEthernet 0/1)# ip ospf source-check-ignore

B	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip ospf network point-to-point B(config-if-GigabitEthernet 0/1)# ip ospf source-check-ignore
Verification	On Router A, verify that the OSPF neighbor information is correct.
A	A# show ip ospfneighbor OSPF process 1, 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.100.2.2 1 Full/- 00:00:34 192.100.2.2 GigabitEthernet 0/1

2.4.12 Disabling MTU Verification

Configuration Effect

- The unicast routing service can be provided even if the MTUs of interfaces on neighbor routers are different.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Disabling MTU Verification

- (Optional) MTU verification is disabled by default. You are advised to retain the default configuration.
- This configuration is performed on two routers with different interface MTUs.

Verification

The adjacency can be set up between routers with different MTUs.

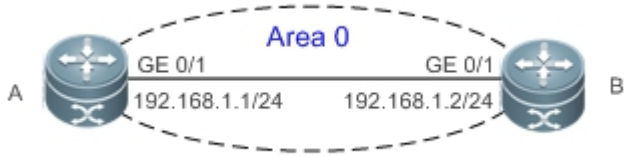
Related Commands

Disabling MTU Verification

Command	**ip ospf mtu-ignore**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	On receiving the database description packet, OSPF checks whether the MTU of the interface on the neighbor is the same as the MTU of its own interface. If the interface MTU specified in the received database description packet is greater than the MTU of the local interface, the adjacency cannot be set up. To resolve this problem, you can disable MTU verification.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 20	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Configure different MTUs for interfaces on two routers. ● Disable MTU verification on all routers. (By default, the function of disabling MTU verification is enabled.)
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip mtu 1400 A(config-if-GigabitEthernet 0/1)# ip ospf mtu-ignore
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip mtu 1600 B(config-if-GigabitEthernet 0/1)# ip ospf mtu-ignore
Verification	● On Router A, verify that the OSPF neighbor information is correct.
A	A# show ip ospfneighbor OSPF process 1, 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.2 1 Full/DR 00:00:34 192.168.1.2 GigabitEthernet 0/1

2.4.13 Enabling Two-Way Maintenance

Configuration Effect

- Non-Hello packets can also be used to maintain the adjacency.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Enabling Two-Way Maintenance

- (Optional) This function is enabled by default. You are advised to retain the default configuration.
- This configuration is performed on all routers.

Verification

Non-Hello packets can also be used to maintain the adjacency.

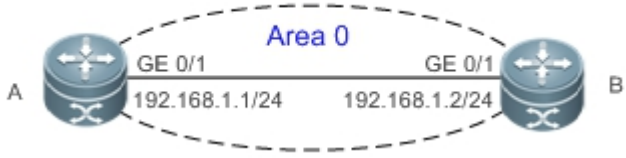
Related Commands

Enabling Two-Way Maintenance

Command	**two-way-maintain**
Parameter Description	N/A
Command Mode	OSPF routing process configuration mode
Usage Guide	On a large network, a lot of packets may be sent or received, occupying too much CPU and memory. As a result, some packets are delayed or discarded. If the processing time of Hello packets exceeds the dead interval, the adjacency will be destroyed due to timeout.If the two-way maintenance function is enabled, in addition to the Hello packets, the DD, LSU, LSR, and LSAck packets can also be used to maintain the bidirectional communication between neighbors when a large number of packets exist on the network. This prevents termination of the adjacency caused by delayed or discarded Hello packets.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 21	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, enable the two-way maintenance function. (This function is enabled by default.)
A	A# configure terminal A(config)#routerospf 1 A(config-router)#two-way-maintain
Verification	When the adjacency is being set up, Router A checks the neighbor dead interval and updates the dead interval without waiting for Router B to send a Hello packet.
A	A# show ip ospfneighbor OSPF process 1, 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Address Interface 192.168.1.2 1 Full/BDR 00:00:40 192.168.1.2 GigabitEthernet 0/1

2.4.14 Enabling GR

Configuration Effect

- When a distributed router switches services from the active board to the standby board, data forwarding continues and is not interrupted.
- When the OSPF process is being restarted, data forwarding continues and is not interrupted.

Notes

- The OSPF basic functions must be configured.
- The neighbor router must support the GR helper function.
- The grace period cannot be shorter than the neighbor dead time of the neighbor router.

Configuration Steps

Configuring the OSPF GR Function

- (Optional) This function is enabled by default. You are advised to retain the default configuration.
- This configuration is performed on all routers.

Configuring the OSPF GR Helper Function

- (Optional) This function is enabled by default. You are advised to retain the default configuration.
- This configuration is performed on all routers.

Verification

- When a distributed router switches services from the active board to the standby board, data forwarding continues and is not interrupted.
- When the OSPF process is being restarted, data forwarding continues and is not interrupted.

Related Commands

Configuring the OSPF GR Function

Command	**graceful-restart** [**grace-period** *grace-period* | **inconsistent-lsa-checking**]
Parameter Description	**grace-period** *grace-period*: Indicates the grace period, which is the maximum time from occurrence of an OSPF failure to completion of the OSPF GR. The value of the graceperiod varies from 1s to 1800s. The default value is 120s. **inconsistent-lsa-checking**: Enables topological change detection. If any topological change is detected, OSPF exits the GR process to complete convergence.After GR is enabled, topological change detection is enabled by default.
Command Mode	OSPF routing process configuration mode
Usage Guide	The GR function is configured based on the OSPF process. You can configure different parameters for different OSPF processes based on the actual conditions. This command is used to configure the GR restarter capability of a device. The grace period is the maximum time of the entire GR process, during which link status is rebuilt so that the original state of the OSPF process is restored. After the grace period expires, OSPF exits the GR state and performs common OSPF operations. Run the**graceful-restart** command to set the grace period to 120s. The **graceful-restart grace-period** command allows

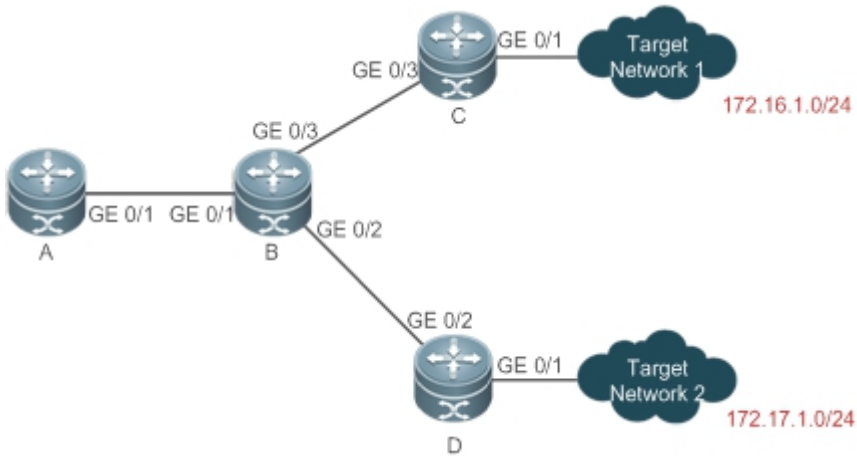
	you to modify the grace period explicitly. The precondition for successful execution of GR and uninterrupted forwarding is that the topology remains stable.If the topology changes, OSPF quickly converges without waiting for further execution of GR, thus avoiding long-time forwarding black-hole. Disabling topology detection: If OSPF cannot converge in time when thetopology changes during the hot standby process, forwarding black-hole may appear in a long time. Enabling topology detection: Forwarding may be interrupted when topology detection is enabled, but the interruption time is far shorter than that when topology detection is disabled. In most cases, it is recommended that topology detection be enabled. In special scenarios, topology detection can be disabled if the topology changes after the hot standby process, but it can be ensured that the forwarding black-hole will not appearin a long time. This can minimize the forwarding interruption time during the hot standby process. If the Fast Hello function is enabled, the GR function cannot be enabled.

Configuring the OSPF GR Helper Function

Command	**graceful-restart helper** { **disable | strict-lsa-checking | internal-lsa-checking**}
Parameter Description	**disable**: Prohibits a device from acting as a GR helper for another device. **strict-lsa-checking**: Indicates that changes in Type 1 to Type 5 and Type 7 LSAs will be checked during the period that the device acts as a GR helper to determine whether the network changes. If the network changes, the device will stop acting as the GR helper. **internal-lsa-checking:** Indicates that changes in Type 1 to Type 3 LSAs will be checked during the period that the device acts as a GR helper to determine whether the network changes. If the network changes, the device will stop acting as the GR helper.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command is used to configure the GR helper capability of a router. When a neighbor router implements GR, it sends a Grace-LSA to notify all neighbor routers. If the GR helper function is enabled on the local router, the local router becomes the GR helper on receiving the Grace-LSA, and helps the neighbor to complete GR. The **disable** option indicates that GR helper is not provided for any device that implements GR. After a device becomes the GR helper, the network changes are not detected by default. If any change takes place on the network, the network topology converges after GR is completed. If you wish that network changes can be quickly detected during the GR process, you can configure **strict-lsa-checking** to check Type 1 to 5 and Type 7 LSAs that indicate the network information or **internal-lsa-checking** to check Type 1 to 3 LSAs that indicate internal routes of the AS domain. When the network scale is large, it is recommended that you disable the LSA checking options (**strict-lsa-checking** and **internal-lsa-checking**) because regional network changes may trigger termination of GR and consequently reduce the convergence of the entire network.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 22	 **Remarks**: The interface IP addresses are as follows: A: GE 0/1 192.168.1.1 B: GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 GE 0/3 192.168.3.1 C: GE 0/1 192.168.4.2 GE 0/3 192.168.3.2 D: GE 0/1 192.168.5.2 GE 0/2 192.168.2.2
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, Router C, and Router D, enable the GR helper function. (This function is enabled by default.) ● On Router B, enable the GR function.
B	B# configure terminal B(config)# router ospf1 B(config-router)# graceful-restart
Verification	● Trigger a hot standby switchover on Router B, and verify that the routing tables of destination networks 1 and 2 remain unchanged on Router A during the switchover. ● Trigger a hot standby switchover on Router B, ping destination network 1 from Router A, and verify that data forwarding is not interrupted during the switchover.

Common Errors

- Traffic forwarding is interrupted during the GR process because the configured grace period is shorter than the neighbor dead time of the neighbor router.

2.4.15 Configuring the Network Management Function

Configuration Effect

- Use the network management software to manage OSPF parameters and monitor the OSPF running status.

Notes

- The OSPF basic functions must be configured.
- You must enable the MIB function of the SNMP-Server before enabling the OSPF MIB function.

- You must enable the Trap function of the SNMP-Server before enabling the OSPF Trap function.
- You must enable the logging function of the device before outputting the OSPF logs.

Configuration Steps

Binding the MIB with the OSPF Process

- (Optional) This configuration is required if you want to use the network management software to manage parameters of a specified OSPF process.
- This configuration is performed on all routers.

Enabling the Trap Function

- (Optional) This configuration is required if you want to use the network management software to monitor the OSPF running status.
- This configuration is performed on all routers.

Configuring the Logging Function

- (Optional) This function is enabled by default. You are advised to retain the default configuration. If you want to reduce the log output, disable this function.
- This configuration is performed on all routers.

Verification

- Use the network management software to manage the OSPF parameters.
- Use the network management software to monitor the OSPF running status.

Related Commands

Binding the MIB with the OSPF Process

Command	**enable mib-binding**
Parameter Description	N/A
Command Mode	OSPF routing process configuration mode
Usage Guide	The OSPFv2 MIB does not have the OSPFv2 process information. Therefore, you must perform operations on a single OSPFv2 process through SNMP. By default, the OSPFv2 MIB is bound with the OSPFv2 process with the smallest process ID, and all user operations take effect on this process. If you wish to perform operations on a specified OSPFv2 through SNMP, run this command to bind the MIB with the process.

Enabling the Trap Function

Command	**enable traps**[**error** [**IfAuthFailure | IfConfigError | IfRxBadPacket | VirtIfAuthFailure | VirtIfConfigError | VirtIfRxBadPacket**] | **lsa** [**LsdbApproachOverflow | LsdbOverflow | MaxAgeLsa | OriginateLsa**] | **retransmit** [**IfTxRetransmit | VirtIfTxRetransmit**] | **state-change**[**IfStateChange | NbrRestartHelperStatusChange | NbrStateChange | NssaTranslatorStatusChange | RestartStatusChange | VirtIfStateChange | VirtNbrRestartHelperStatusChange| VirtNbrStateChange**]]

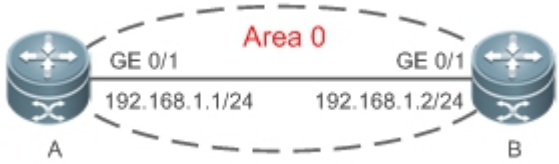
Parameter Description	**IfAuthFailure**: Indicates that an interface authentication failure occurs. **IfConfigError:** Indicates that an interface parameter configuration error occurs. **IfRxBadPacket:** Indicates that the interface receives a bad packet. **IfRxBadPacket:** Indicates that the interface receives a bad packet. **VirtIfAuthFailure:** Indicates that a virtual interface authentication failure occurs. **VirtIfConfigError:** Indicates that a virtual interface parameter configuration error occurs. **VirtIfRxBadPacket:** Indicates that the virtual interface receives a bad packet. **LsdbApproachOverflow:** Indicates that the number of external LSAs has reached 90% of the upper limit. **LsdbOverflow:** Indicates that the number of external LSAs has reached the upper limit. **MaxAgeLsa:** Indicates that the LSA aging timer expires. **OriginateLsa:** Indicates that a new LSA is generated. **IfTxRetransmit:** Indicates that a packet is retransmitted on the interface. **VirtIfTxRetransmit:** Indicates that a packet is retransmitted on the virtual interface. **IfStateChange:** Indicates that interface state changes. **NbrRestartHelperStatusChange:**Indicates that the state of the neighbor GR process changes. **NbrStateChange:** Indicates that the neighbor state changes. **NssaTranslatorStatusChange:** Indicates that the NSSA translation state changes. **RestartStatusChange:** Indicates that the GR state of the local device changes. **VirtIfStateChange:** Indicates that the virtual interface state changes. **VirtNbrRestartHelperStatusChange:** Indicates that the GR state of the virtual neighbor changes. **VirtNbrStateChange:** Indicates that the virtual neighbor state changes.
Command Mode	OSPF routing process configuration mode
Usage Guide	The function configured by this command is restricted by the **snmp-server** command. You can configure **snmp-server enable traps ospf** and then **enable traps** command before the corresponding OSPF traps can be correctly sent out. This command is not restricted by the MIB bound with the process. The trap function can be enabled concurrently for different processes.

Configuring the Logging Function

Command	**log-adj-changes**[**detail**]
Parameter Description	**detail**: Records all status change information.
Command Mode	OSPF routing process configuration mode
Usage Guide	N/A

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Scenario Figure 2- 23	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Bind the MIB with the OSPF process on Router A. ● Enable the trap function on Router A.
A	A# configure terminal A(config)# snmp-server host 192.168.2.2 traps version 2c public A(config)# snmp-server community public rw A(config)# snmp-server enable traps A(config)# router ospf 10 A(config-router)# enable mib-binding A(config-router)# enable traps
Verification	Use the MIB tool to read and set the OSPF parameters and display the OSPF running status.

Common Errors

Configurations on the SNMP-Server are incorrect. For example, the MIB or trap function is not enabled.

2.4.16 Modifying Protocol Control Parameters

Configuration Effect

Modify protocol control parameters to change the protocol running status.

Notes

- The OSPF basic functions must be configured.
- The neighbor dead time cannot be shorter than the Hello interval.

Configuration Steps

Configuring the Hello Interval

- (Optional) You are advised to retain the default configuration.
- This configuration is performed on routers at both end of a link.

Configuring the Dead Interval

- (Optional) You are advised to retain the default configuration. This configuration can be adjusted if you wish to accelerate OSPF convergence when a link fails.

- This configuration is performed on routers at both end of a link.

Configuring LSU Retransmission Interval

- (Optional) You are advised to adjust this configuration if a lot of routes exist in the user environment and network congestion is serious.

Configuring the LSA Generation Time

- (Optional) You are advised to retain the default configuration.

Configuring the LSA Group Refresh Time

- (Optional) You are advised to retain the default configuration. This configuration can be adjusted if a lot of routes exist in the user environment.
- This configuration is performed on an ASBR or ABR.

Configuring LSA Repeated Receiving Delay

- (Optional) You are advised to retain the default configuration.

Configuring the SPF Computation Delay

- (Optional) This configuration can be adjusted if network flapping frequently occurs.

Configuring the Inter-Area Route Computation Delay

- (Optional) You are advised to retain the default configuration.
- This configuration is performed on all routers.

Configuring the External Route Computation Delay

- (Optional) You are advised to retain the default configuration.
- This configuration is performed on all routers.

Verification

Run the **show ip ospf**and**show ip ospf neighbor** commands to display the protocol running parameters and status.

Related Commands

Configuring the Hello Interval

Command	**ip ospf hello-interval***seconds*
Parameter Description	*seconds*: Indicates the interval at which OSPF sends the Hello packet. The unit is second. The value ranges from 1 to 65,535.
Command Mode	Interface configuration mode
Usage Guide	The Hello interval is contained in the Hello packet. A shorter Hello interval indicates that OSPF can detect topological changes more quickly, but the network traffic increases. The Hello interval must be the same on all routers in the same network segment. If you want to manually modify the neighbor dead interval, ensure that the neighbor dead interval is longer than the Hello interval.

Configuring the Dead Interval

Command	**ip ospf dead-interval** *seconds*
Parameter Description	*seconds*: Indicates the time that the neighbor is declared lost. The unit is second. The value ranges from 0 to 2,147,483,647.
Command Mode	Interface configuration mode
Usage Guide	The OSPF dead interval is contained in the Hello packet. If OSPF does not receive a Hello packet from a neighbor within the dead interval, it declares that the neighbor is invalid and deletes this neighbor record form the neighbor list. By default, the dead interval is four times the Hello interval. If the Hello interval is modified, the dead interval is modified automatically. When using this command to manually modify the dead interval, pay attention to the following issues: 1. The dead interval cannot be shorter than the Hello interval. 2. The dead interval must be the same on all routers in the same network segment.

Configuring the LSU Transmission Delay

Command	**ip ospf transmit-delay***seconds*
Parameter Description	*seconds*: Indicates the LSU transmission delay on the OSPF interface. The unit is second. The value ranges from 0 to 65,535.
Command Mode	Interface configuration mode
Usage Guide	Before an LSU packet is transmitted, the Age fields in all LSAs in this packet will increase based on the amount specified by the **ip ospf transmit-delay** command. Considering the transmit and line propagation delays on the interface, you need to set the LSU transmission delay to a greater value for a low-speed line or interface. The LSU transmission delay of a virtual link is defined by the **transmit-delay** parameter in the **area virtual-link** command. If the value of the Age field of an LSA reaches 3600, the packet will be retransmitted or a retransmission will be requested. If the LSA is not updated in time, the expired LSA will be deleted from the LSDB.

Configuring LSU Retransmission Interval

Command	**ip ospf retransmit-interval***seconds*
Parameter Description	*seconds*: Indicates the LSU retransmission interval. The unit is second. The value ranges from 1 to 65,535. This interval must be longer than the round-trip transmission delay of data packets between two neighbors.
Command Mode	Interface configuration mode
Usage Guide	After a router finishes sending an LSU packet, this packet is still kept in the transmit buffer queue. If an acknowledgment from the neighbor is not received within the time defined by the **ip ospf retransmit-interval** command, the router retransmits the LSU packet. The retransmission delay can be set to a greater value on a serial line or virtual link to prevent unnecessary retransmission. The LSU retransmission delay of a virtual link is defined by the **retransmit-interval** parameter in the **area virtual-link** command.

Configuring the LSA Generation Time

Command	**timers throttle lsa all** *delay-time hold-time max-wait-time*
Parameter	*delay-time*: Indicates the minimum delay for LSA generation. The first LSA in the database is always generated instantly.

Description	The value ranges from 0 to 600,000. The unit is ms. *hold-time*: Indicates the minimum interval between the first LSA update and the second LSA update. The value ranges from 1 to 600,000. The unit is ms. *max-wait-time*: Indicates the maximum interval between two LSA updates when the LSA is updated continuously. This interval is also used to determine whether the LSA is updated continuously. The value ranges from 1 to 600,000. The unit is ms.
Command Mode	OSPF routing process configuration mode
Usage Guide	If a high convergence requirement is raised when a link changes, you can set **delay-time** to a smaller value. You can also appropriately increase values of the preceding parameters to reduce the CPU usage. When configuring this command, the value of **hold-time** cannot be smaller than the value of **delay-time,** and the value of **max-wait-time** cannot be smaller than the value of **hold-time**.

Configuring the LSA Group Refresh Time

Command	**timers pacinglsa-group** *seconds*
Parameter Description	*seconds*: Indicates the LSA group pacing interval. The value ranges from 10 to 1,800. The unit is second.
Command Mode	OSPF routing process configuration mode
Usage Guide	Every LSA has a time to live (LSA age). When the LSA age reaches 1800s, a refreshment is needed to prevent LSAs from being cleared because their ages reaching the maximum. If LSA update and aging computation are performed for every LSA, the device will consume a lot of CPU resources. In order to use CPU resources effectively, you can refresh LSAs by group on the device. The interval of group refreshment is called group pacing interval. The group refreshment operation is to organize the LSAs generated within a group pacing interval into a group and refresh the group as a whole. If the total number of LSAs does not change, a larger group pacing interval indicates that more LSAs need to be processed after timeout. To maintain the CPU stability, the number of LSAs processes upon each timeout cannot be too large. If the number of LSAs is large, you are advised to reduce the group pacing interval. For example, if there are 1000 LSAs in the database, you can reduce the pacing interval; if there are 40 to 100 LSAs, you can set the pacing interval to 10-20 minutes.

Configuring the LSA Group Refresh Interval

Command	**timers pacing lsa-transmit** *transmit-time transmit-count*
Parameter Description	*transmit-time*: Indicates the LSA group transmission interval. The value ranges from 10 to 1,000. The unit is ms. *transmit-count*: Indicates the number of LS-UPD packets in a group. The value ranges from 1 to 200.
Command Mode	OSPF routing process configuration mode
Usage Guide	If the number of LSAs is large and the device load is heavy in an environment, properly configuring **transimit-time** and **transimit-count** can limit the number of LS-UPD packets flooded on a network. If the CPU usage is not high and the network bandwidth load is not heavy, reducing the value of **transimit-time** and increasing the value of **transimit-count** can accelerate the environment convergence.

Configuring LSA Repeated Receiving Delay

Command	**timers lsa arrival** *arrival-time*
Parameter	*arrival-time*: Indicates the delay after which the same LSA is received. The value ranges from 0 to 600,000. The unit is ms.

Description	
Command Mode	OSPF routing process configuration mode
Usage Guide	No processing is performed if the same LSA is received within the specified time.

Configuring the Inter-Area Route Computation Delay

Command	**timers throttle route inter-area** *ia-delay*
Parameter Description	*ia-delay*: Indicates the inter-area route computation delay. The unit is ms. The value ranges from 0 to 600,000.
Command Mode	OSPF routing process configuration mode
Usage Guide	This delay cannot be modified if strict requirements are raised for the network convergence time.

Configuring the External Route Computation Delay

Command	**timers throttle route ase** *ase-delay*
Parameter Description	*ase-delay*: Indicates the external route computation delay. The unit is ms. The value ranges from 0 to 600,000.
Command Mode	OSPF routing process configuration mode
Usage Guide	This delay cannot be modified if strict requirements are raised for the network convergence time.

Configuring the SPF Computation Delay

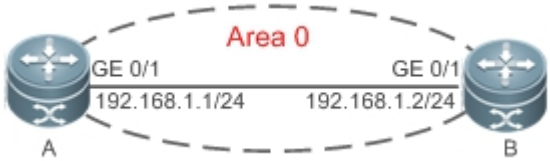
Command	**timers throttle spf** *spf-delay spf-holdtime spf-max-waittime*
Parameter Description	*spf-delay*: Indicates the SPF computation delay. The unit is ms. The value ranges from 1 to 600,000. When detecting a topological change, the OSPF routing process triggers the SPF computation at least after **spf-delay** elapses. *spf-holdtime*: Indicates the minimum interval between two SPF computations. The unit is ms. The value ranges from 1 to 600,000. *spf-max-waittime*: Indicates the maximum interval between two SPF computations. The unit is ms. The value ranges from 1 to 600,000. *number*: indicates the metric of the summarized route.
Command Mode	OSPF routing process configuration mode
Usage Guide	**spf-delay** indicates the minimum time between the occurrence of the topological change and the start of SPF computation. **spf-holdtime** indicates the minimum interval between the first SPF computation and the second SPF computation. After that, the interval between two SPF computations must be at least twice of the previous interval. When the interval reaches **spf-max-waittime**, the interval cannot increase again. If the interval between two SPF computations already exceeds the required minimum value, the interval is computed by starting from **spf-holdtime**. You can set **spf-dela**y and **spf-holdtime** to smaller values to accelerate topology convergence, and set **spf-max-waittime** to a larger value to reduce SPF computation. Flexible settings can be used based on stability of the network topology. Compared with the **timers spf** command, this command supports more flexible settings to accelerate the convergence speed of SPF computation and further reduce the system resources consumed by SPF computation when the topology continuously changes. Therefore, you are advised to use the **timers throttle spf** command for configuration.

	1. The value of **spf-holdtime** cannot be smaller than the value of **spf-delay**; otherwise, **spf-holdtime** will be automatically set to the value of **spf-delay**. 2. The value of **spf-max-waittime** cannot be smaller than the value of **spf-holdtime**; otherwise, **spf-max-waittime** will be automatically set to the value of **spf-holdtime**. 3. The configurations of **timers throttle spf** and **timers spf** are mutually overwritten. 4. When both **timers throttle spf** and **timers spf** are not configured, the default values of **timers throttle spf** prevail.

Configuration Example

The following configuration examples assume that the OSPF basic functions have been configured. For details about the OSPF basic functions, see section 2.4.1 "Configuring OSPF Basic Functions."

Configuring the Hello Interval and Dead Interval

Scenario Figure 2- 24	
Configuration Steps	● Configure the interface IP addresses on all routers. (Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Configure the Hello interval and dead interval on all routers.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip ospf hello-interval 15 A(config-if-GigabitEthernet 0/1)# ip ospf dead-interval 50
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip ospf hello-interval 15 A(config-if-GigabitEthernet 0/1)# ip ospf dead-interval 50
Verification	Check the interface parameters on Router A. Verify that the Hello interval is 10s and the dead interval is 50s.

A	A# show ip ospf interface GigabitEthernet 0/1 is up, line protocol is up Internet Address 192.168.1.1/24, Ifindex 2, Area 0.0.0.0, MTU 1500 Matching network config: 192.168.1.0/24 Process ID 1, Router ID 192.168.1.2, Network Type POINTOMULTIPOINT, Cost: 1 Transmit Delay is 1 sec, State Point-To-Point Timer intervals configured, Hello 15, Dead 50, Wait 40, Retransmit 5 Hello due in 00:00:02 Neighbor Count is 1, Adjacent neighbor count is 0 Crypt Sequence Number is 4787 Hello received 465 sent 466, DD received 8 sent 8 LS-Req received 2 sent 2, LS-Upd received 8 sent 21 LS-Ack received 14 sent 7, Discarded 3

Common Errors

- The configured neighbor dead time is shorter than the Hello interval.

2.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears and resets an OSPF process.	**clear ip ospf** [*process-id*] **process**

Displaying

Description	Command
Displays the OSPF process configurations.	**show ip ospf** [*process-id*]
Displays the OSPF internal routing table, including routes to ABRs and ASBRs.	**show ip ospf**[*process-id*] **border-routers**
Displays information about the OSPF LSDB.	**show ip ospf** [*process-id area-id*] **database** [{ **asbr-summary** | **external** | **network** | **nssa-external** | **opaque-area** | **opaque-as** | **opaque-link** | **router** | **summary** }][{ **adv-router** *ip-address*| **self-originate** } |*link-state-id* | **brief**][**database-summary** | **max-age** | **detail**]
Displays OSPF-enabled interfaces.	**show ip ospf** [*process-id*] **interface** [*interface-type interface-number* | **brief**]
Displays the OSPF neighbor list.	**show ip ospf** [*process-id*] **neighbor** [**detail**] [*interface-typeinterface-number*] [*neighbor-id*]
Displays the OSPF routing table.	**show ip ospf** [*process-id*] **route**[**count**]
Displays the number of times SPT is computed in the OSPF area.	**show ip ospf** [*process-id*] **spf**
Displays the summarized route of OSPF redistributed routes.	**show ip ospf**[*process-id*] **summary-address**

Description	Command
Displays OSPF virtual links.	**show ip ospf** [*process-id*] **virtual-links** [*ip-address*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs OSPF events.	**debug ip ospf events** [**abr**|**asbr**|**lsa**|**nssa**|**os**|**restart**| **router**|**slink**| **vlink**]
Debugs OSPF interfaces.	**debug ip ospf ifsm** [**events**|**status**|**timers**]
Debugs OSPF neighbors.	**debug ip ospf nfsm** [**events** | **status** | **timers**]
Debugs the OSPF NSM.	**debug ip ospf nsm** [**interface** | **redistribute** | **route**]
Debugs OSPF LSAs.	**debug ip ospf lsa** [**flooding** | **generate** | **install** | **maxage** | **refresh**]
Debugs OSPF packets.	**debug ip ospf packet** [**dd**|**detail**|**hello**|**ls-ack**|**ls-request**|**ls-update**|**recv**|**send**]
Debugs OSPF routes.	**debug ip ospf route** [**ase** | **ia** | **install** | **spf** | **time**]

3 Configuring OSPFv3

3.1 Overview

Open Shortest Path First (OSPF) is an Interior Gateway Protocol (IGP) that is used within the Autonomous System (AS) to allow routers to obtain a route to a remote network.

OSPF Version 2 (OSPFv2) is applicable to IPv4, and OSPF Version 3 (OSPFv3) is applicable to IPv6. The protocol running mechanism and most configurations are the same.

OSPF has the following characteristics:

- Wide scope of application: OSPF is applicable to a larger-scale network that supports hundreds of routers.
- Fast convergence: Once the network topology changes, notifications can be quickly sent between routers to update routes.
- No self-loop: Only the link status information is synchronized between routers. Each router computes routes independently, and a self-loop will not occur.
- Area division: A large routing domain is divided into multiple small areas to save system resources and network bandwidth and ensure stability and reliability of routes.
- Route classification: Routes are classified into several types to support flexible control.
- Equivalent routes: OSPF supports equivalent routes.
- Authentication: OSPF supports packet authentication to ensure security of protocol interaction.
- Multicast transmission: Protocol packets are sent using the multicast address to avoid interfering with irrelevant entities and save system resources.

In this chapter, the term "router" refers to any network device that supports the routing function. These network devices can be L3 switches, routers, or firewall.

Unless otherwise specified, "OSPF" in the following descriptions refers to OSPFv3.

Protocols and Standards

RFC2740	This document describes the modifications to OSPF to support version 6 of the Internet Protocol (IPv6).
draft-ietf-ospf-ospfv3-graceful-restart	This document describes the OSPFv3 graceful restart. The OSPFv3 graceful restart is identical to OSPFv2 except for the differences described in this document. These differences include the format of the grace Link State Advertisements (LSA) and other considerations.
draft-ietf-ospf-ospfv3-mib-11	This memo defines a portion of the Management Information Base (MIB) for use with network management protocols in IPv6-based internets. In particular, it defines objects for managing the Open Shortest Path First Routing Protocol for IPv6.

3.2 Applications

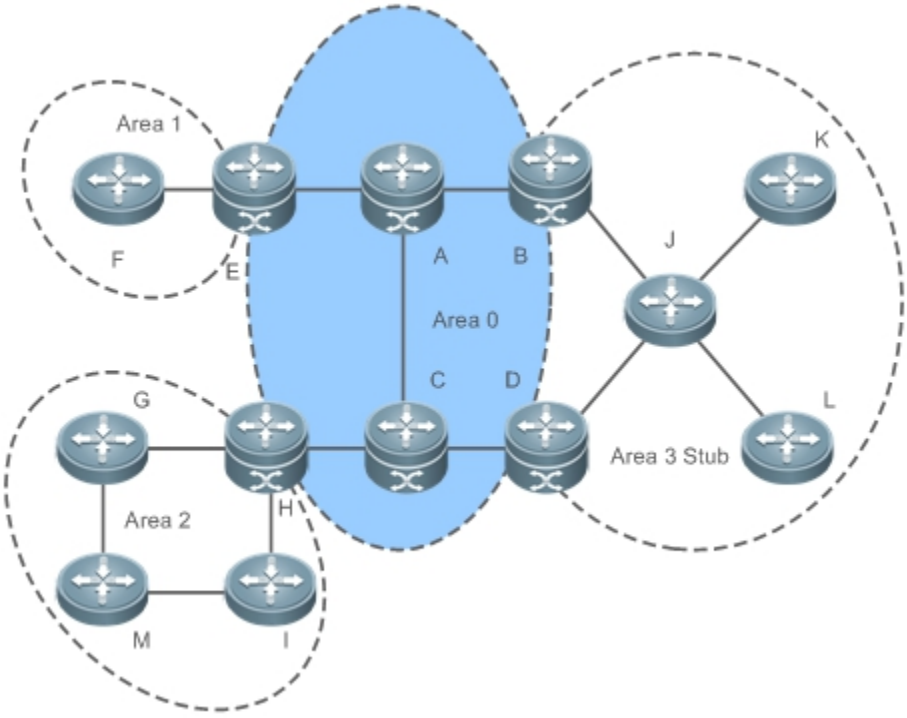
Application	Description
Intra-Domain Interworking	OSPF runs within the AS, which is divided into several areas.
Inter-Domain Interworking	Several ASs are interconnected. OSPF runs within each AS, and BGP runs between ASs.

3.2.1 Intra-Domain Interworking

Scenario

OSPF runs within the AS. If the number of routers exceeds 40, it is recommended that the AS be divided into several areas. Generally, high-end devices featuring reliable performance and fast processing speed are deployed in a backbone area, and low-end or medium-range devices with relatively lower performance can be deployed in a normal area. All normal areas must be connected to the backbone area. It is recommended that a normal area located on the stub be configured as a stub area. As shown in Figure 3- 1, the network is divided into four areas. Communication between these areas must go through the backbone area, that is, area 0.

Figure 3- 1 Division of the OSPF Areas



Remarks	A, B, C, D, E, and H are located in the backbone area, and are backbone routers. Area 3 is configured as a stub area.

Deployment

- OSPF runs on all routers within the AS to implement unicast routing.

3.3 Features

Basic Concepts

Routing Domain

All routers in an AS must be interconnected and use the same routing protocol. Therefore, an AS is also called a routing domain.

An AS on which OSPF runs is also called OSPF routing domain, or OSPF domain for short.

OSPF Process

OSPF supports multiple instances, and each instance corresponds to an OSPF process.

One or more OSPF processes can be started on a router. Each OSPF process runs OSPF independently, and the processes are mutually isolated.

An OSPF packet header contains the Instance ID field, and multiple OSPF instances can run concurrently on a single link. The process ID is valid only on the local device.

RouterID

The router ID uniquely identifies a router in an OSPF domain. Router IDs of any two routers cannot be the same.

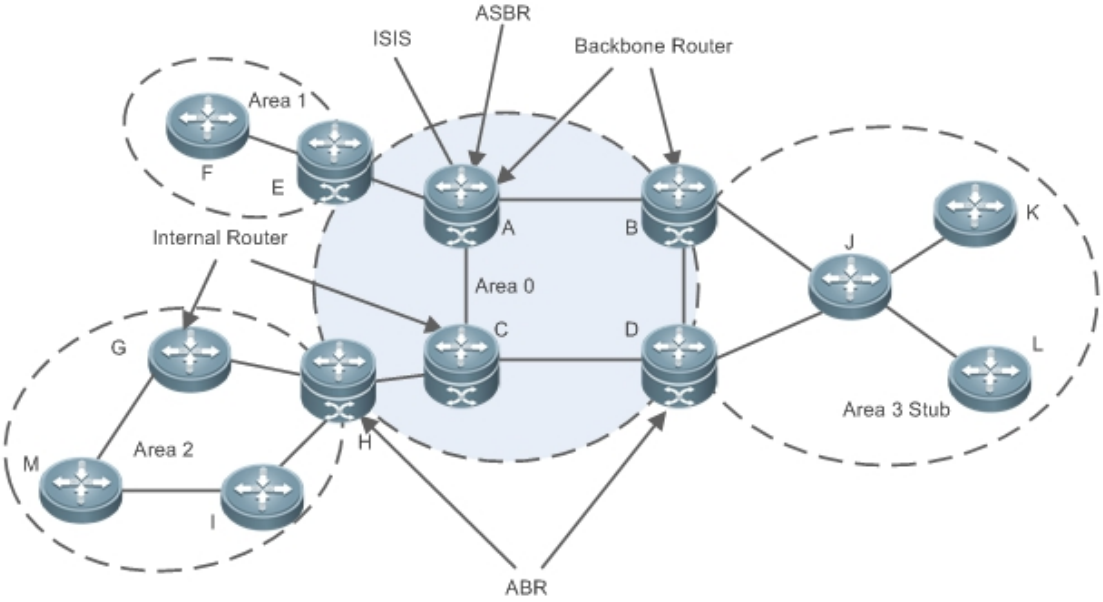
If multiple OSPF processes exist on a router, each OSPF process uses one router ID. Router IDs of any two OSPF processes cannot be the same.

Area

OSPF supports multiple areas. An OSPF domain is divided into multiple areas to ease the computing pressure of a large-scale network. An area is a logical group of routers, and each group is identified by an area ID. The border between areas is a router. A router may belong to one area or multiple areas. One network segment (link) can belong to only one area, or each OSPF-enabled interface must belong to a specified area.

Area 0 is the backbone area, and other areas are normal areas. Normal areas must be directly connected to the backbone area.

Figure 3- 2Division of the OSPF Areas



OSPF Router

The following types of routers are defined in OSPF, and assigned with different responsibilities:

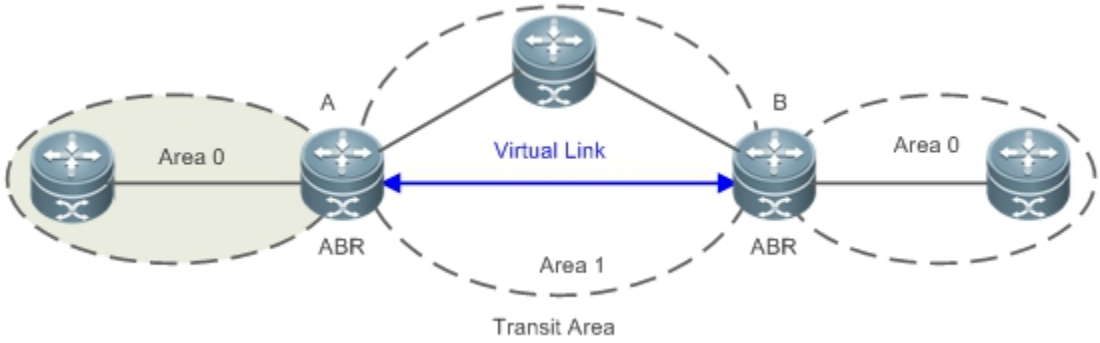
- Internal router
  All interface of an interval router belong to the same OSPF area. As shown in Figure 3- 2, A, C, F, G, I, M, J, K, and L are internal routers.
- Area border router (ABR)
  An ABR is used to connect the backbone area with a normal area. An ABR belongs to two or more areas, and one of the areas must be the backbone area. As shown in Figure 3- 2, B, D, E, and H are ABRs.
- Backbone router
  A backbone router has at least one interface that belongs to the backbone area. All ABRs and all routers in area 0 are backbone routers. As shown in Figure 3- 2, A, B, C, D, E, and H are backbone routers.

- AS boundary router (ASBR)
  An ASBR is used to exchange routing information with other ASs. An ASBR is not necessarily located on the border of an AS. It may be a router inside an area, or an ABR. As shown in Figure 3- 2, A is an ASBR.

Virtual Link

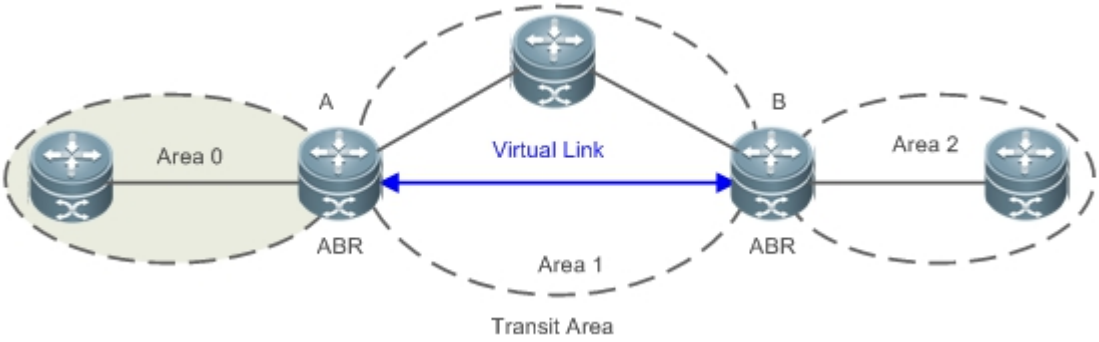
OSPF supports virtual links. A virtual link is a logical link that belongs to the backbone area. It is used to resolve the problems such as a discontinuous backbone area or a failure to directly connect a normal area to the backbone area on the physical network. A virtual link supports traversal of only one normal area, and this area is called transit area. Routers on both ends of a virtual link are ABRs.

Figure 3- 3 Discontinuous Backbone Area on the Physical Network



As shown in Figure 3- 3, a virtual link is set up between A and B to connect two separated parts of Area 0. Area 1 is a transit area, and A and B are ABRs of Area 1.

Figure 3- 4 Failure to Directly Connect a Normal Area to the Backbone Area on the Physical Network



As shown in Figure 3- 4, a virtual link is set up between A and B to extend Area 0 to B so that Area 0 can be directly connected to Area 2 on B. Area 1 is a transit area, A is an ABR of Area 1, and B is an ABR of Area 0 and Area 2.

LSA

OSPF describes the routing information by means of Link State Advertisement (LSA).

LSA Type	Description
Router-LSA(Type1)	This LSA is originated by every router. It describes the link state and cost of the router, and is advertised only within the area where the originating router is located.
Network-LSA(Type2)	This LSA is originated by a designated router (DR). It describes the state of the current link, and is advertised only within the area where the DR is located.
Inter-Area-Prefix-LSA(Type3)	This LSA is originated by an ABR. It describes a route to another area, and is advertised to areas except totally stub areas.
Inter-Area-Router-LSA(Type4)	This LSA is originated by an ABR. It describes a route to an ASBR, and is advertised to areas except areas where the ASBR is located.

LSA Type	Description
AS-external-LSA(Type5)	This LSA is originated by an ABR. It describes a route to a destination outside the AS, and is advertised to all areas except the stub areas.
Link-LSA(Type8)	This LSA is originated by every router. It describes the link-local address and IPv6 prefix address of each link, and provides the link option that will be set in the Network-LSA. It advertised only on the current link.
Intra-Area-Prefix-LSA(Type9)	Every router or DR generates one or more Intra-Area-Prefix-LSAs, which are advertised in the area to which the router or DR belongs. ● The Intra-Area-Prefix-LSA generated by a router describes the IPv6 prefix address associated with the Route-LSA. ● The Intra-Area-Prefix-LSA generated by a DR describes the IPv6 prefix address associated with the Network-LSA.

Stub areas and totally stub areas are special forms of normal areas and help reduce the load of routers and enhance reliability of OSPF routes.

OSPF Packet

The following table lists the protocol packets used by OSPF. These OSPF packets are encapsulated in IP packets and transmitted in multicast or unicast mode.

Packet Type	Description
Hello	Hello packets are sent periodically to discover and maintain OSPF neighbor relationships.
Database Description (DD)	DD packets carry brief information about the local Link-State Database (LSDB) and are used to synchronize the LSDBs between OSPF neighbors.
Link State Request (LSR)	LSR packets are used to request the required LSAs from neighbors. LSR packets are sent only after DD packets are exchanged successfully between OSPF neighbors.
Link State Update (LSU)	LSU packets are used to send the required LSAs to peers.
Link State Acknowledgment (LSAck)	LSAck packets are used to acknowledge the received LSAs.

Overview

Feature	Description
Link-State Routing Protocols	Run OSPF on the router to obtain routes to different destinations on the network.
OSPF Route Management	Properly plan or optimize OSPF routes through manual configuration to implement management of OSPF routes.
Enhanced Security and Reliability	Use functions such as authentication to enhance security, stability, and reliability of OSPF.
Network Management Functions	Use functions such as the MIB and Syslog to facilitate OSPF management.

3.3.1 Link-State Routing Protocols

OSPF is a type of link-state routing protocols. Its working process is as follows:

- Neighbor discovery → Bidirectional communication
  An OSPF neighbor relationship is set up between adjacent routers, and bidirectional communication is maintained.
- Database synchronization → Full adjacency

A router uses LSAs to advertise all its link states. LSAs are exchanged between neighbors and the link state database (LSDB) is synchronized to achieve full adjacency.

- Shortest Path Tree (SPT) computation → Formation of a routing table
  The router computes the shortest path to each destination network based on the LSDB and forms an OSPF routing table.

Working Principle

Neighbor Discovery → Bidirectional Communication

Routers send Hello packets through all OSPF-enabled interfaces (or virtual links). If Hello packets can be exchanged between two routers, and parameters carried in the Hello packets can be successfully negotiated, the two routers become neighbors. Routers that are mutually neighbors find their own router IDs from Hello packets sent from neighbors, and bidirectional communication is set up.

A Hello packet includes, but is not limited to, the following information:

- Router ID of the originating router
- Area ID of the originating router interface (or virtual link)
- Instance ID of the originating router interface (or virtual link)
- Interface ID of the originating router interface (or virtual link)
- Priority of the originating router interface (used for DR/BDR election)
- Hello interval of the originating router interface (or virtual link)
- Neighbor dead interval of the originating router interface (or virtual link)
- IP addresses of the DR and Backup Designated Router (BDR)
- Router ID of the neighbor of the originating router

Database Synchronization → Full Adjacency

After bidirectional communication is set up between neighbor routers, the DD, LSR, LSU, and LSAck packets are used to exchange LSAs and set up the adjacency. The brief process is as follows:

- A router generates an LSA to describe all link states on the router.
- The LSA is exchanged between neighbors. When a router receives the LSA from its neighbor, it copies the LSA and saves the copy in the local LSDB, and then advertises the LSA to other neighbors.
- When the router and its neighbors obtain the same LSDB, full adjacency is achieved.

OSPF will be very quiet without changes in link costs or network addition or deletion. If any change takes place, the changed link states are advertised to quickly synchronize the LSDB.

SPT Computation → Formation of a Routing Table

After the complete LSDB is obtained from the router, the Dijkstra algorithm is run to generate an SPT from the local router to each destination network. The SPT records the destination networks, next-hop addresses, and costs. OSPF generates a routing table based on the SPT.

If changes in link costs or network addition or deletion take place, the LSDB will be updated. The router again runs the Dijkstra algorithm, generates a new SPT, and updates the routing table.

The Dijkstra algorithm is used to find a shortest path from a vertex to other vertices in a weighted directed graph.

OSPF Network Types

A router does not necessarily need to exchange LSAs with every neighbor and set up an adjacency with every neighbor. To improve efficiency, OSPF classifies networks that use various link layer protocols into five types so that LSAs are exchanged in different ways to set up an adjacency:

- Broadcast

  Neighbors are discovered, and the DR and BDR are elected.

  The DR (or BDR) exchanges LSAs with all other routers to set up an adjacency. Except the DR and BDR, all other routers do not exchange LSAs with each other, and the adjacency is not set up.

  Ethernet and fiber distributed data interface (FDDI) belong to the broadcast network type by default.

- Non-broadcast multiple access (NBMA)

  Neighbors are manually configured, and the DR and BDR are elected.

  The DR (or BDR) exchanges LSAs with all other routers to set up an adjacency. Except the DR and BDR, all other routers do not exchange LSAs with each other, and the adjacency is not set up.

  X.25, frame relay, and ATM belong to NBMA networks by default.

- Point-to-point (P2P)

  Neighbors are automatically discovered, and the DR or BDR is not elected.

  LSAs are exchanged between routers at both ends of the link, and the adjacency is set up.

  PPP, HDLC, and LAPB belong to the P2P network type by default.

- Point-to-multipoint(P2MP)

  Neighbors are automatically discovered, and the DR or BDR is not elected.

  LSAs are exchanged between any two routers, and the adjacency is set up.

  Networks without any link layer protocol belong to the P2MP network type by default.

- P2MP broadcast

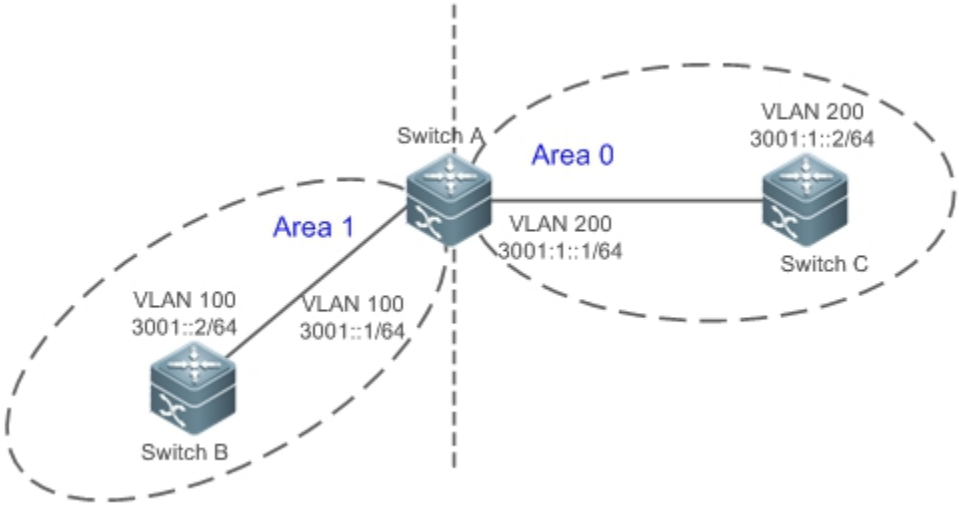
  Neighbors are manually configured, and the DR or BDR is not elected.

  LSAs are exchanged between any two routers, and the adjacency is set up.

  Networks without any link layer protocol belong to the P2MP network type by default.

OSPF Route Types

Figure 3- 5



Display the OSPF routes (marked in red) in the routing table of Router C.

```
C#show ipv6 route ospf
IPv6 routing table name is Default(0) global scope - 7 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
        I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
        O - OSPF intra area, OI - OSPF inter area,   OE1 - OSPF external type 1, OE2 - OSPF external type 2

        [*] - NOT in hardware forwarding table

L       ::1/128    via Loopback, local host
OI      3001::/64    [110/2] via FE80::21A:A9FF:FE15:4CB9, VLAN 200
C       3001:1::/64    via VLAN 200, directly connected
L       3001:1::2/128    via VLAN 200, local host
L       FE80::/10    via ::1, Null0
C       FE80::/64    via VLAN 200, directly connected
L       FE80::21A:A9FF:FE01:FB1F/128    via VLAN 200, local host
```

A mark is displayed in front of each OSPF route to indicate the type of the route. There are six types of OSPF routes:

- O: Intra-area route

  This type of route describes how to arrive at a destination network in the local area. The cost of this type of route is equal to the cost of the route from the local router to the destination network.

- OI: Inter-area route

  This type of route describes how to arrive at a destination network in another area. The cost of this type of route is equal to the cost of the route from the local router to the destination network.

- OE1: Type 1 external route

  This type of route describes how to arrive at a destination network outside the AS. The cost of this type of route is equal to the cost of the route from the local router to the ASBR plus the cost of the route from the ASBR to the destination network. This type of route does not exist on routers in the stub area.

- OE2: Type 2 external route

  This type of route describes how to arrive at a destination network outside the AS. The cost of this type of route is equal to the cost of the route from the ASBR to the destination network. This type of route does not exist on routers in the stub area.

Reliability of OE2 and ON2 routes is poor. OSPF believes that the cost of the route from the ASBR to a destination outside an AS is far greater than the cost of the route to the ASBR within the AS. Therefore, when the route cost is computed, only the cost of the route from the ASBR to a destination outside an AS is considered.

Related Configuration

Enabling OSPF

OSPF is disabled by default.

Run the **ipv6 router ospf** 1 command to create an OSPF process on the router.

Run the **ipv6 ospfarea** command to enable OSPF on an interface and specify the area ID.

Run the **area virtual-link** command to create a virtual link on the router. The virtual link can be treated as a logical interface.

Router ID

By default, the OSPF process elects the largest IPv4 address among the IPv4 addresses of all the loopback interfaces as the router ID. If the loopback interfaces configured with IPv4 addresses are not available, the OSPF process elects the largest IPv4 address among the IPv4 addresses of all the physical ports as the router ID.

Alternatively, you can run the **router-id** command to manually specify the router ID.

Protocol Control Parameters

Run the **ipv6 ospf hello-interval** command to modify the Hello interval on the interface. The default value is 10s (or 30s for NBMA networks).

Run the **ipv6 ospf dead-interval** command to modify the neighbor dead interval on the interface. The default value is four times the Hello interval.

Use the **poll-interval** parameter in the **ipv6 ospf neighbor** command to modify the neighbor polling interval on the NBMA interface. The default value is 120s.

Run the **ipv6 ospf transmit-delay** command to modify the LSU packet transmission delay on the interface. The default value is 1s.

Run the **ipv6 ospf retransmit-interval** command to modify the LSU packet retransmission interval on the interface. The default value is 5s.

Use the **hello-interval** parameter in the **area virtual-link** command to modify the Hello interval on the virtual link. The default value is 10s.

Use the **dead-interval** parameter in the **area virtual-link** command to modify the neighbor dead interval on the virtual link. The default value is four times the Hello interval.

Use the **transmit-delay** parameter in the **area virtual-link** command to modify the LSU packet transmission delay on the virtual link. The default value is 1s.

Use the **retransmit-interval** parameter in the **area virtual-link** command to modify the LSU packet retransmission interval on the virtual link. The default value is 5s.

Run the **timers throttle lsa all** command to modify parameters of the exponential backoff algorithm that generates LSAs. The default values of these parameters are 0 ms, 5000 ms, and 5000 ms.
Run the **timers pacing lsa-group** command to modify the LSA group update interval. The default value is 30s.
Run the **timers pacing lsa-transmit** command to modify the LS-UPD packet sending interval and the number of sent LS-UPD packets. The default values are 40 ms and 1.
Run the **timers lsa arrival** command to modify the delay after which the same LSA is received. The default value is 1000 ms.
Run the **timers throttle spf** command to modify the SPT computation delay, minimum interval between two SPT computations, and maximum interval between two SPT computations. The default values are 1000 ms, 5000 ms, and 10000 ms.

OSPF Network Types

By default, Ethernet and FDDI belong to the broadcast type, X.25, frame relay, and ATM belong to the NBMA type, and PPP, HDLC, and LAPB belong to the P2P type.
Run the **ipv6 ospf network** command to manually specify the network type of an interface.
Run the **ipv6 ospf neighbor** command to manually specify a neighbor. For the NBMA and P2MP non-broadcast types, you must manually specify neighbors.
Run the **ipv6 ospf priority** command to adjust the priorities of interfaces, which are used for DR/BDR election. The DR/BDR election is required for the broadcast and NBMA types. The router with the highest priority wins in the election, and the router with the priority of 0 does not participate in the election. The default value is 1.

3.3.2 OSPF Route Management

Properly plan or optimize OSPF routes through manual configuration to implement management of OSPF routes.

Working Principle

(Totally) Stub Area

The (totally) stub areas help reduce the protocol interaction load and the size of the routing table.

- If an appropriate area is configured as a (totally) stub area, advertisement of a large number of Type 5 and Type 3 LSAs can be avoided within the area.

Area	Type 1 and Type 2 LSAs	Type 3 LSA	Type 4 LSA	Type 5 LSA	Type 7 LSA
Non (totally) stub area	Allowed	Allowed	Allowed	Allowed	Not allowed
Stub area	Allowed	Allowed (containing one default route)	Not allowed	Not allowed	Not allowed
Totally stub area	Allowed	Only one default route is allowed.	Not allowed	Not allowed	Not allowed

The ABR uses Type 3 LSAs to advertise a default route to the (totally) stub area.

- If an area is appropriately configured as a (totally) stub area, a large number of OE1, OE2, and OI routes will not be added to the routing table of a router in the area.

Area	Routes Available in the Routing Table of a Router Inside the Area
Non (totally) stub area	O: a route to a destination network in the local area OI: a route to a destination network in another area OE1 or OE2: a route or default route to a destination network segment outside the AS (via any ASBR in the AS)

Area	Routes Available in the Routing Table of a Router Inside the Area
Stub area	O: a route to a destination network in the local area OI: a route or a default route to a destination network in another area
Totally stub area	O: a route to a destination network in the local area OI: a default route

Route Redistribution

Route redistribution refers to the process of introducing routes of other routing protocols, routes of other OSPF processes, static routes, and direct routes that exist on the device to an OSPF process so that these routes can be advertised to neighbors using Type 5 and Type 7 LSAs. A default route cannot be introduced during route redistribution.

Route redistribution is often used for interworking between ASs. You can configure route redistribution on an ASBR to advertise routes outside an AS to the interior of the AS, or routes inside an AS to the exterior of the AS.

Default Route Introduction

By configuring a command on an ASBR, you can introduce a default route to an OSPF process so that the route can be advertised to neighbors using Type 5 and Type 7 LSAs.

Default route introduction is often used for interworking between ASs. One default route is used to replace all the routes outside an AS.

Route Summarization

Route summarization is a process of summarizing routing information with the same prefix into one route, and advertising the summarized route (replacing a large number of individual routes) to neighbors. Route summarization helps reduce the protocol interaction load and the size of the routing table.

By default, the ABR advertises inter-area routing information by using Type3 LSAs within a network segment, and advertises redistributed routing information by using Type 5 and Type 7 LSAs.If continuous network segments exist, it is recommended that you configure route summarization.

Route Filtering

OSPF supports route filtering to ensure security and facilitate control when the routing information is being learned, exchanged, or used. Using configuration commands, you can configure route filtering for the following items:

- Interface: The interface is prevented from sending routing information (any LSAs) or exchanging routing information (any LSAs) with neighbors.
- Routing information outside an AS: Only the routing information that meets the filtering conditions can be redistributed to the OSPF process (Type 5 and Type 7 LSAs).
- LSAs received by a router: In the OSPF routing table, only the routes that are computed based on the LSAs meeting the filtering conditions can be advertised.

Route Cost

If redundancy links or devices exist on the network, multiple paths may exist from the local device to the destination network. OSPF selects the path with the minimum total cost to form an OSPF route. The total cost of a path is equal to the sum of the costs of individual links along the path.The total cost of a path can be minimized by modifying the costs of individual links along the path. In this way, OSPF selects this path to form a route.

Using configuration commands, you can modify the following link costs:

- Cost from an interface to a directly connected network segment and cost from the interface to a neighbor

- Cost from an ABR to the default network segment
- Cost from an ASBR to an external network segment and cost from the ASBR to the default network segment

Both the cost and the metric indicate the cost and are not differentiated from each other.

OSPF Administrative Distance

The administrative distance (AD) evaluates reliability of a route, and the value is an integer ranging from 0 to 255. A smaller AD value indicates that the route is more trustworthy. If multiples exist to the same destination, the route preferentially selects a route with a smaller AD value. The route with a greater AD value becomes a floating route, that is, a standby route of the optimum route.

By default, the route coming from one source corresponds to an AD value. The AD value is a local concept. Modifying the AD value affects route selection only on the current router.

Route Source	Directly-connected network	Static route	OSPF Route	RIP Route	Unreachable Route
Default AD	0	1	110	120	255

Related Configuration

Stub Area

By default, no stub or NSSA area is configured.

Run the **area stub** command to configure a specified area as a stub area.

A backbone area cannot be configured as a stub.

A transit area (with virtual links going through) cannot be configured as a stub.

An area containing an ASBR cannot be configured as a stub area.

Route Redistribution and Default Route Introduction

By default, routes are not redistributed and the default route is not introduced.

Run the **redistribute** command to configure route redistribution.

Run the **default-information originate** command to introduce a default route.

After configuring route redistribution and default route introduction, the router automatically becomes an ASBR.

Route Summarization

By default, routes are not summarized. If route summarization is configured, a discard route will be automatically added.

Run the **area range** command to summarize routes (Type 3 LSA) distributed between areas on the ABR.

Run the **summary-prefix** command to summarize redistributed routes (Type 5 and Type 7 LSAs) on the ASBR.

Route Filtering

By default, routes are not filtered.

Run the **passive-interface** command to configure a passive interface. Routing information (any LSAs) cannot be exchanged on a passive interface.

Use the **route-map** parameter in the **redistribute** command, or use the **distribute-list out** command to filter the external routing information of the AS on the ASBR. Only the routing information that meets the filtering conditions can be redistributed to the OSPF process (Type 5 LSAs).

Run the **distribute-list in** command to filter LSAs received by the router. In the OSPF routing table, only the routes that are computed based on the LSAs meeting the filtering conditions can be advertised.

Route Cost

- Cost from the interface to the directly-connected network segment (cost on the interface)
  The default value is the auto cost. Auto cost = Reference bandwidth/Interface bandwidth
  Run the **auto-cost reference-bandwidth** command to set the reference bandwidth of the auto cost. The default value is 100 Mbps.
  Run the **ipv6 ospf cost** command to manually set the cost of the interface. The configuration priority of this item is higher than that of the auto cost.
- Cost from the interface to a specified neighbor (that is, cost from the local device to a specified neighbor)
  The default value is the auto cost.
  Use the **cost** parameter in the **ipv6 ospf neighbor** command to modify the cost from the interface to a specified neighbor. The configuration priority of this item is higher than that of the cost of the interface.
  This configuration item is applicable only to P2MP-type interfaces.
- Cost from the ABR to the default network segment (that is, the cost of the default route that is automatically advertised by the ABR to the stub areas)
  The default value is 1.
  Run the **area default-cost** command to modify the cost of the default route that the ABR automatically advertise to the stub areas.
- Cost from the ASBR to an external network segment (that is, the metric of an external route)
  By default, the metric of other types of redistributed routes is 20, and the route type is Type 2 External.
  Run the **default-metric** command to modify the default metric of the external route.
  Use the **metric**,**metric-type**, and **route-map** parameters in the **redistribute** command to modify the metric and route type of the external route.
- Cost from the ASBR to the default network segment (that is, the metric of the default route that is manually introduced)
  By default, the metric is 1, and the route type is Type 2 External.
  Use the **metric**,**metric-type**, and **route-map** parameters in the **default-information originate** command to modify the metric and route type of the default route that is manually introduced.

OSPF Administrative Distance

By default, the OSPF AD is 110.

Run the **distance** command to set the AD of an OSPF route.

3.3.3 Enhanced Security and Reliability

Use functions such as authentication to enhance security, stability, and reliability of OSPF.

Working Principle

Authentication

OSPFv3 uses the authentication mechanism, that is, IP authentication header (AH) and IP Encapsulating Security Payload (ESP), provided by IPv6 to prevent unauthorized routers that access the network and hosts that forge OSPF packets to participate in OSPF routing. OSPF packets received on the OSPF interface (or at both ends of a virtual link) are authenticated. If authentication fails, the packets are discarded and the adjacency cannot be set up.

Enabling authentication can avoid learning unauthenticated or invalid routes, thus preventing advertising valid routes to unauthenticated devices. In the broadcast-type network, authentication also prevents unauthenticated devices from becoming designated devices, ensuring stability of the routing system and protecting the routing system against intrusions.

MTU Verification

On receiving a DD packet, OSPF checks whether the MTU of the neighbor interface is the same as the MTU of the local interface. If the MTU of the interface specified in the received DD packet is greater than the MTU of the interface that receives the packet, the adjacency cannot be set up. Disabling MTU verification can avoid this problem.

Two-Way Maintenance

OSPF routers periodically send Hello packets to each other to maintain the adjacency. On a large network, a lot of packets may be sent or received, occupying too much CPU and memory. As a result, some packets are delayed or discarded. If the processing time of Hello packets exceeds the dead interval, the adjacency will be destroyed.

If the two-way maintenance function is enabled, in addition to the Hello packets, the DD, LSU, LSR, and LSAck packets can also be used to maintain the bidirectional communication between neighbors, which makes the adjacency more stable.

Concurrent neighbor Interaction Restriction

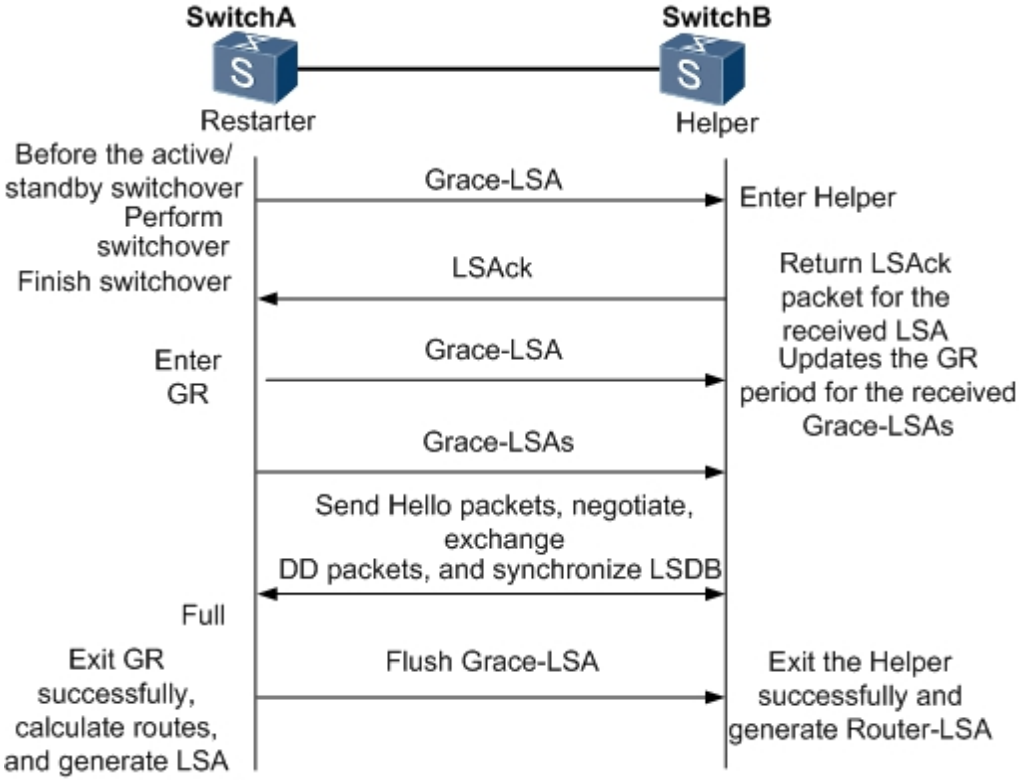
When a router simultaneously exchanges data with multiple neighbors, its performance may be affected. If the maximum number of neighbors that concurrently initiate or accept interaction with the OSPF process, the router can interact with neighbors by batches, which ensures data forwarding and other key services.

GR

The control and forwarding separated technology is widely used among routers. On a relatively stable network topology, when a GR-enabled router is restarted on the control plane, data forwarding can continue on the forwarding plane. In addition, actions (such as adjacency re-forming and route computation) performed on the control plane do not affect functions of the forwarding plane. In this way, service interruption caused by route flapping can be avoided, thus enhancing reliability of the entire network.

Currently, the GR function is used only during active/standby switchover and system upgrade.

Figure 3- 6 Normal OSPF GR Process



- The GR process requires collaboration between the restarter and the helper. The restarter is the router where GR occurs. The helper is a neighbor of the restarter.
- When entering or exiting the GR process, the restarter sends a Grace-LSA to the neighbor, notifying the neighbor to enter or exit the helper state.
- When the adjacency between the restarter and the helper reaches the Full state, the router can exit the GR process successfully.

Fast Hello

After a link fault occurs, it takes a period of time (about 40s) before OSPF can sense the death of the neighbor. Then, OSPF advertises the information and re-computes the SPT. During this period, traffic is interrupted.

- After the fast Hello function is enabled (that is, the neighbor dead interval is set to 1s), OSPF can sense the death of a neighbor within 1s once a link is faulty. This greatly accelerates route convergence and prevents traffic interruption.

Related Configuration

OSPF Packet Authentication

By default, authentication is disabled.

- Run the **area authentication** command to enable authentication in the entire area so that the authentication function takes effect on all interfaces in this area. If authentication is enabled in area 0, the function also takes effect on the virtual link.
- Run the **area encryption** command to enable encryption and authentication in the entire area so that the encryption**and** authentication functions take effect on all interfaces in this area. If encryptionand authentication are enabled in area 0, the functions also take effect on the virtual link.

- Run the **ipv6 ospf authentication** command to enable authentication on an interface. This configuration takes precedence over the area-based configuration.
- Run the **ipv6 ospf encryption** command to enable encryptionand authentication on an interface. This configuration takes precedence over the area-based configuration.
- Use the **authentication** parameter in the **area virtual-link** command to enable authentication at both ends of a virtual link. This configuration takes precedence over the area-based configuration.
- Use the **encryption** parameter in the **area virtual-link** command to enable encryptionand authentication at both ends of a virtual link. This configuration takes precedence over the area-based configuration.

MTU Verification

By default, MTU verification is disabled.

Run the **ipv6 ospf mtu-ignore** command to disable MTU verification on an interface.

Two-Way Maintenance

By default, bidirectional maintenance is enabled.

Run the **two-way-maintain** command to enable two-way maintenance.

Concurrent neighbor Interaction Restriction

Run the **max-concurrent-dd** command to modify the maximum number of neighbors that are concurrently interacting with the current OSPF process. The default value is 5.

Run the **ipv6 router ospf max-concurrent-dd** command to modify the maximum number of neighbors that are concurrently interacting with all OSPF processes on the router. The default value is 10.

GR

By default, the restarter function is disabled, and the helper function is enabled.

Run the **graceful-restart** command to configure the restarter function.

Run the **graceful-restart helper** command to configure the helper function.

Fast Hello

By default, the neighbor dead interval on the interface is 40s.

Run the **ipv6 ospf dead-interval minimal hello-multiplier** command to enable the Fast Hello function on an interface, that is, the neighbor dead interval is 1s.

3.3.4 Network Management Functions

Use functions such as the MIB and Syslog to facilitate OSPF management.

Working Principle

MIB

MIB is the device status information set maintained by a device. You can use the management program to view and set the MIB node.

Multiple OSPF processes can be simultaneously started on a router, but the OSPF MIB can be bound with only one OSPF process.

Trap

A trap message is a notification generated when the system detects a fault. This message contains the related fault information.

If the trap function is enabled, the router can proactively send the trap messages to the network management device.

Syslog

The Syslog records the operations (such as command configuration) performed by users on routers and specific events (such as network connection failures).

If the syslog is allowed to record the adjacency changes, the network administrator can view the logs to learn the entire process that the OSPF adjacency is set up and maintained.

Related Configuration

MIB

By default, the MIB is bound with the OSPF process with the smallest process ID.

Run the **enable mib-binding** command to bind the MIB with the current OSPF process.

Trap

By default, all traps functions are disabled, and the device is not allowed to send OSPF traps.

Run the **snmp-server enable traps ospf** command to allow the device to send OSPF traps.

Run the **enable traps** command to enable a specified trap function for an OSPF process.

Syslog

By default, the Syslog is allowed to record the adjacency changes.

Run the **log-adj-changes** command to allow the Syslog to record the adjacency changes.

3.4 Configuration

Configuration	Description and Command	
Configuring OSPF Basic Functions	(Mandatory)It is used to build an OSPF routing domain.	
	ipv6routerospf	Creates an OSPF process.
	router-id	Configures a router ID.
	ipv6 ospfarea	Enables OSPF on an interface and specifies an area ID.
	area virtual-link	Creates a virtual link.
Setting the Network Type	(Optional) The configurations are mandatory if the physical network is the X.25, frame relay, or ATM network.	
	ipv6 ospf network	Defines the network type.
	ipv6 ospf neighbor	Specifies a neighbor.
	ipv6 ospf priority	Configures the DR priority.
Configuring Route Redistribution and Default Route	(Optional) The configurations are recommended if the OSPF routing domain is connected with an external network.	
	redistribute	Configures route redistribution.

Configuration	Description and Command	
	default-information originate	Introduces a default route.
Configuring the Stub Area	(Optional) It is used to reduce interaction of routing information and the size of routing table, and enhance stability of routes.	
	area stub	Configures a stub area.
Configuring Route Summarization	(Optional) It is used to reduce interaction of routing information and the size of routing table, and enhance stability of routes.	
	area range	Summarizes routes that are advertised between areas.
	summary-prefix	Summarizes routes that are introduced through redistribution.
Configuring Route Filtering	(Optional) It is used to manually control interaction of routing information and filter available OSPF routes.	
	passive-interface	Configures a passive interface.
	distribute-list out	Filters routes that are introduced through redistribution.
	distribute-listin	Filters received LSAs.
Modifying the Route Cost and AD	(Optional) It is used to manually control the shortest route computed by OSPF and determine whether to select an OSPF route preferentially.	
	auto-costreference-bandwidth	Modifies the reference bandwidth of the auto cost.
	ipv6 ospf cost	Modifies the cost in the outbound direction of an interface.
	area default-cost	Modifies the cost of the default route in a stub area.
	default-metric	Modifies the default metric of a redistributed route.
	distance	Modifies the OSPF AD.
Enabling Authentication	(Optional) It is used to prevent routers that illegally access the network and hosts that forge OSPF packets from participating in the OSPF protocol process.	
	area authentication	Enables authentication and sets the authentication mode in an area.
	area encryption	Enables encryption and authentication and sets the authentication mode in an area.
	ipv6 ospf authentication	Enables authentication and sets the authentication mode on an interface.
	ipv6 ospf encryption	Enables encryption and authentication and sets the authentication mode on an interface.
Modifying the Maximum Number of Concurrent Neighbors	(Optional) It is used to prevent the problem of performance deterioration caused by over-consumption of the CPU.	
	max-concurrent-dd	Modifies the maximum number of con current

Configuration	Description and Command	
		neighbors on the current OSPF process.
	ipv6 router ospf max-concurrent-dd	Modifies the maximum number of con current neighbors on all OSPF processes.
Disabling MTU Verification	(Optional) It is used to prevent the problem that the adjacency cannot be set up due to MTU inconsistency on the neighbor interface.	
	ipv6 ospf mtu-ignore	Disables MTU verification on an interface.
Enabling Two-Way Maintenance	(Optional) It is used to prevent termination of the adjacency due to the delay or loss of Hello packets.	
	two-way-maintain	Enables two-way maintenance.
Enabling GR	(Optional) It is used to retain OSPF routing forwarding during restart or active/standby switchover of the OSPF processes to prevent traffic interruption.	
	graceful-restart	Enables the restarter function.
	graceful-restart helper	Enables the helper function.
Enabling Fast Hello	(Optional) It is used to quickly discover the death of a neighbor to prevent traffic interruption when a link is faulty.	
	ipv6 ospf dead-intervalminimal hello-multiplier	Enabling the Fast Hello function on an interface.
Configuring Network Management Functions	(Optional) The configurations enable users to use the SNMP network management software to manage OSPF.	
	enable mib-binding	Bind MIB to the OSPF process.
	enable traps	Enables the trap function of the OSPF process.
	log-adj-changes	Allows the syslogs to record the changes in adjacency status.
Modifying Protocol Control Parameters	(Optional) You are advised not to modify protocol control parameters unless necessary.	
	ipv6 ospf hello-interval	Modifies the Hello interval on an interface.
	ipv6 ospf dead-interval	Modifies the neighbor death interval on an interface.
	ipv6 ospf transmit-delay	Modifies the LSU packet transmission delay on an interface.
	ipv6 ospf retransmit-interval	Modifies the LSU packet retransmission interval on an interface.
	timers throttle lsa all	Modifies parameters of the exponential backoff algorithm that generates LSAs.
	timerspacinglsa-group	Modifies the LSA group update interval.
	timers pacing lsa-transmit	Modifies the LS-UPD packet sending interval.
	timers lsa arrival	Modifies the delay after which the same LSA is received.
	timers throttlespf	Modifies the SPT computation timer.
	timers throttle route inter-area	Modifies the inter-area route computation delay.

Configuration	Description and Command	
	timers throttle route ase	Modifies the inter-area route computation delay.

3.4.1 Configuring OSPF Basic Functions

Configuration Effect

- Set up an OSPF routing domain on the network to provide IPv6 unicast routing service for users on the network.

Notes

- Ensure that the IPv6 routing function is enabled, that is, **ipv6 routing** is not disabled; otherwise, OSPF cannot be enabled.
- IPv6 must be enabled on the interface.
- It is strongly recommended that you manually configure the router ID.

Configuration Steps

Creating an OSPF Process

- Mandatory.
- The configuration is mandatory for every router.

Configuring a Router ID

- (Optional) It is strongly recommended that you manually configure the router ID.
- If the router ID is not configured, OSPF selects an interface IP address. If the IP address is not configured for any interface, or the configured IP addresses have been used by other OSPF instances, you must manually configure the router ID.

Enabling OSPF on an Interface and Specifying an Area ID

- Mandatory.
- The configuration is mandatory for every router.

Verification

- Run the **show ipv6 route ospf** command to verify that the entries of the OSPF routing table are correctly loaded.
- Run the **ping** command to verify that the IPv6 unicast service is correctly configured.

Related Commands

Creating an OSPF Process

Command	**ipv6 router ospf***process-id*
Parameter Description	*process-id:* Indicates the OSPFv3 process ID. If the process ID is not specified, process 1 is enabled.
Command Mode	Global configuration mode
Usage Guide	After enabling the OSPFv3 process, the device enters the routing process configuration mode.

Configuring a Router ID

Command	**router-id***router-id*
Parameter Description	*router-id:* Indicates the ID of the device, which is expressed in the IPv4 address.
Command Mode	OSPF routing process configuration mode
Usage Guide	Every device where OSPFv3 run must be identified by using a router ID. You can configure any IPv4 address as the router ID of the device, and ensure that the router ID is unique in an AS. If multiple OSPFv3 processes run on the same device, the router ID of each process must also be unique. After the router ID changes, OSPF performs a lot of internal processing. Therefore, you are advised not to change the router ID unless necessary. When an attempt is made to modify the router ID, a prompt is displayed, requesting you to confirm the modification. After the OSPFv3 process is enabled, you are advised to specify the router ID before configuring other parameters of the process.

Enabling OSPF on an Interface and Specifying an Area ID

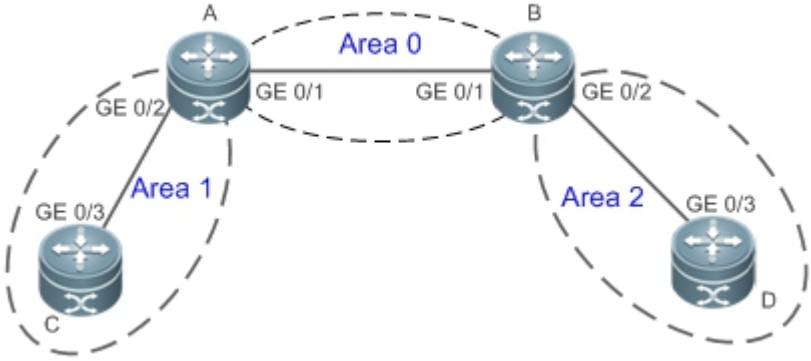
Command	**ipv6 ospf***process-id* **area***area-id* [**instance***instance-id*]
Parameter Description	*process-id*: Indicates the ID of an OSPFv3 process. The value ranges from 1 to 65,535. **Area***area-id:* Indicates the ID of the OSPFv3 area in which the interface participates. It can be an integer or an IPv4 prefix. **Instance***instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	Interface configuration mode
Usage Guide	Run this command in interface configuration mode to enable the interface to participate in OSPFv3, and then run the **ipv6 router ospf** command to configure the OSPFv3 process. After the OSPFv3 process is configured, the interface will automatically participate in the related process. Run the **no ipv6 ospfarea** command so that the specified interface no longer participates in the OSPFv3 routing process. Run the **no ipv6 router ospf** command so that all interfaces no longer participate in the OSPFv3 routing process. The adjacency can be set up only between devices with the same *instance-id*. After this command is configured, all prefix information on the interface will participate in the OSPFv3 process.

Creating a Virtual Link

Command	**area** *area-id***virtual-link***router-id* [**hello-interval** *seconds*] [**dead-interval***seconds*] [**retransmit-interval***seconds*] [**transmit-delay***seconds*] [**instance***instance-id*] [**authenticationipsec spi***spi*[**md5|sha1**] [**0|7**] *key*] [**encryption ipsec spi***spi* **esp** [**null**|[**des | 3des**] [**0** | **7**]*des-key*][**md5|sha1**] [**0|7**] *key*]
Parameter Description	*area-id:* Indicates the ID of the area where the virtual link is located. It can be an integer or an IPv4 prefix. *router-id:* Indicates the router ID of the neighbor connected to the virtual link. **dead-interval***seconds:* Indicates the time that the local interface of the virtual link detects the failure of the neighbor. The unit is second. The value ranges from 1 to 65,535. **hello-interval** *seconds:* Indicates the time that the Hello packet is sent on the local interface of the virtual link. The unit is second. The value ranges from 1 to 65,535. **retransmit-interval** *seconds:* Indicates the interval at which the LSA is retransmitted on the local interface of the

	virtual link. The unit is second. The value ranges from 1 to 65,535. **transmit-delay** *seconds:* Indicates the delay after which the LSA is sent on the local interface of the virtual link. The unit is second. The value ranges from 1 to 65,535. **instance***instance-id:* Indicates the ID of the instance corresponding to the virtual link. The value ranges from 0 to 255. A virtual link cannot be set up between devices with different instance IDs. *spi:* Indicates the security parameter index (SPI). The value ranges from 256 to 4,294,967,295. md5: Enables message digit 5 (MD5) authentication. sha1: Enables Secure Hash Algorithm 1 (SHA1) authentication. 0: Indicates that the key is displayed in plain text. 7: Indicates that the key is displayed in cipher text. *key:* Indicates the authentication key. **null:** Indicates that no encryption mode is used. **des:** Specifies the DES encryption mode. **3des:** Specifies the 3DES encryption mode. *des-key:* Indicates the encryption key.
Command Mode	OSPF routing process configuration mode
Usage Guide	In an OSPFv3 AS, all areas must be connected to the backbone area to properly learn the routing information of the entire OSPFv3 AS. If an area cannot be directly connected to the backbone area, the virtual link can be used to connect this area to the backbone area. The area where the virtual link is located cannot be a stub area. At both ends of neighbors between which the virtual link is set up, settings of **hello-interval, dead-interval**, and **instance** must be consistent; otherwise, the adjacency cannot be set up properly.

Configuration Example

Scenario	 **Remarks**: The interface IP addresses are as follows: A: GE 0/1 2001:1::1/64 GE 0/2 2001:2::1/64 B: GE 0/1 2001:1::2/64 GE 0/2 2001:3::1/64 C: GE 0/3 2001:2::2/64 D: GE 0/3 2001:3::2/64
Configuration Steps	● Configure the interface IP addresses on all routers. ● Enable the IPv4 unicast routing function on all routers. (This function is enabled by default.) ● Configure the OSPF instances and router IDs on all routers. ● Enable OSPF on the interfaces configured on all routers.

A	A#configure terminal A(config)#interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ipv6 enable A(config-if-GigabitEthernet 0/1)#ipv6 address 2001:1::1/64 A(config-if-GigabitEthernet 0/1)#ipv6 ospf 1 area 0 A(config-if-GigabitEthernet 0/1)#exit A(config)#interface GigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)#ipv6 enable A(config-if-GigabitEthernet 0/2)#ipv6 address 2001:2::1/64 A(config-if-GigabitEthernet 0/2)#ipv6 ospf 1 area 1 A(config-if-GigabitEthernet 0/2)#exit A(config)#ipv6 router ospf 1 A(config-router)#router-id1.1.1.1
B	B#configure terminal B(config)#interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)#ipv6 enable B(config-if-GigabitEthernet 0/1)#ipv6 address 2001:1::2/64 B(config-if-GigabitEthernet 0/1)#ipv6 ospf 1 area 0 B(config-if-GigabitEthernet 0/1)#exit B(config)#interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)#ipv6 enable B(config-if-GigabitEthernet 0/2)#ipv6 address 2001:3::1/64 B(config-if-GigabitEthernet 0/2)#ipv6 ospf 1 area 2 B(config-if-GigabitEthernet 0/2)#exit B(config)#ipv6 router ospf 1 B(config-router)#router-id2.2.2.2
C	C#configure terminal C(config)#interface GigabitEthernet 0/3 C(config-if-GigabitEthernet 0/3)#ipv6 enable C(config-if-GigabitEthernet 0/3)#ipv6 address 2001:2::2/64 C(config-if-GigabitEthernet 0/3)#ipv6 ospf 1 area 1 C(config-if-GigabitEthernet 0/3)#exit C(config)#ipv6 router ospf 1

	C(config-router)#router-id3.3.3.3
D	D#configure terminal D(config)#interface GigabitEthernet 0/3 D(config-if-GigabitEthernet 0/3)#ipv6 enable D(config-if-GigabitEthernet 0/3)#ipv6 address 2001:4::2/64 D(config-if-GigabitEthernet 0/3)#ipv6 ospf 1 area 2 D(config-if-GigabitEthernet 0/3)#exit D(config)#ipv6 router ospf 1 D(config-router)#router-id4.4.4.4
Verification	● Verify that the OSPF neighbors are correct on all routers. ● Verify that the routing table is correctly loaded on all routers. ● Verify that 2001:2::2/64 can be pinged successfully on Router D.
A	A#show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 2.2.2.2 1 Full/BDR 00:00:30 0 GigabitEthernet 0/1 3.3.3.31 Full/BDR 00:00:35 0 GigabitEthernet 0/2 A#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O IA2001:3::/64 [110/20] via FE80::2D0:F8FF:FE22:4524, GigabitEthernet 0/1
B	B# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 1.1.1.11 Full/DR 00:00:30 0 GigabitEthernet 0/1 4.4.4.41 Full/BDR 00:00:35 0 GigabitEthernet 0/2

	B#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O IA2001:2::/64 [110/20] via FE80::2D0:F8FF:FE22:4536, GigabitEthernet 0/1
C	C# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 1.1.1.11 Full/DR 00:00:30 0 GigabitEthernet 0/3 C#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O IA2001:1::/64 [110/2] via FE80::2D0:F8FF:FE22:4537, GigabitEthernet 0/3 O IA2001:3::/64 [110/3] via FE80::2D0:F8FF:FE22:4537, GigabitEthernet 0/3
D	D# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 2.2.2.2 1 Full/DR 00:00:30 0 GigabitEthernet 0/3

```
D#show ipv6 route ospf

IPv6 routing table name - Default - 0 entries
Codes:   C - Connected, L - Local, S - Static
         R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route

         E1 - OSPF external type 1, E2 - OSPF external type 2
         SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         IA - Inter area

O   IA2001:1::/64 [110/2] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/3
O   IA2001:2::/64 [110/3] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/3
D#
D#ping 2001:2::2
Sending 5, 100-byte ICMP Echoes to 2001:2::2, timeout is 2 seconds:
< press Ctrl+C to break >
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 5/9/14 ms.
```

Common Errors

- IPv6 is disabled on the interface.
- OSPF cannot be enabled because the IPv6 unicast routing function is disabled.
- The area IDs enabled on adjacent interfaces are inconsistent.
- The same router ID is configured on multiple routers, resulting in a router ID conflict.

3.4.2 Setting the Network Type

Configuration Effect

- If the physical network is X.25, Frame Relay, or ATM, OSPF can also run to provide the IPv6 unicast routing service.

Notes

- The OSPF basic functions must be configured.
- The broadcast network sends multicast OSPF packets, automatically discovers neighbors, and elects a DR and a BDR.
- The P2P network sends multicast OSPF packets and automatically discovers neighbors.
- The NBMA network sends unicast OSPF packets. Neighbors must be manually specified, and a DR and a BDR must be elected.

- The P2MP network (without carrying the **non-broadcast** parameter) sends multicast OSPF packets. Neighbors are automatically discovered.
- The P2MP network (carrying the **non-broadcast** parameter) sends unicast OSPF packets. Neighbors must be manually specified.

Configuration Steps

Configuring the Interface Network Type

- Optional.
- Perform this configuration on routers at both ends of the link.

Configuring a Neighbor

- (Optional)If the interface network type is set to NBMA or P2MP (carrying the **non-broadcast** parameter), neighbors must be configured.
- Neighbors are configured on routers at both ends of the NBMA or P2MP (carrying the **non-broadcast** parameter) network.

Configuring the Interface Priority

- (Optional)You must configure the interface priority if a router must be specified as a DR, or a router cannot be specified as a DR.
- Configure the interface priority on a router that must be specified as a DR, or cannot be specified as a DR.

Verification

- Run the **show ipv6 ospf interface** command to verify that the network type of each interface is correct.

Related Commands

Configuring the Interface Network Type

Command	**ipv6 ospf network** {**broadcast** | **non-broadcast** |**point-to-point** | **point-to-multipoint** [**non-broadcast**]}[**instance***instance-id*]
Parameter Description	**broadcast**: Indicates the broadcast network type. **non-broadcast:** Indicates the non-broadcast network type. **point-to-multipoint**: Indicates the point-to-multipoint (P2MP) network type. **point-to-multipoint non-broadcast:** Indicates the P2MP non-broadcast network type. **point-to-point:** Indicates the point-to-point (P2P) network type. **instance***instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	Interface configuration mode
Usage Guide	You can configure the network type of an interface based on the actual link type and topology.

Configuring a Neighbor

Command	**ipv6 ospf neighbor** *ipv6-address*{ [**cost***cost*] | [**poll-interval***seconds* | **priority***value*] }[**instance***instance-id*]
Parameter Description	*ip-address*: Indicates the link address of theneighborinterface. **cost***cost***:** Indicates the cost required from the P2MP network to each neighbor. The cost is not defined by default. The cost configured on the interface is used. The value ranges from 1 to 65,535. Only a P2MP network supports this option. **poll-interval** *seconds*: Indicates the neighbor polling interval. The unit is second. The value ranges from 1 to

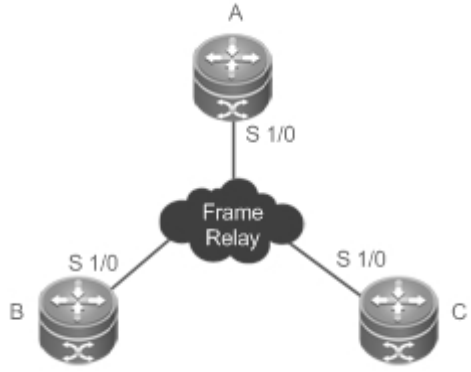
	2,147,483,647. Only the non-broadcast (NBMA) network supports this option. **priority** *value***:** Indicates the priority value of the non-broadcast network neighbor. The value ranges from 0 to 255. Only the non-broadcast network (NBMA) supports this option. **instance***instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	Interface configuration mode
Usage Guide	You can configure neighbor parameters based on the actual network type.

Configuring the Interface Priority

Command	**ipv6 ospf priority** *number-value*[**instance***instance-id*]
Parameter Description	*number-value*: Indicates the priority of the interface. The value ranges from 0 to 255. **instance***instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	Interface configuration mode
Usage Guide	On a broadcast network, a DR or BDR must be elected. During the DR/BDR election, the device with a higher priority will be preferentially elected as a DR or BDR. If the priority is the same, the device with a larger router ID will be preferentially elected as a DR or BDR. A device with the priority 0 does not participate in the DR/BDR election.

Configuration Example

Configuring the Interface Network Type

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers. ● Configure the OSPF basic functions on all routers. ● Set the interface network type to P2MP on all routers.
A	A#configure terminal A(config)# interface Serial1/0 A(config-Serial1/0)# encapsulation frame-relay A(config-Serial1/0)# ipv6 ospf network point-to-multipoint

B	B#configure terminal B(config)# interface Serial1/0 B(config-Serial1/0)# encapsulation frame-relay B(config-Serial1/0)# ipv6 ospf network point-to-multipoint
C	C#configure terminal C(config)# interface Serial1/0 C(config-Serial1/0)# encapsulation frame-relay C(config-Serial1/0)# ipv6 ospf network point-to-multipoint
Verification	● Verify that the interface network type is P2MP.
A	A#show ipv6 ospf interface Serial1/0 Serial1/0 is up, line protocol is up Interface ID 2 IPv6 Prefixes fe80::2d0:f8ff:fe22:3346/64 (Link-Local Address) OSPFv3 Process (1), Area 0.0.0.1, Instance ID 0 Router ID 192.168.22.30,Network Type POINTOMULTIPOINT, Cost: 1 Transmit Delay is 1 sec, State Point-To-Point, Priority 1 Timer interval configured, Hello 30, Dead 120, Wait 40, Retransmit 10 Hello due in 00:00:06 Neighbor Count is 1, Adjacent neighbor count is 1 Hello received 40 sent 40, DD received 17 sent 9 LS-Req received 1 sent 3, LS-Upd received 6 sent 5 LS-Ack received 3 sent 4, Discarded 1

Common Errors

- The network types configured on interfaces at two ends are inconsistent, causing abnormal route learning.
- The network type is set to NBMA or P2MP (non-broadcast), but neighbors are not specified.

3.4.3 Configuring Route Redistribution and Default Route

Configuration Effect

- Introduce unicast routes for other AS domains to the OSPF domain to provide the unicast routing service to other AS domains for users in the OSPF domain.
- In the OSPF domain, inject a default route to another AS domain so that the unicast routing service to another AS domain can be provided for users in the OSPF domain.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Configuring External Route Redistribution

- (Optional)This configuration is mandatory if external routes of the OSPF domain should be introduced to the ASBR.
- Perform this configuration on an ASBR.

Generating a Default Route

- (Optional)Perform this configuration if the default route should be introduced to an ASBR so that other routers in the OSPF domain access other AS domains through this ASBR by default.
- Perform this configuration on an ASBR.

Verification

- On a router inside the OSPF domain, run the **show ipv6 route ospf** command to verify that the unicast routes to other AS domains are loaded.
- On a router inside the OSPF domain, run the **show ipv6 route ospf** command to verify that the default route to the ASBR is loaded.
- Run the **ping** command to verify that the IPv6 unicast service to other AS domains is correct.

Related Commands

Configuring Route Redistribution

Command	**redistribute** {**connected** |**ospf***process-id*| **rip** | **static**}[**match** {**internal | external** [**1**|**2**] } **| metric** *metric-value*|**metric-type** {**1**|**2**} | **route-map** *route-map-name* |**tag***tag-value*]
Parameter Description	**connected:** Indicates redistribution from direct routes. **ospf***process-id:* Indicates redistribution from OSPF.**process-id** specifies an OSPF instance. The value ranges from 1 to 65535. 1-65535 **rip:** Indicates redistribution from RIP. **static:** Indicates redistribution from static routes. **match:** Used only when OSPF routes are redistributed. Only the routes that match the specified criteria are redistributed. By default, all OSPF routes can be redistributed. **metric***metric-value:* Indicates the metric of the OSPF external LSA. *metric-value* specifies the size of the metric. The value ranges from 0 to 16,777,214. **metric-type** {**1**|**2**}: Indicates the external route type, which can be E-1 or E-2. **route-map***route-map-name:* Sets the redistribution filtering rules. **tag***tag-value:* Specifies the tag value of the route that is redistributed into the OSPF routing domain. The value ranges from 0 to 4294967295.
Command Mode	OSPF routing process configuration mode
Usage Guide	When the device supports multiple routing protocols, collaboration between protocols is very important. To run multiple routing protocols concurrently, the device must be able to redistribute routing information of a protocol to another protocol. This applies to all routing protocols.

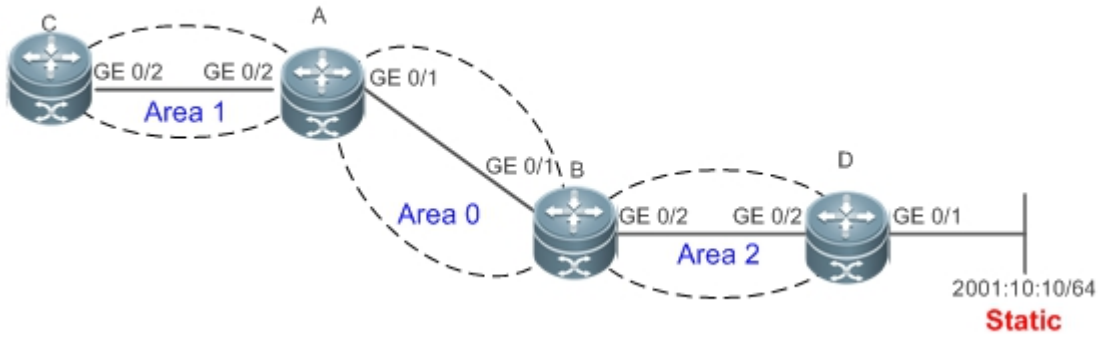
	During redistribution of OSPFv3 routes, **match** can be configured to indicate that OSPFv3 routes of the specified sub-type will be redistributed. By default, all types of OSPFv3 routes are redistributed. For the *match* parameter configured during redistribution of OSPFv3 routes, the routes are matched against the route map only when the sub-type of the routes are correct. During configuration of route redistribution, the **match** rules configured in route map configuration mode areused based on the original information of routes. The priorities of **tag**, **metric** and **metric-type** in the route redistribution configuration are lower than the priority of the**set** rulesconfigured in route map configuration mode. The **set metric** value of the associated routemap should fall into the range of 0 to 16,777,214. If the value exceeds this range, routes cannot be introduced. The configuration rules for the **no** form of the **redistribute** command are as follows: 1. If some parameters are specified in the **no** form of the command, default values of these parameters will be restored. 2. If no parameter is specified in the **no** form of the command, the entire command will be deleted.

Introducing a Default Route

Command	**default-information originate** [**always**] [**metric** *metric*] [**metric-type** *type*] [**route-map***map*]
Parameter Description	**always:** Enables OSPF to generate a default route regardless of whether the local router has a default route. **metric** *metric:* Indicates the initial metric of the default route. The value ranges from 0 to 16,777,214. By default, the metric of the default route is 1. **metric-type***type:* Indicates the type of the default route. OSPF external routes are classified into two types: Type 1: The metric varies with routers; Type 2: The metric is the same for all routers. Type 1 external routes are more trustworthy than Type 2 external routes. **route-map** *map-name:* Indicates the associated route-map name. By default, no route-map is associated.
Command Mode	OSPF routing process configuration mode
Usage Guide	When the **redistribute** or **default-information** command is executed, the OSPFv3-enabled router automatically becomes an ASBR. The ASBR, however, does not automatically generate or advertise a default route to all routers in the OSPF routing domain. To have the ASBR generate a default route, configure the **default-information originate** command. If **always** is specified, the OSPFv3 process advertises an external default route to neighbors no matter whether a default route exists in the core routing table. This default route, however, is not displayed on the local router. To confirm whether the default route is generated, run the **show ipv6 ospf database** command to display the OSPFv3 link status database. On an OSPFv3 neighbor, you can run the **show ipv6 route ospf** command to see the default route. The metric of the external default route can only be defined in the **default-information originate** command, instead of the **default-metric** command. OSPFv3 has two types of external routes. The metric of the Type 1 external route changes, but the metric of the Type 2 external route is fixed. If two parallel paths to the same destination network have the same route metric, the priority of the Type 1 route is higher than that of the Type 2 route. Therefore, the **show ipv6 route ospf** command displays only the Type 1 route. A router in a stub area cannot generate an external default route.

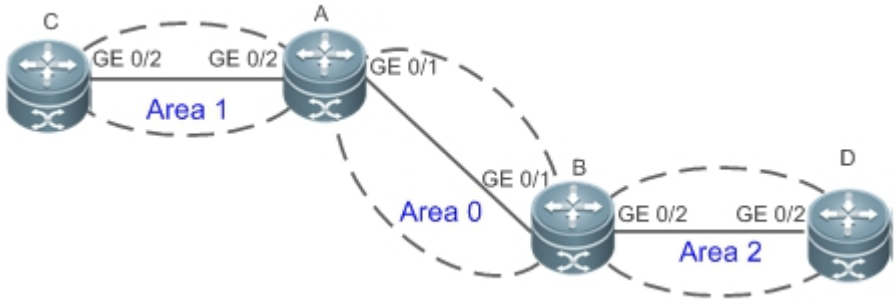
Configuration Example

Configuring Route Redistribution

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers. ● Configure the OSPF basic functions on all routers. ● Introduce an external static route to Router D.
D	D#configure terminal D(config)#ipv6 router ospf 1 D(config-router)# redistribute static
Verification	● On Router D, run the **show ipv6ospf database external brief** command to verify that an LSA corresponding to an external route is generated. ● On Router C, run the **show ipv6 route ospf** command to verify that the external static route has been introduced.
D	D#show ipv6 ospf database external OSPFv3 Router with ID (4.4.4.4) (Process 1) AS-external-LSA LS age: 7 LS Type: AS-External-LSA Link State ID: 0.0.0.6 Advertising Router: 4.4.4.4 LS Seq Number: 0x80000001 Checksum: 0x9C1F Length: 36 Metric Type: 2 (Larger than any link state path) Metric: 20 Prefix: 2001:10:10::/64 Prefix Options: 0 (-|-|-|-)

C	C#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O E2 2001:10:10::/64 [110/20] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/2

Configuring the Default Route

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers. ● Configure the OSPF basic functions on all routers. ● Configure the default route on Router D.
D	D#configure terminal D(config)#ipv6 router ospf 1 D(config-router)#default-information originate always
Verification	● On Router D, run the **show ipv6ospf database external brief** command to verify that an LSA corresponding to the default route is generated. ● On Router C, run the **show ipv6 route ospf** command to verify that the OSPF default route exists.

D	D#show ipv6 ospf database external OSPFv3 Router with ID (4.4.4.4) (Process 1) AS-external-LSA LS age: 3 LS Type: AS-External-LSA Link State ID: 0.0.0.7 Advertising Router: 4.4.4.4 LS Seq Number: 0x80000001 Checksum: 0x1839 Length: 32 Metric Type: 2 (Larger than any link state path) Metric: 1 Prefix: ::/0 Prefix Options: 0 (-|-|-|-) External Route Tag: 1
C	C#show ipv6route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O E2::/0 [110/20] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/2

Common Errors

- A route loop is formed because the **default-information originate always** command is configured on multiple routers.
- Routes cannot be introduced because route redistribution is configured on a router in the stub area.

3.4.4 Configuring the Stub Area

Configuration Effect

- Configure an area located on the stub as a stub area to reduce interaction of routing information and the size of routing table, and enhance stability of routes.

Notes

- The OSPF basic functions must be configured.
- A backbone or transit area cannot be configured as a stub.
- A router in the stub area cannot introduce external routes.

Configuration Steps

Configuring a Stub Area

- (Optional)Perform this configuration if you wish to reduce the size of the routing table on routers in the area.
- Perform this configuration on all routers in the same area.

Verification

Verifying the Stub Area

- On a router in the stub area, run the **show ipv6 route ospf** command to verify that the router is not loaded with any external routes.

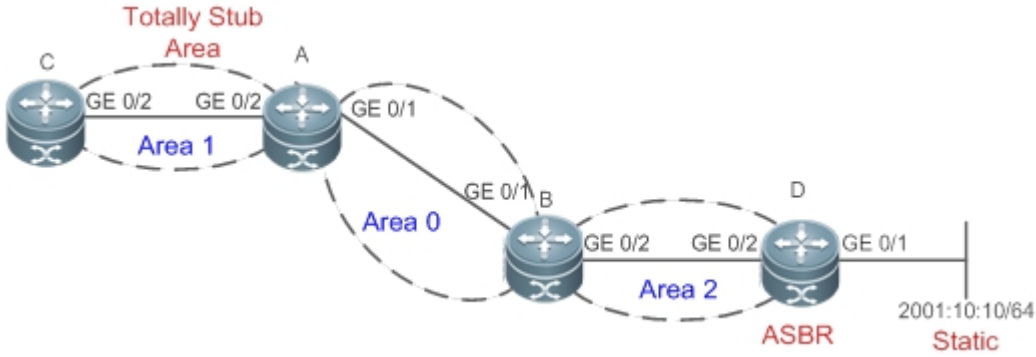
Related Commands

Configuring a Stub Area

Command	**area***area-id***stub** [**no-summary**]
Parameter Description	*area-id:* Indicates the ID of the stub area. The value can be an integer or an IPv4 prefix. **no-summary:** This option is valid only ona the ABR in a stub area. If this option is specified, the ABR only advertises one Type 3 LSA indicating the default route to the stub area, and does not advertise other Type 3 LSAs.
Command Mode	OSPF routing process configuration mode
Usage Guide	An area located on the stub of a network can be configured as a stub area. You must run the **area stub** command on all routers in a stub area. Devices in a stub area cannot learn the external routes (Type 5 LSAs) of the AS. In practice, external routes take up a large proportion of the link status database. Therefore, devices in a stub area can learn only a small amount of routing information, which reduces the amount of system resources required to run the OSPFv3 protocol. By default, an ABR in a stub area will generate a Type 3 LSA indicating the default fault, and advertise the LSA to the stub area. In this way, devices in the stub area can access devices outside the AS. To configure a totally stub area, add the **no-summary** keyword when running the **area stub** command on the ABR.

Configuration Example

Configuring a Stub Area

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Introduce an external static route to Router D. ● Configure area 1 as the stub area on Router A and Router C.
D	D#configure terminal D(config)#ipv6 router ospf 1 D(config-router)#redistribute static
A	A# configure terminal A(config)#ipv6 router ospf 1 A(config-router)#area 1 stubno-summary
C	C#configure terminal C(config)#ipv6 router ospf 1 C(config-router)#area 1 stub
Verification	● On Router C, run the **show ipv6 route ospf** command to display the routing table. Verify that there is only one default inter-area route, and no external static route is introduced from Router D.
C	C#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O IA::/0 [110/3] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/2

Common Errors

- Configurations of the area type are inconsistent on routers in the same area.
- External routes cannot be introduced because route redistribution is configured on a router in the stub area.

3.4.5 Configuring Route Summarization

Configuration Effect

- Summarize routes to reduce interaction of routing information and the size of routing table, and enhance stability of routes.
- Shield or filter routes.

Notes

- The OSPF basic functions must be configured.
- The address range of the summarize route may exceed the actual network range in the routing table. If data is sent to a network beyond the summarization range, a routing loop may be formed and the router processing load may increase. To prevent these problems, a discard route must be added to the routing table or shield or filter routes.

Configuration Steps

Configuring Inter-Area Route Summarization

- (Optional) Perform this configuration when routes of the OSPF area need to be summarized.
- Unless otherwise required, perform this configuration on an ABR in the area where routes to be summarized are located.

Configuring External Route Summarization

- (Optional) Perform this configuration when routes external to the OSPF domain need to be summarized.
- Unless otherwise required, perform this configuration on an ASBR, to which routes that need to be summarized are introduced.

Verification

- Run the **show ipv6 route ospf** command to verify that individual routes do not exist and only the summarized route exists.

Related Commands

Configuring Inter-Area Route Summarization

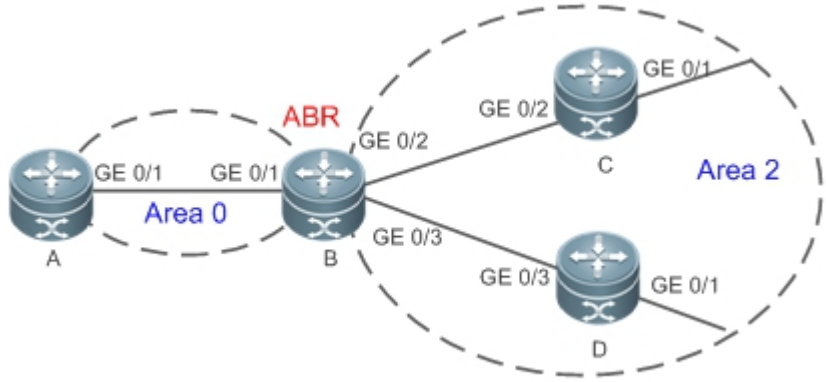
Command	**area***area-id* **range***ipv6-prefix/prefix-length* [**advertise**|**not-advertise**]
Parameter Description	*area-id*: Specifies the ID of the OSPF area to which the summarized route should be injected. The value can be an integer or an IPv4 prefix. *ipv6-prefix/prefix-length*: Indicates the range of IP addresses to be summarized. **advertise** | **not-advertise**: Specifies whether the summarized route should be advertised.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command takes effect only on an ABR, and is used to summarize multiple routes in an area into a route and advertise this route to other areas. Combination of the routing information occurs only on the boundary of an area. Routers inside the area can learn specific routing information, whereas routers in other areas can learn only one summarized route. In addition, you can set **advertise** or **not-advertise** to determine whether to advertise the

	summarized route to shield and filter routes. By default, the summarized route is advertised. You can use the **cost** parameter to set the metric of the summarized route. You can configure route summarization commands for multiple areas. This simplifies routes in the entire OSPF routing domain, and improves the network forwarding performance, especially for a large-sized network. When multiple route summarization commands are configured and have the inclusive relationship with each other, the area range to be summarized is determined based on the maximum match principle.

Configuring External Route Summarization

Command	**summary-prefix***ipv6-prefix/prefix-length* [**not-advertise** | **tag** *number*]
Parameter Description	*ipv6-prefix/prefix-length*: Indicates the range of IP addresses to be summarized. **not-advertise**: Indicates that the summarized route is not advertised. If this parameter is not specified, the summarized route is advertised. **tag***number:* Specifies the tag value of the route that is redistributed into the OSPFv3 routing domain. The value ranges from 0 to 4,294,967,295.
Command Mode	OSPF routing process configuration mode
Usage Guide	When routes are redistributed from other routing processes and injected to the OSPFv3 routing process, each route is advertised to the OSPFv3 routers using an external LSA. If the injected routes are a continuous address space, the ABR can advertise only one summarized route to significantly reduce the size of the routing table. **area range** summarizes the routes between OSPFv3 areas, whereas **summary-prefix** summarizes external routes of the OSPFv3 routing domain.

Configuration Example

Configuration Steps	 **Remarks**: The interface IPv6 addresses are as follows: B: GE0/2 2001:16:2::1/64 GE0/3 2001:16:3::1/64 C: GE0/2 2001:16:2::2/64 GE0/1 2001:16:4::2/64 D: GE0/3 2001:16:3::2/64 GE0/1 2001:16:5::1/64
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Summarize routes of area 2 on Router B.

B	B#configure terminal B(config)#ipv6 router ospf 1 B(config-router)#area 2 range 2001:16::/64
Verification	On Router A, check the routing table and verify that the entry 2001:16::/64 is generated and other routes do not exist.
A	A#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O IA 2001:16::/64 [110/2] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/1

Common Errors

- Inter-area route summarization cannot be implemented because the **area range** command is configured on a non-ABR device.

3.4.6 Configuring Route Filtering

Configuration Effect

- Routes that do not meet filtering conditions cannot be loaded to the routing table, or advertised to neighbors. Network users cannot access specified destination network.

Notes

- The OSPF basic functions must be configured.
- Filtering routes by using the **distribute-list in** command affects forwarding of local routes, but does not affect route computation based on LSAs. Therefore, if route filtering is configured on the ABR, Type 3 LSAs will still be generated and advertised to other areas because routes can still be computed based on LSAs. As a result, black-hole routes are generated. In this case, you can run the **area filter-list** or **area range** (containing the **not-advertise** parameter) command on the ABR to prevent generation of black-hole routes.

Configuration Steps

Configuring Inter-Area Route Filtering

- (Optional) This configuration is recommended if users need to be restricted from accessing the network in a certain OSPF area.
- Unless otherwise required, perform this configuration on an ABR in the area where filtered routes are located.

Configuring Redistributed Route Filtering

- (Optional) Perform this configuration if external routes introduced by the ASBR need to be filtered.
- Unless otherwise required, perform this configurationon an ASBR to which filtered routes are introduced.

Configuring Learned Route Filtering

- (Optional) Perform this configuration if users need to be restricted from accessing a specified destination network.
- Unless otherwise required, perform this configurationon a router that requires route filtering.

Verification

- Run the **show ipv6 route** command to verify that the router is not loaded with routes that have been filtered out.
- Run the **ping** command to verify that the specified destination network cannot be accessed.

Related Commands

Configuring a Passive Interface

Command	**passive-interface** {**default** | *interface-typeinterface-number* }
Parameter Description	*interface-type interface-number*: Indicates the interface that should be configured as a passive interface. **default:** Indicates that all interfaces will be configured as passive interfaces.
Command Mode	OSPF routing process configuration mode
Usage Guide	When an interface is configured as a passive interface, it no longer sends or receives Hello packets. This command takes effect only on an OSPFv3-enabled interface, and not on a virtual link.

Configuring Redistributed Route Filtering

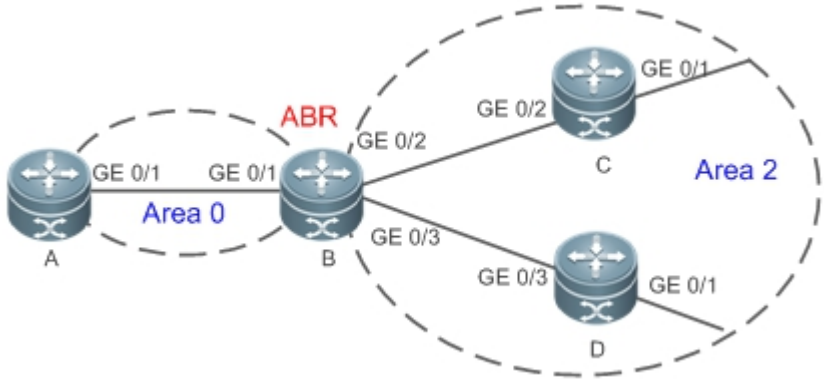
Command	**distribute-list**{*name* | **prefix-list** *prefix-list-name*}**out**[**connected** | **ospf** *process-id*| **rip** | **static**]
Parameter Description	*name*: Uses the ACL for filtering. **prefix** *prefix-list-name*: Uses the prefix list for filtering. **connected | ospf** process-id| **rip** | **static**: Indicates the source of routes to be filtered.
Command Mode	OSPF routing process configuration mode
Usage Guide	**distribute-list out** is similar to **redistribute route-map**, and is used to filter routes that are redistributed from other protocols to OSPFv3. The **distribute-list out** command itself does not redistribute routes, and is generally used together with the **redistribute** command. The ACL and the prefix list filtering rules are mutually exclusive in the configuration. That is, if the ACL is used for filtering routes coming from a certain source, the prefix list cannot be configured to filter the same routes.

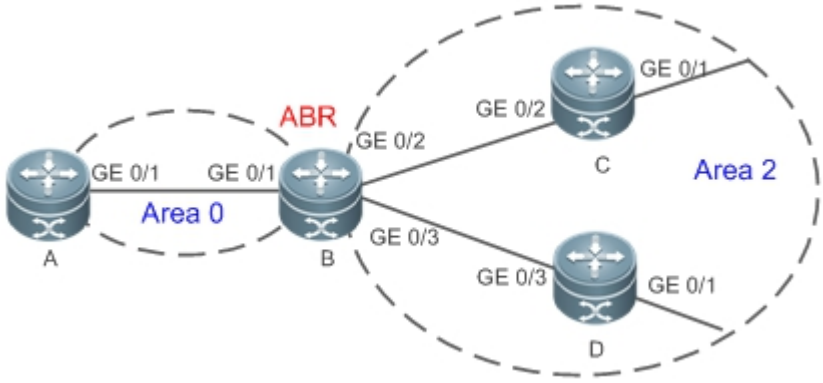
Configuring Learned Route Filtering

Command	**distribute-list**{*name* | **prefix-list** *prefix-list-name*}**in** [*interface-typeinterface-number*]
Parameter Description	*name*: Uses the ACL for filtering. **prefix***prefix-list-name*: Uses the prefix list for filtering. *interface-type interface-number:* Specifies the interface for which LSA routes are filtered.
Command Mode	OSPF routing process configuration mode

Usage Guide	Filter routes that are computed based on received LSAs. Only routes meeting the filtering conditions can be forwarded. The command does not affect the LSDB or the routing tables of neighbors. The ACL and the prefix list filtering rules are mutually exclusive in the configuration. That is, if the ACL is used for filtering routes on a specified interface, the prefix list cannot be configured to filter routes on the same interface. Filtering routes by using the **distribute-list in** command affects forwarding of local routes, but does not affect route computation based on LSAs. Therefore, if route filtering is configured on the ABR, Type 3 LSAs will still be generated and advertised to other areas because routes can still be computed based on LSAs. As a result, black-hole routes are generated. In this case, you can run the **area range** (containing the **not-advertise** parameter) command on the ABR to prevent generation of black-hole routes.

Configuration Example

Scenario	 **Remarks** The interface IPv6 addresses are as follows: B: GE0/2 2001:16:2::1/64 GE0/3 2001:16:3::1/64 C: GE0/2 2001:16:2::2/64 GE0/1 2001:16:4::2/64 D: GE0/3 2001:16:3::2/64 GE0/1 2001:16:5::1/64
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, configure route filtering.
A	A#configure terminal A(config)#ipv6 access-list test A (config-ipv6-acl)#permit ipv6 2001:16:5::/64 any A(config)#ipv6 router ospf 1 A(config-router)#distribute-list test in GigabitEthernet0/1
Verification	● On Router A, check the routing table. Verify that only the entry 2001:16:5::/64 is loaded.
A	A#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2

Scenario	 **Remarks** The interface IPv6 addresses are as follows: B: GE0/2 2001:16:2::1/64 GE0/3 2001:16:3::1/64 C: GE0/2 2001:16:2::2/64 GE0/1 2001:16:4::2/64 D: GE0/3 2001:16:3::2/64 GE0/1 2001:16:5::1/64
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, configure route filtering.
A	A#configure terminal A(config)#ipv6 access-list test A (config-ipv6-acl)#permit ipv6 2001:16:5::/64 any A(config)#ipv6 router ospf 1 A(config-router)#distribute-list test in GigabitEthernet0/1
Verification	● On Router A, check the routing table. Verify that only the entry 2001:16:5::/64 is loaded.
	SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O IA 2001:16:5::/64 [110/2] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/1

Common Errors

- Filtering routes by using the **distribute-list in** command affects forwarding of local routes, but does not affect route computation based on LSAs. Therefore, if route filtering is configured on the ABR, Type 3 LSAs will still be generated and advertised to other areas because routes can still be computed based on LSAs. As a result, black-hole routes are generated.

3.4.7 Modifying the Route Cost and AD

Configuration Effect

- Change the OSPF routes so that the traffic passes through specified nodes or bypasses specified nodes.
- Change the sequence that a router selects routes so as to change the priorities of OSPF routes.

Notes

- The OSPF basic functions must be configured.

- If you run the **ipv6 ospf cost** command to configure the cost of an interface, the configured cost will automatically overwrite the cost that is computed based on the auto cost.

Configuration Steps

Configuring the Reference Bandwidth

- Optional.
- A router is connected with lines with different bandwidths. This configuration is recommended if you wish to preferentially select the line with a larger bandwidth.

Configuring the Cost of an Interface

- Optional.
- A router is connected with multiple lines. This configuration is recommended if you wish to manually specify a preferential line.

Configuring the Default Metric for Redistribution

- Optional.
- This configuration is mandatory if the cost of external routes of the OSPF domain should be specified when external routes are introduced to an ASBR.

Configuring the Maximum Metric

- Optional.
- A router may be unstable during the restart process or a period of time after the router is restarted, and users do not want to forward data through this router. In this case, this configuration is recommended.

Configuring the AD

- Optional.
- Perform this configuration if you wish to change the priorities of OSPF routes on a router that runs multiple unicast routing protocols.

Verification

- Run the **show ipv6 ospf interface** command to verify that the costs of interfaces are correct.
- Run the **show ipv6 route** command to verify that the costs of external routes introduced by the ASBR are correct.
- Restart the router. Within a specified period of time, data is not forwarded through the restarted router.

Related Commands

Configuring the Reference Bandwidth

Command	**auto-costreference-bandwidth** *ref-bw*
Parameter Description	*ref-bw*: Indicates the reference bandwidth. The unit is Mbps. The value ranges from 1 to 4,294,967.
Command Mode	OSPF routing process configuration mode
Usage Guide	You can run the **ipv6 ospf cost** command in interface configuration mode to specify the cost of the interface. The priority

	of this cost is higher than that of the metric computed based on the reference bandwidth.

Configuring the Cost of an Interface

Command	**ipv6 ospf cost** *cost*[**instance***instance-id*]
Parameter Description	*cost*: Indicates the cost of an OSPF interface. The value ranges from 0 to 65,535. **instance***instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	Interface configuration mode
Usage Guide	By default, the cost of an OSPFv3 interface is equal to 100 Mbps/Bandwidth, where Bandwidth is the bandwidth of the interface and configured by the **bandwidth** command in interface configuration mode. The costs of OSPF interfaces on several typical lines are as follows: ● 64 Kbps serial line: The cost is 1562. ● E1 line: The cost is 48. ● 10M Ethernet: The cost is 10. ● 100M Ethernet: The cost is 1. If you run the **ipv6 ospf cost** command to configure the cost of an interface, the configured cost will automatically overwrite the cost that is computed based on the auto cost.

Configuring the Cost of the Default Route in a Stub Area

Command	**area***area-id* **default-cost***cost*
Parameter Description	*area-id*: Indicates the ID of the stub area. *cost*: Indicates the cost of the default summarized route injected to the stub area. The value ranges from 0 to 16,777,215.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command takes effect only on an ABR in a stub area.

Configuring the Default Metric for Redistribution

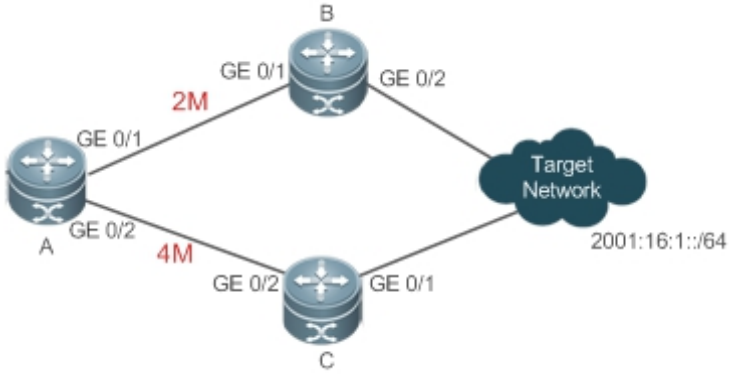
Command	**default-metric** *metric*
Parameter Description	*metric*: Indicates the default metric of the OSPF redistributed route. The value ranges from 1 to 16,777,214.
Command Mode	OSPF routing process configuration mode
Usage Guide	The **default-metric** command must be used together with the **redistribute** command to modify the initial metrics of all redistributed routes. The **default-metric** command does not take effect on external routes that are injected to the OSPF routing domain by the **default-information originate** command. The default metric of a redistributed direct route is always 20.

Configuring the AD

Command	**distance** { *distance* | **ospf**{ [**intra-area***distance*] [**inter-area***distance*] [**external** *distance*]} }
Parameter Description	*distance*: Indicates the AD of a route. The value ranges from 1 to 255. **intra-area***distance*: Indicates the AD of an intra-area route. The value ranges from 1 to 255. **inter-area** *distance*: Indicates the AD of an inter-area route. The value ranges from 1 to 255. **external** *distance:* Indicates the AD of an external route. The value ranges from 1 to 255.

Command Mode	OSPF routing process configuration mode
Usage Guide	Use this command to specify different ADs for different types of OSPF routes. The AD allows different routing protocols to compare route priorities. A smaller AD indicates a higher route priority. The priorities of routes generated by different OSPFv3 processes must be compared based on ADs. If the AD of a route entry is set to 255, the route entry is not trustworthy and does not participate in packet forwarding.

Configuration Example

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, configure the cost of each interface.
A	A#configure terminal A(config)#interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ipv6 ospf cost 10 A(config)#interface GigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)#ipv6 ospf cost 20
Verification	● On Router A, check the routing table. The next hop of the optimum path to 2001:16:1::/64 is Router B.
A	A#show ipv6 route ospf IPv6 routing table name - Default - 0 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area O E2 2001:16:1::/64 [110/2] via FE80::2D0:F8FF:FE22:4547, GigabitEthernet 0/1

Common Errors

- If the cost of an interface is set to 0 in the **ipv6 ospf cost** command, a route computation error may occur. For example, a routing loop is obtained.

3.4.8 Enabling Authentication

Configuration Effect

- All routers connected to the OSPF network must be authenticated to ensure stability of OSPF and protect OSPF against intrusions.

Notes

- The OSPF basic functions must be configured.
- If authentication is configured for an area, the configuration takes effect on all interfaces that belong to this area.
- If authentication is configured for both an interface and the area to which the interface belongs, the configuration for the interface takes effect preferentially.

Configuration Steps

Configuring Authentication

- Optional.
- Perform this configuration if a router accesses a network that requires authentication.

Configuring Encryption

- Optional.
- Perform this configuration if a router accesses a network that requires encryption.

Configuring Virtual Link Authentication

- Optional.
- Perform this configuration if a router accesses a network that requires authentication.

Configuring Virtual Link Encryption

- Optional.
- Perform this configuration if a router accesses a network that requires encryption.

Verification

- If routers are configured with different authentication keys, run the **show ipv6 ospf neighbor** command to verify that there is no OSPF neighbor.
- If routers are configured with the same authentication key, run the **show ipv6 ospf neighbor** command to verify that there are OSPF neighbors.

Related Commands

Configuring Area-based Authentication

Command	**area***area-id***authenticationipsec spi***spi*[**md5|sha1**] [**0 | 7**] *key*
Parameter Description	*area-id*: Indicates the area ID.The value can be an integer or an IPv4 prefix. *spi*: Indicates the SPI. The value ranges from 256 to 4,294,967,295. **md5:** Enables MD5 authentication. **sha1:** Enables SHA1 authentication. **0:** Indicates that the key is displayed in plain text. **7:** Indicates that the key is displayed in cipher text. *key:* Indicates the authentication key.
Command Mode	OSPF routing process configuration mode
Usage Guide	The FSOS supports three authentication types: ● No authentication ● MD5 authentication ● SHA1 authentication Configuration of area-based authentication for OSPFv3 takes effect on all interfaces (except virtual links) in the area, but the interface-based authentication configuration takes precedence over the area-based configuration.

Configuring Area-based Encryption and Authentication

Command	**area***area-id***encryption ipsec spi***spi* **esp**[**null**|[**des** | **3des**][**0 | 7**] *des-key*][**md5|sha1**] [**0|7**] *key*
Parameter Description	*area-id*: Indicates the area ID.The value can be an integer or an IPv4 prefix. *spi*: Indicates the SPI. The value ranges from 256 to 4,294,967,295. **null:** Indicates that no encryption mode is used. **des:** Indicates that the Data Encryption Standard (DES) mode is used. **3des:** Indicates that the Triple DES (3DES) mode is used. des-key: Indicates the encryption key. **md5**: Enables MD5 authentication. **sha1**: Enables SHA1 authentication. **0**: Indicates that the key is displayed in plain text. **7**: Indicates that the key is displayed in cipher text. key: Indicates the authentication key.
Command Mode	OSPF routing process configuration mode
Usage Guide	The FSOS supports two encryption modes and two authentication modes. The two encryption modes are as follows: ● DES ● 3DES The two authentication modes are as follows: ● MD5 ● SHA1 Configuration of area-based encryption and authentication for OSPFv3 takes effect on all interfaces (except virtual links) in the area, but the interface-based encryption and authentication configuration takes precedence over the area-based configuration.

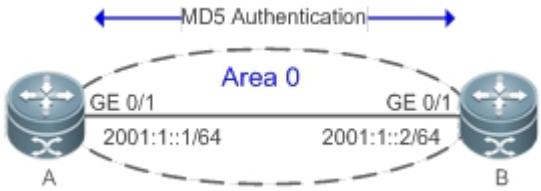
Configuring Interface-based Authentication

Command	**ipv6 ospfauthentication**[**null | ipsec spi***spi*[**md5|sha1**] [**0|7**]*key*][**instance***instance-id*]
Parameter Description	*area-id*: Indicates the area ID.The value can be an integer or an IPv4 prefix. *spi*: Indicates the SPI. The value ranges from 256 to 4,294,967,295. **md5:** Enables MD5 authentication. **sha1:** Enables SHA1 authentication. **0:** Indicates that the key is displayed in plain text. **7:** Indicates that the key is displayed in cipher text. *key:* Indicates the authentication key. **instance** *instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	OSPF routing process configuration mode
Usage Guide	The FSOS supports three authentication types: ● No authentication ● MD5 authentication ● SHA1 authentication OSPFv3 authentication parameters configured on interconnected interfaces must be consistent.

Configuring Interface-based Encryptionand Authentication

Command	**ipv6 ospfencryption ipsec spi***spi* **esp**[**null**|[**des** | **3des**][**0 | 7**] *des-key*][**md5|sha1**] [**0|7**] *key*[**instance***instance-id*]
Parameter Description	*spi*: Indicates the SPI. The value ranges from 256 to 4,294,967,295. **null:** Indicates that no encryption mode is used. **des:** Indicates that the DES mode is used. **3des:** Indicates that the3DES mode is used. des-key: Indicates the encryption key. **md5**: Enables MD5 authentication. **sha1**: Enables SHA1 authentication. **0**: Indicates that the key is displayed in plain text. **7**: Indicates that the key is displayed in cipher text. key: Indicates the authentication key. **instance** *instance-id:* Indicates the ID of a specified OSPFv3 process of the interface. The value ranges from 0 to 255.
Command Mode	OSPF routing process configuration mode
Usage Guide	The FSOS supports two encryption modes and two authentication modes. The two encryption modes are as follows: ● DES ● 3DES The two authentication modes are as follows: ● MD5 ● SHA1 OSPFv3 encryption and authentication parameters configured on the local interface must be consistent with those configured on the interconnected interfaces.

Configuration Example

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Configure MD5 authentication for interfaces of all routers.
A	A#configure terminal A(config)#interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ipv6 ospf authentication ipsec spi 256 md5 01234567890123456789012345678912
B	B# configure terminal B(config)#interface GigabitEthernet 0/3 B(config-if-GigabitEthernet 0/3)#ipv6 ospf authentication ipsec spi 256 md5 01234567890123456789012345678912
Verification	● On Router A and Router B, verify that the OSPF neighbor status is correct.
A	A# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 2.2.2.2 1 Full/DR 00:00:38 0 GigabitEthernet 0/1
B	B# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 1.1.1.1 1 Full/BDR 00:00:38 0 GigabitEthernet 0/1

Common Errors

- The configured authentication modes are inconsistent.
- The configured authentication keys are inconsistent.

3.4.9 Modifying the Maximum Number of Concurrent Neighbors

Configuration Effect

- Control the maximum number of concurrent neighbors on the OSPF process to ease the pressure on the device.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Configuring the Maximum Number of Concurrent Neighbors on the OSPF Process

- (Optional) This configuration is recommended if you wish to set up the OSPF adjacency more quickly when a router is connected with a lot of other routers.
- Perform this configuration on a core router.

Verification

- Run the **show ipv6 ospf neighbor** command to display the number of neighbors that are concurrently interacting with the OSPF process.

Related Commands

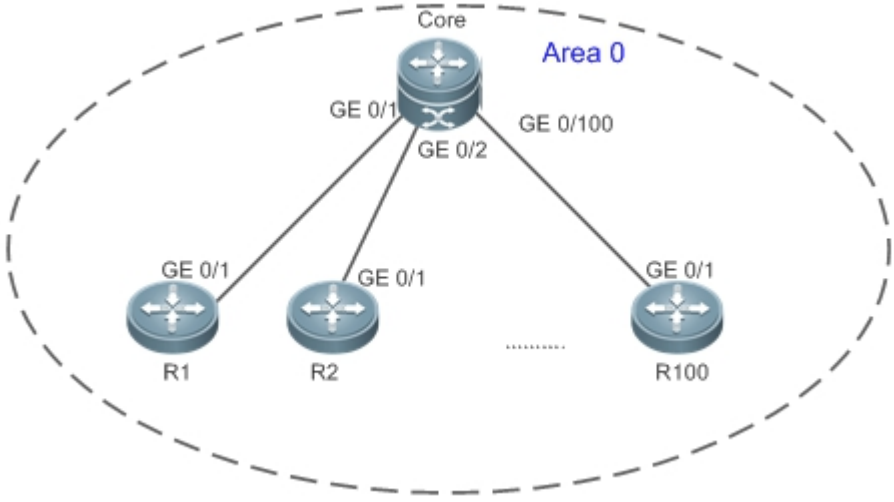
Configuring the Maximum Number of Concurrent Neighbors on the Current Process

Command	**max-concurrent-dd***number*
Parameter Description	*number:* Specifies the maximum number of neighbors that are concurrently interacting with the OSPF process. The value ranges from 1 to 65,535.
Command Mode	OSPF routing process configuration mode
Usage Guide	When the performance of a router is affected because the router exchanges data with multiple neighbors, you can configure this command to restrict the maximum of neighbors with which each OSPF process can concurrently initiate or accept interaction.

Configuring the Maximum Number of Concurrent Neighbors on All Processes

Command	**Ipv6 router ospf max-concurrent-dd***number*
Parameter Description	*number:* Specifies the maximum number of neighbors that are concurrently interacting with the OSPF process. The value ranges from 1 to 65,535.
Command Mode	Global configuration mode
Usage Guide	When the performance of a router is affected because the router exchanges data with multiple neighbors, you can configure this command to restrict the maximum of neighbors with which all OSPF processes can concurrently initiate or accept interaction.

Configuration Example

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On the Router Core, set the maximum number of concurrent neighbors to 4.
Core	Core# configure terminal Core(config)# ipv6 router ospf max-concurrent-dd 4
Verification	● On the Router Core, check the neighbor status and verify that at most eight neighbors concurrently interact with the OSPF process.

Common Errors

N/A

3.4.10 Disabling MTU Verification

Configuration Effect

- The unicast routing service can be provided even if the MTUs of interfaces on neighbor routers are different.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Disabling MTU Verification

- (Optional) MTU verification is disabled by default. You are advised to retain the default configuration.
- Perform this configuration on two routers with different interface MTUs.

Verification

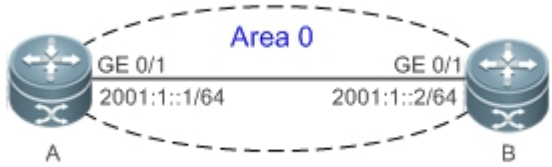
- The adjacency can be set up between routers with different MTUs.

Related Commands

Disabling MTU Verification

Command	**Ipv6 ospf mtu-ignore**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	On receiving the database description packet, OSPF checks whether the MTU of the interface on the neighbor is the same as the MTU of its own interface. If the interface MTU specified in the received database description packet is greater than the MTU of the local interface, the adjacency cannot be set up. To resolve this problem, you can disable MTU verification.

Configuration Example

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Configure different MTUs for interfaces on two routers. ● Disable MTU verification on all routers. (By default, the function of disabling MTU verification is enabled.)
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ipv6 mtu 1400 A(config-if-GigabitEthernet 0/1)#ipv6 ospf mtu-ignore
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ipv6 mtu 1600 B(config-if-GigabitEthernet 0/1)# ipv6 ospf mtu-ignore
Verification	● On Router A, verify that the OSPF neighbor information is correct.
A	A# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 2.2.2.2 1 Full/DR 00:00:38 0 GigabitEthernet 0/1

Common Errors

N/A

3.4.11 Enabling Two-Way Maintenance

Configuration Effect

- Non-Hello packets can also be used to maintain the adjacency.

Notes

- The OSPF basic functions must be configured.

Configuration Steps

Enabling Two-Way Maintenance

- (Optional) This function is enabled by default. You are advised to retain the default configuration.
- Perform this configuration on all routers.

Verification

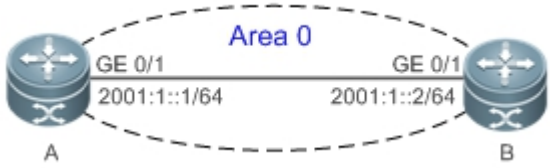
- Non-Hello packets can also be used to maintain the adjacency.

Related Commands

Enabling Two-Way Maintenance

Command	**two-way-maintain**
Parameter Description	N/A
Command Mode	OSPF routing process configuration mode
Usage Guide	On a large network, a lot of packets may be sent or received, occupying too much CPU and memory. As a result, some packets are delayed or discarded. If the processing time of Hello packets exceeds the dead interval, the adjacency will be destroyed due to timeout.If the two-way maintenance function is enabled, in addition to the Hello packets, the DD, LSU, LSR, and LSAck packets can also be used to maintain the bidirectional communication between neighbors when a large number of packets exist on the network. This prevents termination of the adjacency caused by delayed or discarded Hello packets.

Configuration Example

Scenario	
Configuration Steps	• Enable IPv6 on interfaces of all routers.(Omitted) • Configure the OSPF basic functions on all routers. (Omitted) • On Router A, enable the two-way maintenance function. (This function is enabled by default.)
A	A# configure terminal A(config)# ipv6 routerospf 1 A(config-router)#two-way-maintain

Verification	● When the adjacency is being set up, Router A checks the neighbor dead interval and updates the dead interval without waiting for Router B to send a Hello packet.
A	A# show ipv6 ospfneighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 2.2.2.2 1 Full/DR 00:00:38 0 GigabitEthernet 0/1

Common Errors

N/A

3.4.12 Enabling GR

Configuration Effect

- When a distributed route switches services from the active board to the standby board, traffic forwarding continues and is not interrupted.
- When the OSPF process is being restarted, data forwarding continues and is not interrupted.

Notes

- The OSPF basic functions must be configured.
- The neighbor router must support the GR helper function.
- The grace period cannot be shorter than the neighbor dead time of the neighbor router.

Configuration Steps

Configuring the OSPF GR Function

- (Optional) This function is enabled by default. You are advised to retain the default configuration.
- Perform this configuration on routers where hot standby switchover is triggered or the OSPF process is restarted.

Configuring the OSPF GR Helper Function

- (Optional) This function is enabled by default. You are advised to retain the default configuration.
- Perform this configuration on a router if hot standby switchover is triggered or the OSPF process is restarted on a neighbor of this router.

Verification

- When a distributed router switches services from the active board to the standby board, data forwarding continues and the traffic is not interrupted.
- When the OSPF process is being restarted, data forwarding continues and the traffic is not interrupted.

Related Commands

Configuring the OSPF GR Function

Command	**graceful-restart** [**grace-period***grace-period*|**inconsistent-lsa-checking**]

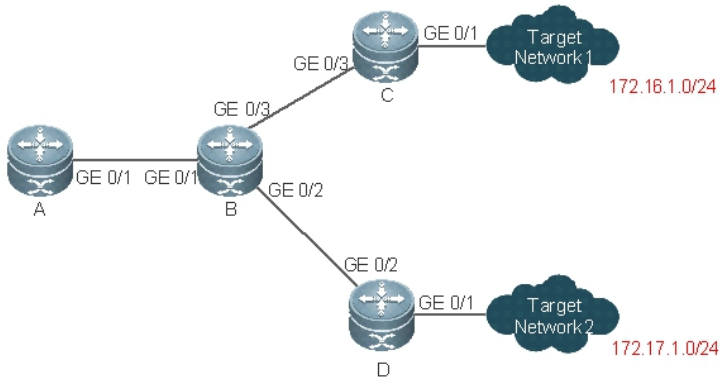
Parameter Description	**grace-period** *grace-period*: Indicates the grace period, which is the maximum time from occurrence of an OSPF failure to completion of the OSPF GR. The value of the grace period varies from 1s to 1800s. The default value is 120s. **inconsistent-lsa-checking**: Enables topological change detection. If any topological change is detected, OSPF exits the GR process to complete convergence.After GR is enabled, topological change detection is enabled by default.
Command Mode	OSPF routing process configuration mode
Usage Guide	The GR function is configured based on the OSPF process. You can configure different parameters for different OSPF processes based on the actual conditions. This command is used to configure the GR restarter capability of a device. The grace period is the maximum time of the entire GR process, during which link status is rebuilt so that the original state of the OSPF process is restored. After the grace period expires, OSPF exits the GR state and performs common OSPF operations. Run the **graceful-restart** command to set the grace period to 120s. The **graceful-restart grace-period** command allows you to modify the grace period explicitly. The precondition for successful execution of GR and uninterrupted forwarding is that the topology remains stable. If the topology changes, OSPF quickly converges without waiting for further execution of GR, thus avoiding long-time forwarding black-hole. ● Disabling topology detection: If OSPF cannot converge in time when the topology changes during the hot standby process, forwarding black-hole may appear in a long time. ● Enabling topology detection: Forwarding may be interrupted when topology detection is enabled, but the interruption time is far shorter than that when topology detection is disabled. In most cases, it is recommended that topology detection be enabled. In special scenarios, topology detection can be disabled if the topology changes after the hot standby process, but it can be ensured that the forwarding black-hole will not appear in a long time. This can minimize the forwarding interruption time during the hot standby process. If the Fast Hello function is enabled, the GR function cannot be enabled.

Configuring the OSPF GR Helper Function

Command	**graceful-restart helper** { **disable |strict-lsa-checking | internal-lsa-checking**}
Parameter Description	**disable**: Prohibits a device from acting as a GR helper for another device. **strict-lsa-checking**: Indicates that changes in Type 1 to Type 5 and Type 7 LSAs will be checked during the period that the device acts as a GR helper to determine whether the network changes. If the network changes, the device will stop acting as the GR helper. **internal-lsa-checking:** Indicates that changes in Type 1 to Type 3 LSAs will be checked during the period that the device acts as a GR helper to determine whether the network changes. If the network changes, the device will stop acting as the GR helper.
Command Mode	OSPF routing process configuration mode
Usage Guide	This command is used to configure the GR helper capability of a router. When a neighbor router implements GR, it sends a Grace-LSA to notify all neighbor routers. If the GR helper function is enabled on the local router, the local router becomes the GR helper on receiving the Grace-LSA, and helps the neighbor to complete GR. The **disable** option indicates that GR helper is not provided for any device that implements GR. After a device becomes the GR helper, the network changes are not detected by default. If any change takes place on the network, the network topology converges after GR is completed. If you wish that network changes can be quickly detected during the GR process, you can configure **strict-lsa-checking** to check Type 1 to 5 and Type 7 LSAs that indicate the network information or **internal-lsa-checking** to check Type 1 to 3 LSAs that indicate internal routes of the

	AS domain. When the network scale is large, it is recommended that you disable the LSA checking options (**strict-lsa-checking** and **internal-lsa-checking**) because regional network changes may trigger termination of GR and consequently reduce the convergence of the entire network.

Configuration Example

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● On Router A, Router C, and Router D, enable the GR helper function. (This function is enabled by default.) ● On Router B, enable the GR function.
B	B# configure terminal B(config)# ipv6 router ospf1 B(config-router)# graceful-restart
Verification	● Trigger a hot standby switchover on Router B, and verify that the routing tables of destination Network 1 and Network 2 remain unchanged on Router A during the switchover. ● Trigger a hot standby switchover on Router B, ping destination Network 1 from Router A, and verify that traffic forwarding is not interrupted during the switchover.

Common Errors

- Traffic forwarding is interrupted during the GR process because the configured grace period is shorter than the neighbor dead time of the neighbor router.

3.4.13 Configuring Network Management Functions

Configuration Effect

- Use the network management software to manage OSPF parameters and monitor the OSPF running status.

Notes

- The OSPF basic functions must be configured.
- You must enable the MIB function of the SNMP server before enabling the OSPF MIB function.
- You must enable the trap function of the SNMP server before enabling the OSPF trap function.
- You must enable the logging function of the device before outputting the OSPF logs.

Configuration Steps

Binding the MIB with the OSPF Process

- (Optional) This configuration is required if you want to use the network management software to manage parameters of a specified OSPF process.
- Perform this configuration on all routers.

Enabling the Trap Function

- (Optional) This configuration is required if you want to use the network management software to monitor the OSPF running status.
- Perform this configuration on all routers.

Configuring the Logging Function

- (Optional) This function is enabled by default. You are advised to retain the default configuration. If you want to reduce the log output, disable this function.
- Perform this configuration on all routers.

Verification

- Use the network management software to manage the OSPF parameters.
- Use the network management software to monitor the OSPF running status.

Related Commands

Binding the MIB with the OSPF Process

Command	**enable mib-binding**
Parameter Description	N/A
Command Mode	OSPF routing process configuration mode
Usage Guide	The OSPFv2 MIB does not have the OSPFv3 process information. Therefore, you can perform operations only on a single OSPFv2 process through SNMP. By default, the OSPFv3 MIB is bound with the OSPFv3 process with the smallest process ID, and all user operations take effect on this process. If you wish to perform operations on a specified OSPFv3 process through SNMP, run this command to bind the MIB with the process.

Enabling the Trap Function

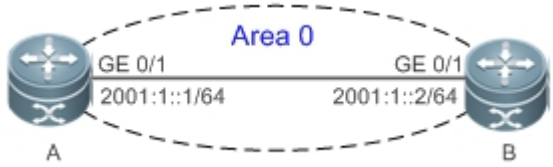
Command	**enable traps**[**error** [**IfConfigError**| **IfRxBadPacket** | **VirtIfConfigError** | **VirtIfRxBadPacket**] | **state-change**[**IfStateChange** | **NbrStateChange** | **VirtIfStateChange** | **VirtNbrStateChange** | **RestartStatusChange** | **NbrRestartHelperStatusChange** | **VirtNbrRestartHelperStatusChange**]]
Parameter Description	**IfConfigError:** Indicates that an interface parameter configuration error occurs. **IfRxBadPacket:** Indicates that the interface receives a bad packet. **VirtIfConfigError:** Indicates that a virtual interface parameter configuration error occurs. **VirtIfRxBadPacket:** Indicates that the virtual interface receives a bad packet. **IfStateChange:** Indicates that interface state changes.

	NbrStateChange: Indicates that the neighbor state changes. **VirtIfStateChange:** Indicates that the virtual interface state changes. **VirtNbrStateChange:** Indicates that the virtual neighbor state changes. **RestartStatusChange:** Indicates that the GR state of the local device changes. **NbrRestartHelperStatusChange:** Indicates that the state of the neighbor GR process changes. **VirtNbrRestartHelperStatusChange:** Indicates that the GR state of the virtual neighbor changes.
Command Mode	OSPF routing process configuration mode
Usage Guide	The function configured by this command is restricted by the **snmp-server** command. You can configure **snmp-server enable traps ospf** and then **enable traps** command before the corresponding OSPF traps can be correctly sent out. This command is not restricted by the MIB bound with the process. The trap function can be enabled concurrently for different processes.

Configuring the Logging Function

Command	**log-adj-changes**[**detail**]
Parameter Description	**detail**: Records all status change information.
Command Mode	OSPF routing process configuration mode
Usage Guide	N/A

Configuration Example

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Bind the MIB with the OSPF process on Router A. ● Enable the trap function on Router A.
A	A# configure terminal A(config)#snmp-server host 192.168.2.2 traps version 2c public A(config)#snmp-server community public rw A(config)#snmp-server enable traps A(config)# A(config)# ipv6 routerospf 10 A(config-router)# enable mib-binding A(config-router)# enable traps
Verification	● Use the MIB tool to read and set the OSPF parameters and display the OSPF running status.

Common Errors

N/A

3.4.14 Modifying Protocol Control Parameters

Configuration Effect

- Modify protocol control parameters to change the protocol running status.

Notes

- The OSPF basic functions must be configured.
- The neighbor dead time cannot be shorter than the Hello interval.

Configuration Steps

Configuring the Hello Interval

- (Optional) You are advised to retain the default configuration.
- Perform this configuration on routers at both end of a link.

Configuring the Dead Interval

- (Optional) You are advised to retain the default configuration. This configuration can be adjusted if you wish to accelerate OSPF convergence when a link fails.
- Perform this configuration on routers at both end of a link.

Configuring the LSU Retransmission Interval

- (Optional) You are advised to adjust this configuration if a lot of routes exist in the user environment and network congestion is serious.

Configuring the LSA Generation Time

- (Optional) You are advised to retain the default configuration.

Configuring the LSA Group Refresh Time

- (Optional) You are advised to retain the default configuration. This configuration can be adjusted if a lot of routes exist in the user environment.
- Perform this configuration on an ASBR or ABR.

Configuring LSA Repeated Receiving Delay

- (Optional) You are advised to retain the default configuration.

Configuring the SPF Computation Delay

- (Optional) This configuration can be adjusted if network flapping frequently occurs.

Configuring the Inter-Area Route Computation Delay

- (Optional) You are advised to retain the default configuration.

- Perform this configuration on all routers.

Configuring the Inter-Area Route Computation Delay

- (Optional) You are advised to retain the default configuration.
- Perform this configuration on all routers.

Verification

- Run the **show ipv6 ospf** and **show ipv6 ospf neighbor** commands to display the protocol running parameters and status.

Related Commands

Configuring the Hello Interval

Command	**ipv6 ospf hello-interval***seconds*
Parameter Description	*seconds*: Indicates the interval at which OSPF sends the Hello packet. The unit is second. The value ranges from 1 to 65,535.
Command Mode	Interface configuration mode
Usage Guide	The Hello interval is contained in the Hello packet. A shorter Hello interval indicates that OSPF can detect topological changes more quickly, but the network traffic increases. The Hello interval must be the same on all routers in the same network segment. If you want to manually modify the neighbor dead interval, ensure that the neighbor dead interval is longer than the Hello interval.

Configuring the Dead Interval

Command	**ipv6 ospf dead-interval** *seconds*
Parameter Description	*seconds*: Indicates the time that the neighbor is declared lost. The unit is second. The value ranges from 1 to 65,535.
Command Mode	Interface configuration mode
Usage Guide	The OSPF dead interval is contained in the Hello packet. If OSPF does not receive a Hello packet from a neighbor within the dead interval, it declares that the neighbor is invalid and deletes this neighbor record form the neighbor list. By default, the dead interval is four times the Hello interval. If the Hello interval is modified, the dead interval is modified automatically. When using this command to manually modify the dead interval, pay attention to the following issues: 1. The dead interval cannot be shorter than the Hello interval. 2. The dead interval must be the same on all routers in the same network segment.

Configuring the LSU Transmission Delay

Command	**ipv6 ospf transmit-delay***seconds*
Parameter Description	*seconds*: Indicates the LSU transmission delay on the OSPF interface. The unit is second. The value ranges from 0 to 65,535.
Command Mode	Interface configuration mode
Usage Guide	Before an LSU packet is transmitted, the Age fields in all LSAs in this packet will increase based on the amount specified by the **ip ospf transmit-delay** command. Considering the transmission delay and line propagation delay on the

	interface, you need to set the LSU transmission delay to a greater value for a low-speed line or interface. The LSU transmission delay of a virtual link is defined by the **transmit-delay** parameter in the **area virtual-link** command. If the value of the Age field of an LSA reaches 3600, the packet will be retransmitted or a retransmission will be requested. If the LSA is not updated in time, the expired LSA will be deleted from the LSDB.

Configuring the LSU Retransmission Interval

Command	**ipv6 ospf retransmit-interval** *seconds*
Parameter Description	*seconds*: Indicates the LSU retransmission interval. The unit is second. The value ranges from 0 to 65,535. This interval must be longer than the round-trip transmission delay of data packets between two neighbors.
Command Mode	Interface configuration mode
Usage Guide	After a router finishes sending an LSU packet, this packet is still kept in the transmit buffer queue. If an acknowledgment from the neighbor is not received within the time defined by the **ip ospf retransmit-interval** command, the router retransmits the LSU packet. The retransmission delay can be set to a greater value on a serial line or virtual link to prevent unnecessary retransmission. The LSU retransmission delay of a virtual link is defined by the **retransmit-interval**parameter in the **area virtual-link** command.

Configuring the LSA Generation Time

Command	**timers throttle lsa all** *delay-time hold-time max-wait-time*
Parameter Description	*delay-time*: Indicates the minimum delay for LSA generation. The first LSA in the database is always generated instantly. The value ranges from 0 to 600,000. The unit is ms. *hold-time*: Indicates the minimum interval between the first LSA update and the second LSA update. The value ranges from 1 to 600,000. The unit is ms. *max-wait-time*: Indicates the maximum interval between two LSA updates when the LSA is updated continuously. This interval is also used to determine whether the LSA is updated continuously. The value ranges from 1 to 600,000. The unit is ms.
Command Mode	OSPF routing process configuration mode
Usage Guide	If a high convergence requirement is raised when a link changes, you can set **delay-time** to a smaller value. You can also appropriately increase values of the preceding parameters to reduce the CPU usage. When configuring this command, the value of **hold-time** cannot be smaller than the value of **delay-time,** and the value of **max-wait-time** cannot be smaller than the value of **hold-time**.

Configuring the LSA Group Refresh Time

Command	**timers pacinglsa-group** *seconds*
Parameter Description	*seconds*: Indicates the LSA group pacing interval. The value ranges from 10 to 1,800. The unit is second.
Command Mode	OSPF routing process configuration mode
Usage Guide	Every LSA has a time to live (LSA age). When the LSA age reaches 1800s, a refreshment is needed to prevent LSAs from being cleared because their ages reaching the maximum. If LSA update and aging computation are performed for every LSA, the device will consume a lot of CPU resources. In order to use CPU resources effectively, you can refresh LSAs by group on the device. The interval of group refreshment is called group pacing interval. The group refreshment operation

	is to organize the LSAs generated within a group pacing interval into a group and refresh the group as a whole. If the total number of LSAs does not change, a larger group pacing interval indicates that more LSAs need to be processed after timeout. To maintain the CPU stability, the number of LSAs processes upon each timeout cannot be too large. If the number of LSAs is large, you are advised to reduce the group pacing interval. For example, if there are 1000 LSAs in the database, you can reduce the pacing interval; if there are 40 to 100 LSAs, you can set the pacing interval to 10-20 minutes.

Configuring the LSA Group Refresh Interval

Command	**timers pacing lsa-transmit** *transmit-time transmit-count*
Parameter Description	*transmit-time*: Indicates the LSA group transmission interval. The value ranges from 10 to 600,000. The unit is ms. *transmit-count*: Indicates the number of LS-UPD packets in a group. The value ranges from 1 to 200.
Command Mode	OSPF routing process configuration mode
Usage Guide	If the number of LSAs is large and the device load is heavy in an environment, properly configuring **transimit-time** and **transimit-count** can limit the number of LS-UPD packets flooded on a network. If the CPU usage is not high and the network bandwidth load is not heavy, reducing the value of **transimit-time** and increasing the value of **transimit-count** can accelerate the environment convergence.

Configuring LSA Repeated Receiving Delay

Command	**timers lsa arrival***arrival-time*
Parameter Description	*arrival-time*: Indicates the delay after which the same LSA is received. The value ranges from 0 to 600,000. The unit is ms.
Command Mode	OSPF routing process configuration mode
Usage Guide	No processing is performed if the same LSA is received within the specified time.

Configuring the SPF Computation Delay

Command	**timers throttle spf** *spf-delay spf-holdtime spf-max-waittime*
Parameter Description	*spf-delay*: Indicates the SPF computation delay. The unit is ms. The value ranges from 1 to 600,000. When detecting a topological change, the OSPF routing process triggers the SPF computation at least after **spf-delay** elapses. *spf-holdtime*: Indicates the minimum interval between two SPF computations. The unit is ms. The value ranges from 1 to 600,000. *spf-max-waittime*: Indicates the maximum interval between two SPF computations. The unit is ms. The value ranges from 1 to 600,000. *number*: Indicates the metric of the summarized route.
Command Mode	OSPF routing process configuration mode
Usage Guide	**spf-delay** indicates the minimum time between the occurrence of the topological change and the start of SPF computation. **spf-holdtime** indicates the minimum interval between the first SPF computation and the second SPF computation. After that, the interval between two SPF computations must be at least twice of the previous interval. When the interval reaches **spf-max-waittime**, the interval cannot increase again. If the interval between two SPF computations already exceeds the required minimum value, the interval is computed by starting from **spf-holdtime**. You can set **spf-dela**y and **spf-holdtime** to smaller values to accelerate topology convergence, and set **spf-max-waittime** to a larger value to reduce SPF computation. Flexible settings can be used based on stability of the

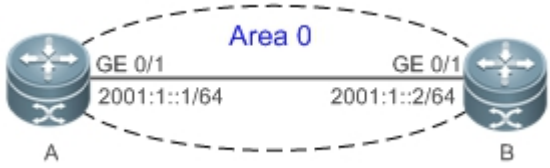
	network topology. Compared with the **timers spf** command, this command supports more flexible settings to accelerate the convergence speed of SPF computation and further reduce the system resources consumed by SPF computation when the topology continuously changes. Therefore, you are advised to use the **timers throttle spf** command for configuration. 1. The value of **spf-holdtime** cannot be smaller than the value of **spf-delay**; otherwise, **spf-holdtime** will be automatically set to the value of **spf-delay**. 2. The value of **spf-max-waittime** cannot be smaller than the value of **spf-holdtime**; otherwise, **spf-max-waittime** will be automatically set to the value of **spf-holdtime**. 3. The configurations of **timers throttle spf** and **timers spf** are mutually overwritten. 4. When both **timers throttle spf** and **timers spf** are not configured, the default values of **timers throttle spf** prevail.

Configuring the Computation Delays of Inter-Area Routes and External Routes

Command	**timers throttle route**{**inter-area***ia-delay*|**ase***ase-delay*}
Parameter Description	**inter-area***ia-delay*: Indicates the inter-area route computation delay. The unit is ms. The value ranges from 0 to 600,000. **ase***ase-delay:* Indicates the external route computation delay. The unit is ms. The value ranges from 0 to 600,000.
Command Mode	OSPF routing process configuration mode
Usage Guide	If a strict requirement is raised for the network convergence time, use the default value. If a lot of inter-area or external routes exist on the network and the network is not stable, adjust the delays and optimize route computation to reduce the load on the device.

Configuration Example

Configuring the Hello Interval and Dead Interval

Scenario	
Configuration Steps	● Enable IPv6 on interfaces of all routers.(Omitted) ● Configure the OSPF basic functions on all routers. (Omitted) ● Configure the Hello interval and dead interval on all routers.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ipv6 ospf hello-interval 15 A(config-if-GigabitEthernet 0/1)# ipv6 ospf dead-interval 50
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ipv6 ospf hello-interval 15 A(config-if-GigabitEthernet 0/1)# ipv6 ospf dead-interval 50

Verification	● Check the interface parameters on Router A and Router B. Verify that the Hello interval is 10s and the dead interval is 50s. ● On Router A and Router B, verify that the OSPF neighbor information is correct.
A	A# show ipv6 ospf interface GigabitEthernet 0/1 is up, line protocol is up Interface ID 2 IPv6 Prefixes fe80::2d0:f8ff:fe22:3346/64 (Link-Local Address) OSPFv3 Process (1), Area 0.0.0.0, Instance ID 0 Router ID 1.1.1.1,Network Type BROADCAST, Cost: 1 Transmit Delay is 1 sec, State DR, Priority 1 Timer interval configured, Hello 15, Dead 50, Wait 40, Retransmit 10 Hello due in 00:00:06 Neighbor Count is 1, Adjacent neighbor count is 1 Hello received 40 sent 40, DD received 17 sent 9 LS-Req received 1 sent 3, LS-Upd received 6 sent 5 LS-Ack received 3 sent 4, Discarded 1 A# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 2.2.2.21 Full/BDR 00:00:30 0 GigabitEthernet 0/1

B	B# show ipv6 ospf interface GigabitEthernet 0/1 is up, line protocol is up Interface ID 2 IPv6 Prefixes fe80::2d0:f8ff:fe22:3446/64 (Link-Local Address) OSPFv3 Process (1), Area 0.0.0.0, Instance ID 0 Router ID 2.2.2.2,Network Type BROADCAST, Cost: 1 Transmit Delay is 1 sec, State BDR, Priority 1 Timer interval configured, Hello 15, Dead 50, Wait 40, Retransmit 10 Hello due in 00:00:06 Neighbor Count is 1, Adjacent neighbor count is 1 Hello received 40 sent 40, DD received 17 sent 9 LS-Req received 1 sent 3, LS-Upd received 6 sent 5 LS-Ack received 3 sent 4, Discarded 1 B# show ipv6 ospf neighbor OSPFv3 Process (1), 1 Neighbors, 1 is Full: Neighbor ID Pri State Dead Time Instance ID Interface 1.1.1.11 Full/DR 00:00:38 0 GigabitEthernet 0/1

Common Errors

- The configured neighbor dead time is shorter than the Hello interval.

3.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears and resets an OSPF process.	**clear ipv6 ospf** [*process-id*]**process**

Displaying

Description	Command
Displays the OSPF process configurations.	**show ipv6 ospf** [*process-id*]
Displays information about the OSPF LSDB.	**show ipv6 ospf**[*process- id*] **database**[*lsa-type* [**adv-router***router-id*]]
Displays OSPF-enabled interfaces.	**show ipv6 ospf** [*process-id*] **interface** [*interface-type interface-number* | **brief**]

Description	Command
Displays the OSPF neighbor list.	**show ipv6 ospf**[*process- id*] **neighbor**[*interface-type interface-number*[**detail**]| *neighbor-id* |**detail**]
Displays the OSPF routing table.	**show ipv6 ospf** [*process-id*] **route**[**count**]
Displays the summarized route of OSPF redistributed routes.	**showipv6ospf**[*process-id*]**summary-prefix**
Displays the OSPF network topology information.	**show ipv6 ospf**[*process- id*] **topology** [**area***area-id*]
Displays OSPF virtual links.	**show ipv6 ospf** [*process-id*] **virtual-links**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs OSPF events.	**debug ipv6 ospf events** [**abr**|**asbr**|**os**| **router**| **vlink**]
Debugs OSPF interfaces.	**debug ipv6 ospf ifsm** [**events**|**status**|**timers**]
Debugs OSPF neighbors.	**debug ipv6 ospf nfsm** [**events** | **status** | **timers**]
Debugs the OSPF NSM.	**debug ipv6 ospf nsm** [**interface** | **redistribute** | **route**]
Debugs OSPF LSAs.	**debug ipv6 ospf lsa** [**flooding** | **generate** | **install** | **maxage** | **refresh**]
Debugs OSPF packets.	**debug ipv6 ospf packet** [**dd**|**detail**|**hello**|**ls-ack**|**ls-request**|**ls-update**|**recv**|**send**]
Debugs OSPF routes.	**debug ipv6 ospf route** [**ase** | **ia** | **install** | **spf** | **time**]

4 Configuring IS-IS

4.1 Overview

Intermediate System to Intermediate System (IS-IS) is an extensible, robust, and easy-to-use Interior Gateway Protocol (IGP) for route selection and applicable to an IP-ISO CLNS dual environment network (ISO CLNS is short for International Organization for Standardization Connectionless Network Service).

IS-IS has the common characteristics of a link state protocol. It sends Hello packets to discover and maintain neighbor relationships, and sends Link State Protocol Data Units (LSPs) to neighbors to advertise its link state. IS-IS supports Level-1 routing and Level-2 routing. All devices at the same Level maintain the same Link State Database (LSDB), which stores the LSPs generated by the devices to notify each other of the Level's network topology. Each device uses the Dijkstra Shortest Path First (SPF) algorithm to perform best-route calculation, path selection, and fast convergence.

Protocols and Standards

- RFC1142: OSI IS-IS Intra-domain Routing Protocol
- RFC1195: Use of OSI IS-IS for routing in TCP/IP and dual environments
- RFC3786: Extending the Number of Intermediate System to Intermediate System (IS-IS) Link State PDU (LSP) Fragments Beyond the 256 Limit
- RFC3373: Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies
- RFC3358: Optional Checksums in Intermediate System to Intermediate System (ISIS)
- RFC3784: Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)
- RFC2763: Dynamic Hostname Exchange Mechanism for IS-IS
- RFC6119(draft-ietf-isis-ipv6-te-00): IPv6 Traffic Engineering in IS-IS
- RFC 2966: Domain-wide Prefix Distribution with Two-Level IS-IS

4.2 Applications

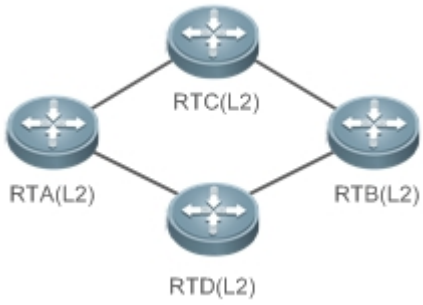
Application	Description
Planar Topology	A planar topology is applicable to a small-scale network. At the initial stage of large-scale network construction, core devices are deployed to form an area based on a planar topology.
Hierarchical Topology	A hierarchical topology is applicable to a large-scale network with frequent link flapping.

4.2.1 Planar Topology

Scenario

A planar topology is formed by devices in the same area. See Figure 4- 1.

Figure 4- 1 Planar Topology



Deployment

- To facilitate future extension and reduce device burden, configure the devices in a planar topology as Level-2 devices.

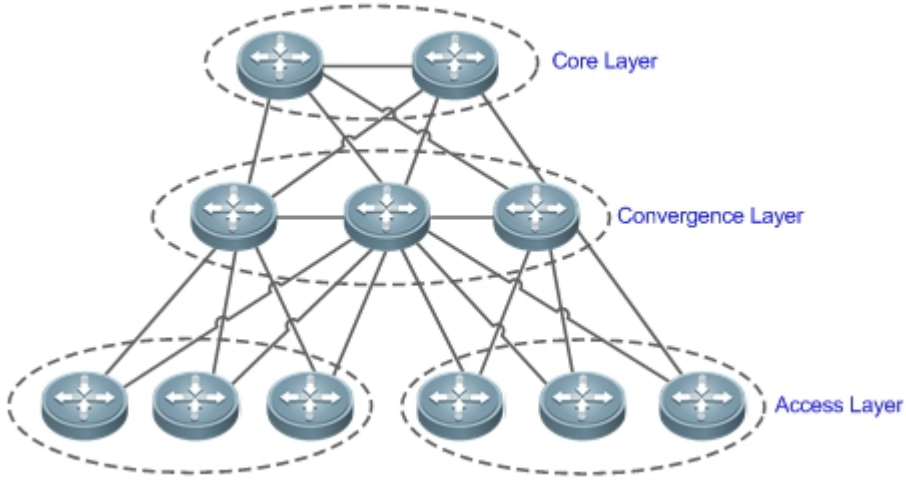
4.2.2 Hierarchical Topology

Scenario

A hierarchical topology divides the network into the core layer, convergence layer, and access layer. See Figure 4- 2.

- Route summarization at the convergence layer is facilitated by address planning.
- When primary and secondary routes exist, devices at the convergence layer leak Level-2 routes to Level-1 areas.

Figure 4- 2 Hierarchical Topology



Remarks	Devices at the core layer must be connected consecutively.

Deployment

- Design the network topology starting from the core layer.
- Configure devices at the core layer as Level-2 devices.
- Configure devices at the convergence layer as Level-1/Level-2 devices.
- Configure devices at the access layer as Level-1 devices.

4.3 Features

Basic Concepts

End System (ES)

An ES is a non-router device, for example, a host.

Intermediate System (IS)

An IS is a router, which is the basic unit used to transmit routing information and generate routes in IS-IS.

End System to Intermediate System Routing Exchange Protocol (ES-IS)

ES-IS is the protocol used for communication between ESs and ISs in Open System Interconnection (OSI) to dynamically discover Level-2 neighbor relationships.

Domain

A set of ISs in the same routing domain (RD) use the same routing protocol to exchange routing information.

Area

An RD can be divided into multiple areas.

Complete Serial Number PDU (CSNP)

CSNPs are sent by a Designated Intermediate System (DIS) every 10s to synchronize link states in a broadcast network.

Partial Sequence Number PDU (PSNP)

PSNPs are sent by a point-to-point (P2P) link to confirm LSPs, or request LSPs in a broadcast network.

Connectionless Network Protocol (CLNP)

CLNP is an OSI protocol used to transmit data and error messages at the network layer. It is similar to the IP protocol.

Connectionless Network Service (CLNS)

The CLNS is a type of unreliable connection and requires no circuit setup before data transmission.

Designated Intermediate System (DIS)

Similar to a DIS router (DR) in Open Shortest Path First (OSPF), a DIS propagates LSPs to other machines in a Local Area Network (LAN). Neighbor relationships are established not only between DISs and other machines but also between those machines. This characteristic is not possessed by OSPF.

Hello Packet

Hello packets are used to establish and maintain neighbor relationships.

LSP

LSPs describe link states, similar to link-state advertisement (LSA) in OSPF, but the former do not depend on TCP/IP information. LSPs are classified into Level-1 LSPs and Level-2 LSPs, depending on different route types.

Network Selector (NSEL)

An NSEL (sometimes referred to as SEL) specifies the target network-layer protocol service. It is similar to the TCP/UDP port for the Upper Layer Service in the IP protocol. In IS-IS, SEL is typically set to 00 to indicate a device.

Network Service Access Point (NSAP)

An NSAP is the CLNS complete address, including the OSI address and high-layer processes. It consists of an area ID, a system ID, and SEL. When SEL is set to 00, the NSAP is a Network Entity Title (NET), similar to an IP address plus a protocol number.

Sub-Network Point of Attachment (SNPA)

An SNPA provides physical connections and network-layer services. It is similar to a MAC address used in the IP protocol, a Data Link Connection Identifier (DLCI) used by frame relay (FR), or High-Level Data Link Control (HDLC) in a wide area network (WAN).

Level-1 Route

A Level-1 route is an intra-area route that only receives relevant information within the area. To reach other areas, you need to store in Level-1 a default route destined for the closest Level-2.

Level-2 Route

A Level-2 route is an inter-area backbone route. Level-1 and Level-2 cannot be connected directly.

Level-1/Level-2 Route

A Level-1/Level-2 route is a border route connecting a Level-1 route and a Level-2 route. It maintains two databases for the Level-1 and Level-2 routes respectively. It is similar to an area border router (ABR) in OSPF.

Pseudonode

A pseudonode identifies a broadcast subnet (LAN) and allows a broadcast medium to work as a virtual device, which has a route as its interface. The route-pseudonode relationship is managed by a DIS.

Network Entity Title (NET)

A NET is part of an OSI address and describes the area ID and system ID, but it does not define the NSEL, which is contained in the NSAP of the specified system.

Circuit

Circuit is an interface-related term used in IS-IS. Whereas NSAP and NET indicate whole devices, a circuit indicates an interface. The circuit ID of a P2P interface is one byte long. For example, the circuit ID of HDLC is 0x00. In a broadcast network (for example, a LAN), the circuit ID is seven bytes long, including the system ID, for example, 1921.6800.0001.01.

For details about terms related to IS-IS, see ISO 10589 and RFC1195.

Overview

Feature	Description
IS-IS Network Hierarchy	An IS-IS network is divided into Level-1 and Level-2. The nodes on which devices exchange information in the same area form one Level (Level-1).

Feature	Description
IS-IS Address Coding Mode	An IS-IS address is called a NET, which consists of an area ID, a system ID, and an NSAP identifier.
IS-IS Packet Types	There are three types of IS-IS packets: LSP, IS-IS Hello packet (IIH PDU), and serial number packet (SNP) classified into CSNP and PSNP.
DIS Election	A DIS simulates multiple access links as a pseudonode and generates LSPs for the pseudonode. The pseudonode sets up a relationship with each device in the local network and forbids direct communication between the devices.
IS-IS Supported TLV Types	IS-IS supports 21 types of Type-Length-Value (TLV).
LSP Fragment Extension	IS-IS floods LSPs to advertise link states. The size of an LSP is limited by the Maximum Transmission Unit (MTU) size of the link. When the content to be advertised exceeds one LSP, IS-IS will create LSP fragments to carry new link state information.
IS-IS VRF	VPN Routing and Forwarding (VRF) is mainly used for local routing and packet separation. It avoids route conflict caused by use of the same prefix by multiple VPNs.
IS-IS MTR	Multi-topology Routing (MTR) is mainly used to separate IPv4 unicast route calculation and IPv6 unicast route calculation based on topologies.
IS-IS Neighbor	Conditions for establishing IS-IS neighbor relationships under different configurations.

4.3.1 IS-IS Network Hierarchy

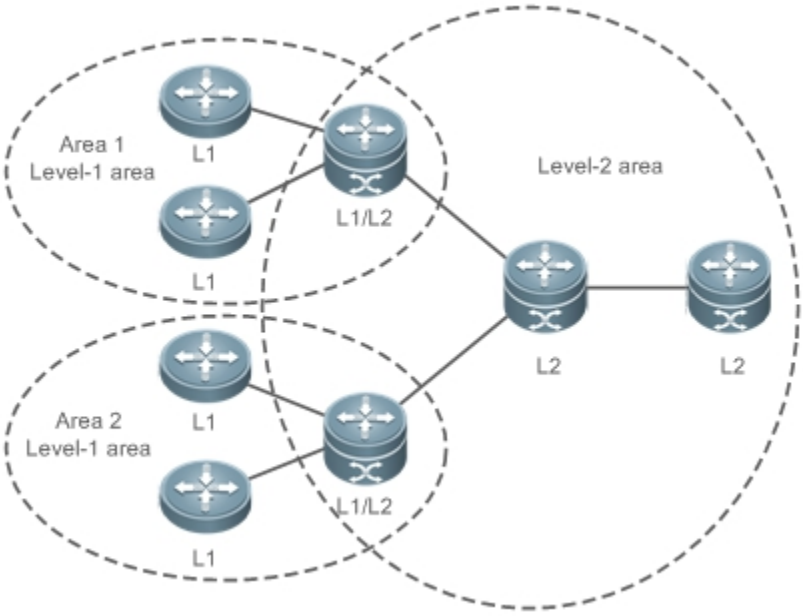
An IS-IS network is divided into Level-1 and Level-2. The nodes on which devices exchange information in the same area form one Level (Level-1).

Working Principle

All devices in an area know the area's network topology and exchange data within the area. A Level-1/Level-2 device is a border device that belongs to different areas and provides inter-area connections. Areas are connected by Level-2 devices. The border devices in various areas form a Level-2 backbone network for inter-area data exchange.

Level-1 devices are only interested in the local area's topology, including all nodes in the local area and the next-hop devices destined for the nodes. Level-1 devices access other areas through Level-2 devices and forward packets from a target network outside of the local area to the closest Level-2 device.

Figure 4- 3 IS-IS Network Topology



Related Configuration

Setting the Circuit Type of an IS-IS Interface

By default, **circuit-type** is set to Level-1/Level-2.

Run the **isis circuit-type** command to change the Level of an interface.

If **circuit-type** is set to Level-1 or Level-2-only, IS-IS will only send PDUs of the corresponding Level.

Specifying the IS-IS Level

By default, **is-type** is set to Level-1/Level-2 if no IS-IS instance runs at Level-2 (including Level-1/Level-2). **is-type** is set to Level-1 if there are IS-IS instances running at Level-2 (including Level-1/Level-2).

Run the **is-type** command to specify the Level at which IS-IS will run.

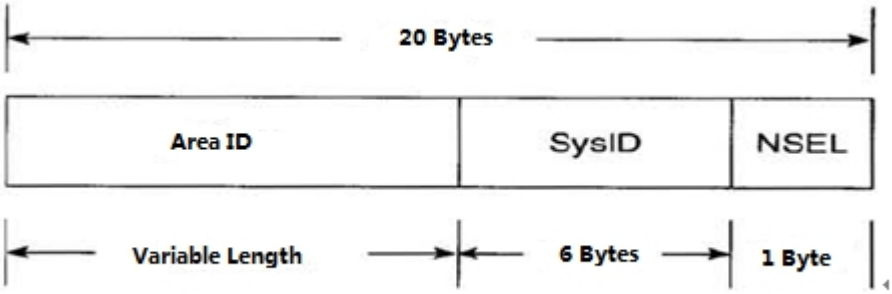
Changing the **is-type** value will enable or disable the routes of a certain Level. A device can have only one instance running at Level-2 (including Level-1/Level-2).

4.3.2 IS-IS Address Coding Mode

An IS-IS address is called a NET, which consists of an area ID, a system ID, and an NSAP identifier, ranging from eight to 20 bytes.

Working Principle

Figure 4- 4 NET Address Format



- The area ID identifies the RD length in an area and is fixed relative to the RD. It ranges from one to 13 bytes.
- The system ID is unique in an autonomous system (AS).
- The NSAP is a network selector and sometimes called SEL. In IS-IS, SEL is typically set to 00 to indicate a device.

Related Configuration

Configuring a NET Address in IS-IS

By default, no NET address is configured in IS-IS.

Run the **net** command to configure a NET address in IS-IS.

The command configures an area ID and a system ID in IS-IS. Different NET addresses must have the same system ID.

4.3.3 IS-IS Packet Types

There are three types of IS-IS packets:

- LSP
- IIH PDU
- SNP (classified into CSNP and PSNP)

Working Principle

There are three types of IS-IS packets:

- LSP

LSPs are used to transmit link state records within an area and are classified into Level-1 LSPs and Level-2 LSPs. LSPs are only flooded to the corresponding Level.

- IIH PDU

IIH PDUs are used to maintain neighbor relationships. They carry multicast MAC addresses used to determine whether other systems run IS-IS.

- SNP (classified into CSNP and PSNP)

CSNPs are used for LSDB synchronization. By default, a DIS sends a CSNP every 10s in a broadcast network. In a P2P network, a CSNP is sent only after a neighbor relationship is established.

PSNPs are also used for LSDB synchronization.

Related Configuration

Configuring the LSP Interval on an IS-IS Interface

By default, the LSP interval is 33 ms. If no Level is specified, the interval takes effect for Level-1 and Level-2 LSPs.

Run the **isis lsp-interval** command to configure the LSP interval on an IS-IS interface, in the unit of seconds.

The command changes the LSP interval.

Configuring the Hello Packet Interval on an IS-IS Interface

By default, the Hello packet interval is 10s for Level-1 and Level-2.

Run the **isis hello-interval** command to configure the Hello packet interval on an IS-IS interface, in the unit of seconds.

The command changes the Hello packet interval. A DIS sends Hello packets at a frequency three times that by non-DIS devices in a broadcast network. If an IS is elected as the DIS on the interface, by default, the interface sends a Hello packet every 3.3s.

Configuring the Minimum PSNP Interval

By default, the minimum PSNP interval is not configured, and the default interval 2s takes effect for Level-1 and Level-2 PSNPs.

Run the **isis psnp-interval** command to configure the minimum PSNP interval, in the unit of seconds.

PSNPs are mainly used to request LSPs that are absent locally or respond to received LSPs (in a P2P network). The PSNP interval should be minimized. If many LSPs exist and the device performance is low, you can increase the PSNP interval and LSP retransmission interval to reduce the device burden.

Configuring the CSNP Broadcast Interval on an IS-IS Interface

By default, CSNPs are sent at 10s intervals in a broadcast network. No CSNPs are sent in a P2P network. When you configure a new CSNP interval without Level-1 or Level-2 specified, the interval takes effect for Level-1 and Level-2 CSNPs.

Run the **isis csnp-interval** command to specify the CSNP broadcast interval on an IS-IS interface, in the unit of seconds.

The command changes the CSNP interval. By default, a DIS sends a CSNP every 10s in a broadcast network. In a P2P network, a CSNP is sent only after a neighbor relationship is established. An interface set to **mesh-groups** can be configured to periodically send CSNPs. No CSNPs are sent if the CSNP interval is set to 0.

4.3.4 DIS Election

A DIS is a designated device in a broadcast network and works like a DR in OSPF.

A pseudonode is generated by a DIS and sets up a relationship with each device in the local network.

Working Principle

A DIS simulates multiple access links as a pseudonode and generates LSPs for the pseudonode. The pseudonode sets up a relationship with each device in the local network and forbids direct communication between the devices. A broadcast subnet and a non-broadcast multiple access (NBMA) network are considered as pseudonodes externally. Non-DIS devices report their link states to the DIS in the same network, and the DIS maintains the link states reported by all ISs in the network. Like DR election in OSPF, a DIS is elected to reduce unnecessary neighbor relationships and route information exchanges.

DIS election in IS-IS is preemptive. The election result can be manually controlled through interface priority configuration. The device with a higher interface priority is more likely to be elected as the DIS.

Related Configuration

Configuring the Priority for DIS Election in a LAN

By default, Priority 64 takes effect for Level-1 and Level-2.

Run the **isis priority** command to configure the priority for DIS election in a LAN.

The command changes the priority carried in Hello packets in a LAN. The device with a lower priority is less likely to be elected as the DIS. The command is invalid on a P2P network interface. The **no isis priority** command, with or without parameters, restores the priority to its default value. To change the configured priority, run the **isis priority** command with the priority specified to overwrite the existing configuration, or you can first restore the priority to its default value and then configure a new priority.

4.3.5 IS-IS Supported TLV Types

IS-IS supports 26 types of TLV.

Working Principle

The following table lists the IS-IS supported TLV types:

TLV Code	Description
Code = 1	Area ID
Code = 2	Priority of an IS neighbor
Code = 3	ES neighbor
Code = 6	MAC address of an IS neighbor
Code = 8	Filling field
Code = 9	LSP entity
Code = 10	Verification field
Code = 14	Size of the source LSP buffer
Code = 22	Extended IS reachability
Code = 128	IP internal reachability information
Code = 129	Supported protocol
Code = 130	IP external reachability information
Code = 131	Inter-domain routing protocol information
Code = 132	IP interface address
Code = 133	Verification information
Code = 135	Extended IP reachability TLV
Code = 137	Dynamic host name
Code = 222	Multi-Topology (MT) IS reachability
Code = 229	MT TLV
Code = 211	GR

TLV Code	Description
Code=232	IPv6 interface
Code = 235	IPv4 MT IP reachability TLV
Code =236	IPv6 IP reachability TLV
Code = 237	IPv6 MT IP reachability TLV
Code = 240	P2P three-way handshake TLV

Related Configuration

Configuring the Neighbor Detection Protocol Carried in Hello Packets

By default, neighbor detection is enabled.

Run the **adjacency-check** command to configure the neighbor detection protocol carried in Hello packets.

4.3.6 LSP Fragment Extension

IS-IS floods LSPs to advertise link states. The size of an LSP is limited by the MTU size of the link. When the content to be advertised exceeds one LSP, IS-IS will create LSP fragments to carry new link state information. According to ISO standards, an LSP fragment is identified by a one-byte LSP number. An IS-IS device can generate up to 256 LSP fragments.

Working Principle

The 256 LSP fragments are insufficient in any of the following situations:

8. New applications (such as traffic engineering [TE]) extend new TLV or Sub-TLV.
9. The network is expanded continuously.
10. Routes with reduced granularity are advertised, or other routes are redistributed to IS-IS.

After LSP fragments are used up, new routing information and neighbor information will be discarded, causing network exceptions such as routing black holes or loops. LSP fragments must be extended to carry more link state information, thus ensuring normal network operation.

You can configure an additional system ID and enable fragment extension to allow IS-IS to advertise more link state information in extended LSP fragments. Each virtual system can be considered as a virtual device that establishes a neighbor relationship (with the path value being 0) with the originating system. Extended LSPs are published by the neighbor of the originating system, that is, the virtual system.

The following terms are related to fragment extension:

Normal System ID

The system ID defined by ISO is used to establish neighbor relationships and learn routes. It is further defined as the normal system ID in order to be distinguished from the additional system ID introduced to fragment extension.

Additional System ID

The additional system ID is configured by an administrator to generate extended LSPs. The additional system ID shares the usage rules of the normal system ID (for example, the additional system ID must be unique in the entire area), except that the additional system ID is not carried in Hello packets for neighbor relationship establishment.

Originating System

An originating system is an IS-IS-enabled device and maps a virtual system identified by the additional system ID.

Virtual System (Virtual IS)

A virtual system is identified by the additional system ID and used to generate extended LSPs. The virtual system concept is proposed by RFC for distinguishing from the originating system concept. Each virtual system can generate up to 256 LSP fragments. The administrator can configure multiple additional system IDs (virtual systems) to generate more LSP fragments.

Original LSP

An original LSP is the LSP whose system ID contained in the LSP ID is a normal system ID. Original LSPs are generated by an originating system.

Extended LSP

An extended LSP is the LSP whose system ID contained in the LSP ID is an additional system ID. Extended LSPs are generated by a virtual system.

Related Configuration

Enabling Fragment Extension

By default, fragment extension is disabled. If you do not specify a Level when enabling fragment extension, it will take effect for Level-1 and Level-2 LSPs.

Run the **lsp-fragments-extend** command to enable fragment extension.

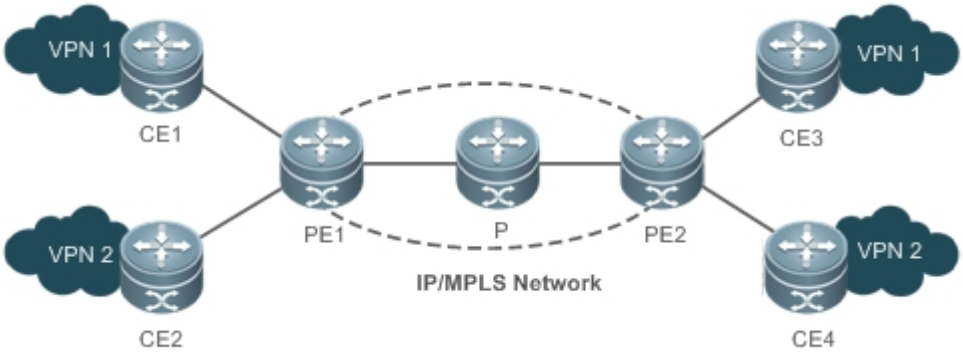
There are up to 256 LSP fragments. When the fragments are used up, subsequent link state information, including neighbor information and IP route information, will be discarded, causing a network exception. To solve this problem, enable fragment extension at the specified Level and configure an additional system ID by using the **virtual-system** command.

4.3.7 IS-IS VRF

VRF is mainly used for local routing and packet separation. It avoids route conflict caused by use of the same prefix by multiple VPNs. IPv4 VPN and IPv6 VPN combine Multiprotocol Label Switching (MPLS) advantages in terms of Quality of Service (QoS) and security assurance, and are the primary solutions for interconnecting the geographically different office branches of an enterprise or industry user.

Working Principle

Figure 4- 5 Separation of Different VPNs by VRF Tables Configured on Provider Edge (PE) Devices



In Figure 4- 5, the following configuration requirements exist: Configure the two sites (CE1 and CE3) in VPN1 to access each other and the two sites (CE2 and CE4) in VPN2 to access each other, and forbid access between the sites in VPN1 and those in VPN2, because VPN1 and VPN2 belong to different customers or departments and may have identical IP addresses.

The customer edge (CE) devices connect the customer network to the PEs to exchange VPN routing information with the PEs, that is, advertise local routes to the PEs and learn remote routes from the PEs.

Each PE learns routes from directly connected CEs and exchanges the learned VPN routes with the other PE through the Border Gateway Protocol (BGP). The PEs provide access to the VPN service.

The Provider (P) device in the Service Provider (SP) network is not directly connected to the CEs. The P device only needs the MPLS forwarding capability and does not maintain VPN information.

The IS-IS protocol running between the PEs and CEs requires the VRF capability to separate routing information between VPN1 and VPN2. That is, IS-IS only learns routes through VRF.

Related Configuration

↘ Binding an IS-IS Instance with a VRF Table

By default, an IS-IS instance is not bound with any VRF table.

Run the **VRF** command to bind an IS-IS instance with a VRF table.

Note the following constraints or conventions for the binding operation:

- The IS-IS instances bound with the same non-default VRF table must be configured with different system IDs. The IS-IS instances bound with different VRF tables can be configured with the same system ID.
- One IS-IS instance can be bound with only one VRF table, but one VRF table can be bound to multiple IS-IS instances.
- When the VRF table bound to an IS-IS instance is changed, all IS-IS interfaces associated with the instance will be deleted. That is, the **ip** (or **ipv6**) **router isis** [*tag*] interface configuration and the redistribution configuration in routing process mode will be deleted.

4.3.8 IS-IS MTR

IS-IS MTR is an extended feature used to separate IPv4 unicast route calculation and IPv6 unicast route calculation based on topologies. It complies with the specification of IS-IS MT extension defined in RFC 5120. New TLV types are introduced to IIH PDUs and LSPs to transmit IPv6 unicast topology information. One physical network can be divided into an IPv4 unicast logical topology and an IPv6 unicast logical topology. The two topologies perform SPF calculation separately and maintain independent IPv4 and IPv6 unicast routing tables respectively. In this way, IPv4 unicast service traffic and IPv6 unicast service traffic are forwarded by different paths. The IS-IS MTR

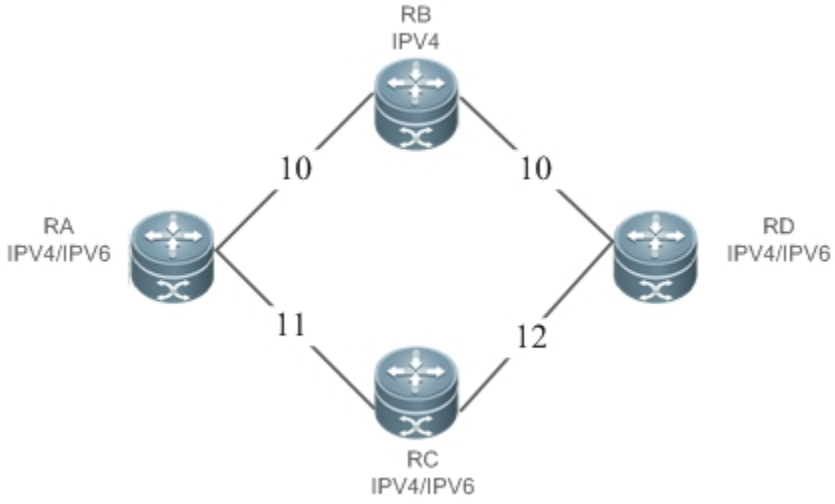
technique helps users deploy IPv6 unicast networks without the constraint on consistency between IPv4 and IPv6 unicast topology information.

IS-IS MTR is derived from IS-IS MT, which is used to separate IPv4 and IPv6 unicast topologies, unicast and multicast topologies, and topologies using different protocol stacks (such as IPv4 and Pv6). IS-IS MTR separates IPv4 and IPv6 unicast topologies based on IS-IS MT.

Working Principle

Figure 4- 6 shows a typical networking application. The following implementation requirements exist: Deploy an IPv6 unicast topology in incremental mode, and upgrade some devices to support IPv4 and IPv6 dual protocol stacks while keeping other IPv4-enabled devices unchanged.

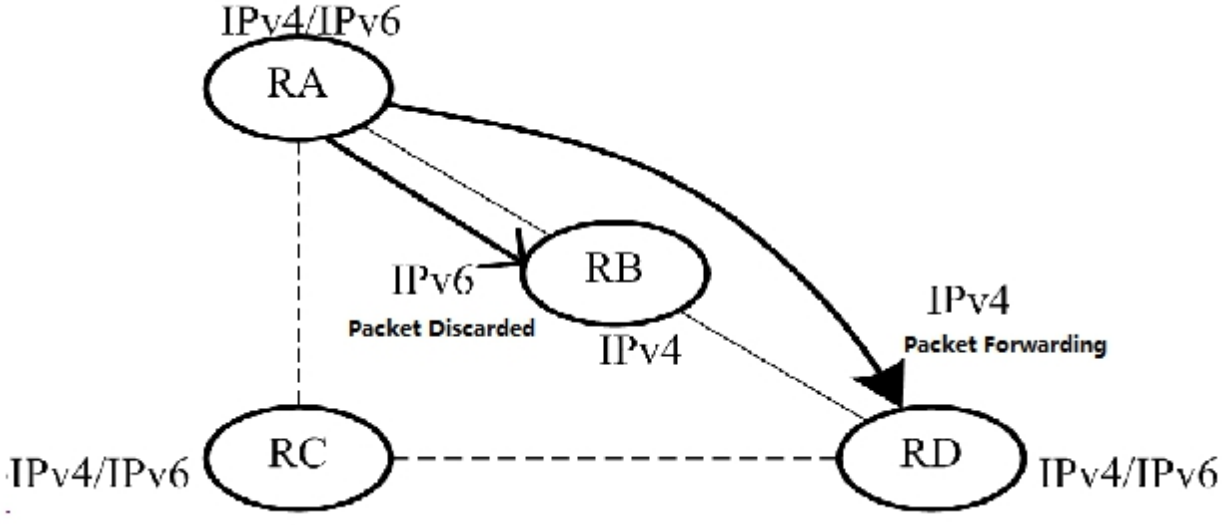
Figure 4- 6 Physical Topology for IPv4-IPv6 Hybrid Deployment



In Figure 4- 6, each link is marked by a number indicating its metric. RB only supports the IPv4 protocol stack, whereas other devices support IPv4 and IPv6 dual protocol stacks.

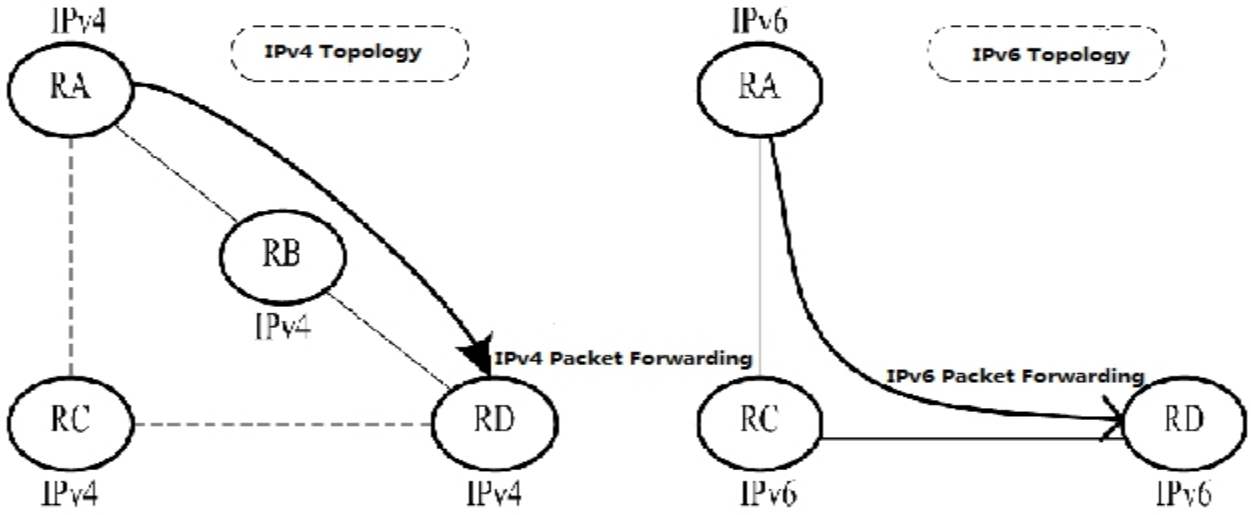
The networking constraint on consistency between IPv4 and IPv6 unicast topologies must be canceled to retain the use of RB; otherwise, RB cannot establish a neighbor relationship with RA or RD, which will cause new problems.

Figure 4- 7 IPv4-IPv6 Hybrid Topology



In Figure 4- 7, without IS-IS MTR support, the SPF calculations performed by RA, RB, RC, and RD only take into account the single hybrid topology. The calculated shortest path is RA -> RB -> RD, with the overhead being 20. RB will discard IPv6 packets because it does not support IPv6.

Figure 4- 8 Separation of IPv4 and IPv6 Topologies



In Figure 4- 8, the IS-IS MTR technique is used to separate IPv4 and IPv6 unicast topologies. RA, RB, RC, and RD establish neighbor relationships based on the IPv4 unicast topology and IPv6 unicast topology respectively. The left part shows the IPv4 topology formed by IPv4-enabled routers. The calculated IPv4 shortest path is RA -> RB -> RC, which realizes IPv4 packet forwarding. The right part shows the IPv6 topology formed by IPv6-enabled routers. The calculated IPv6 shortest path is RA -> RC -> RD, which realizes IPv6 packet forwarding.

IS-IS MTR must be deployed to avoid routing black holes when some devices support only one protocol. IS-IS MTR is not required when all devices support IPv4 and IPv6 dual protocol stacks.

- Deployment of a new network: IS-IS MTR is not required when devices only support the IPv4 protocol stack. For devices that only support the IPv6 protocol stack or devices that support IPv4 and IPv6 dual protocol stacks, enable the MT mode of IS-IS MTR. You are advised not to enable Multi-Topology Transition (MTT); otherwise, loops may occur.

- Reconstruction of an existing network with devices supporting only one protocol stack: Enable the MTT mode of IS-IS MTR on devices that support IPv4 and IPv6 dual protocol stacks in sequence (starting from the device closest to a device supporting only one protocol stack in the network topology). After the MTT mode is enabled on all new devices, switch the MTT mode to the MT mode on these devices in sequence (starting from the device farthest from a device supporting only one protocol stack in the network topology).

Related Configuration

Enabling MTR for IS-IS Instances

By default, IS-IS instances are not enabled with MTR.

Run the **multi-topology** command to configure IS-IS to support IPv6 unicast topologies. After that, IPv4 and IPv6 unicast routes in IS-IS will be calculated based on different topologies.

Note the following constraints or conventions when you use the **multi-topology** command:

1. Set **metric-style** to **Wide** or **Transition** before you run the command.

2. The MTR feature will be disabled if **metric-style** is set to **Narrow** or only one Level is configured to support the Wide or Transition mode.

4.3.9 IS-IS Neighbor

When IS-IS MTR is not configured, the following conditions must be met for two routing devices to establish a neighbor relationship:

- The interface addresses on both routing devices are in the same network segment.
- The interface Levels on both routing devices match.
- The routing devices are authenticated by each other.
- The routing devices support the same protocol.

When IS-IS MTR is configured, the following conditions must be met for routing devices to establish a neighbor relationship:

- The interface addresses on both routing devices are in the same network segments.
- The interface Levels on both routing devices match.
- The routing devices are authenticated by each other.
- The routing devices have at least one consistent MT ID when P2P links are configured.
- There are no constraints on the MT IDs that the routing devices support when LAN links are configured.

4.4 Configuration

Configuration	Description and Command	
Enabling IS-IS	(Mandatory) It is used to enable IS-IS on specified interfaces. You need to create an IS-IS routing process in advance.	
	router isis [*tag*]	Starts an IS-IS routing process. *tag* indicates the process name.
	net *areaAddress.SystemId.00*	Configures a NET address in IS-IS.
	ip router isis [*tag*]	Enables IS-IS on an interface. *tag* indicates the name of the IS-IS routing process.
Configuring IS-IS Hello Packets	(Optional) It is used to configure the IS-IS Hello packet holdtime.	
	isis hello-interval { *interval* | **minimal** } [**level-1** | **level-2**]	Configures the Hello packet interval on an interface. The value range is 1 to 65,535, in the unit of seconds.
	isis hello-multiplier *multiplier-number* [**level-1** | **level-2**]	Configures the Hello packet holdtime multiplier on an IS-IS interface. The value range is 2 to 100. The default value is 3.
Configuring IS-IS LSPs	(Optional) It is used to perform time-related LSP configuration, determine whether to ignore LSP checksum errors, and enable/disable LSP fragment extension.	
	isis lsp-interval *interval* [**level-1** | **level-2**]	Configures the minimum LSP interval on an interface. The value range is 1 to 4,294,967,295, in the unit of milliseconds.

Configuration	Description and Command	
	isis retransmit-interval *interval* [**level-1** | **level-2**]	Configures the LSP retransmission interval by P2P links on an interface. The value range is 0 to 65,535, in the unit of seconds.
	lsp-gen-interval [**level-1** | **level-2**] *maximum-interval*	Configures LSP generation cycle. *maximum-interval*: Indicates the maximum interval for generating two consecutive LSP packets. The value range is **1** to **65535** (in seconds). The default value is **5**. *initial-interval*: Indicates the waiting time for generating an LSP packet for the first time. The value range is **0** to **60000** (in milliseconds). The default value is **50**. *hold-interval*: Indicates the minimum interval for generating an LSP packet for the second time. The value range is **10** to **60000** (in milliseconds). The default value is **200**.
	lsp-refresh-interval *interval*	Configures the LSP refresh interval. The value range is 1 to 65,535, in the unit of seconds.
	max-lsp-lifetime *value*	Configures the LSP lifetime. The value range is 1 to 65,535, in the unit of seconds.
	ignore-lsp-errors	Configures to ignore LSP checksum errors.
	lsp-fragments-extend [**level-1** | **level-2**] [**compatible rfc3786**]	Enables fragment extension.
	virtual-system *system-id*	Configures an additional system ID.
Configuring IS-IS SNPs	(Optional) It is used to configure the CSNP broadcast interval.	
	isis csnp-interval *interval* [**level-1** | **level-2**]	Configures the CSNP interval on an interface. The value range is 0 to 65,535, in the unit of seconds. The default value is 10s. No CSNPs are sent if the CSNP interval is set to 0.
Configuring the IS-IS Level Type	(Optional) It is used to configure the system type or interface circuit type in IS-IS.	
	is-type { **level-1** | **level-1-2** | **level-2-only** }	Configures the system type.
	isis circuit-type { **level-1** | **level-1-2** | **level-2-only** [**external**] }	Configures the interface circuit type.
Configuring IS-IS Authentication	(Optional) It is used to configure interface authentication, area authentication, and RD authentication.	

Configuration	Description and Command	
	isis password [**0** | **7**] *password* [**send-only**] [**level-1** | **level-2**]	Configures the password for plaintext authentication of Hello packets on an interface. When **send-only** is included, the authentication password is only used to authenticate sent Hello packets. Received Hello packets are not authenticated. If no Level is specified, the configured authentication and password take effect for all Levels. This command does not take effect if the **isis authentication mode** command is executed. Both commands are used to configure IS-IS interface authentication, but the **isis password** command has a lower priority. Before you run the **isis password** command, delete the **isis authentication mode** command configuration.
	isis authentication mode { **text** | **md5** } [**level-1** | **level-2**]	Specifies authentication as plaintext or MD5. If no Level is specified, the authentication mode takes effect for all Levels. If you use this command after the **isis password** *password* [level-1 | level-2] command is executed, the previous command configuration will be overwritten. Both commands are used to configure IS-IS interface authentication, but the **isis authentication mode** command has a higher priority.
	isis authentication key-chain *name-of-chain* [**level-1** | **level-2**]	Configures the password for interface authentication. If no Level is specified, the configured key chain takes effect for all Levels. This command must be used with the **isis authentication mode** command to configure IS-IS interface authentication.

Configuration	Description and Command	
	isis authentication send-only [**level-1** | **level-2**]	(Optional) Specifies that interface authentication is performed only on sent packets. Received packets are not authenticated. If no Level is specified, the send-only authentication mode takes effect for all Levels. This command is used to avoid network flapping caused by a temporary authentication failure when IS-IS authentication is configured. Before you deploy IS-IS authentication in the entire network, run the **isis authentication mode** { **text** | **md5** } [**level-1** | **level-2**] and **isis authentication key-chain** *name-of-chain* [**level-1** | **level-2**] commands on each device. After that, run the **no isis authentication send-only** command to restore the authentication of received packets. This realizes smooth authentication deployment and avoids network flapping.
	area-password [**0** | **7**] *password* [**send-only**]	Configures the password for area (Level-1) plaintext authentication. When **send-only** is included, the authentication password is only used to authenticate sent packets. Received packets are not authenticated. This command does not take effect if the **authentication mode** command is executed. Both commands are used to configure IS-IS area authentication, but the **area-password** command has a lower priority. Before you run the **area-password** command, delete the **authentication mode** command configuration.
	authentication mode { **text** | **md5** } **level-1**	Specifies the IS-IS area authentication mode. If you use this command after the **area-password** *password* command is executed, the previous command configuration will be overwritten. Both commands are used to configure IS-IS area authentication, but the **authentication mode** command has a higher priority.
	authentication key-chain *name-of-chain* **level-1**	Configures the key chain for IS-IS area authentication. This command must be used with the **authentication mode** command to configure IS-IS area authentication.

Configuration	Description and Command	
	authentication send-only level-1	(Optional) Specifies that IS-IS area authentication is performed only on sent packets. Received packets are not authenticated. This command is used to avoid network flapping caused by a temporary authentication failure when IS-IS authentication is configured. Before you deploy IS-IS authentication in the entire area, run the **authentication mode** { **text** | **md5** } **level-1** and **authentication key-chain** *name-of-chain* **level-1** commands on each device. After that, run the **no authentication send-only** command to restore the authentication of received packets. This realizes smooth authentication deployment and avoids network flapping.
	domain-password [**0** | **7**] *password* [**send-only**]	Configures the password for RD (Level-2) plaintext authentication. When **send-only** is included, the authentication password is only used to authenticate sent packets. Received packets are not authenticated. This command does not take effect if the **authentication mode** command is executed. Both commands are used to configure IS-IS RD authentication, but the **domain-password** command has a lower priority. Before you run the **domain-password** command, delete the **authentication mode** command configuration.
	authentication mode { **text** | **md5** } **level-2**	Specifies the IS-IS RD authentication mode. If you use this command after the **domain-password** *password* command is executed, the previous command configuration will be overwritten. Both commands are used to configure IS-IS RD authentication, but the **authentication mode** command has a higher priority.
	authentication key-chain *name-of-chain* **level-2**	Configures the password for IS-IS RD authentication. This command must be used with the **authentication mode** command to configure IS-IS RD authentication.

Configuration	Description and Command	
	authentication send-only level-2	(Optional) Specifies that IS-IS RD authentication is performed only on sent packets. Received packets are not authenticated. This command is used to avoid network flapping caused by a temporary authentication failure when IS-IS authentication is configured. Before you deploy IS-IS authentication in the entire RD, run the **authentication mode** { **text** | **md5** } **level-2** and **authentication key-chain** *name-of-chain* **level-2** commands on each device. After that, run the **no authentication send-only** command to restore the authentication of received packets. This realizes smooth authentication deployment and avoids network flapping.
Configuring IS-IS GR	(Optional) It is used to enable IS-IS GR.	
	graceful-restart	Enables the GR Restart capability on the device that works as a Restarter. By default, the GR Restart capability is enabled.
	graceful-restart grace-period *seconds*	(Optional) Configures the IS-IS GR time on the device that works as a Restarter. The default value is 300s.
	no graceful-restart helper disable	Enables the IS-IS GR Help capability on the device that works as a Helper. By default, the GR Help capability is enabled.
Configuring BFD Support for IS-IS	(Optional) It is used to enable BFD support for IS-IS.	
	bfd all-interfaces [**anti-congestion**]	Enables BFD support for IS-IS on all interfaces.
	isis bfd [**disable | anti-congestion**]	Enables or disables BFD support for IS-IS on the current interface.
Setting the IS-IS Overload Bit	(Optional) It is used to set the overload bit in LSPs.	
	set-overload-bit [**on-startup** *seconds*] [**suppress** { [**interlevel**] [**external**] }] [**level-1** | **level-2**]	Sets the overload bit.
Configuring IS-IS VRF	(Optional) It is used to bind an IS-IS instance with a VRF table.	
	vrf *vrf-name*	Binds an IS-IS instance with a VRF table.
Configuring IS-IS MTR	(Optional) It is used to calculate IPv4 and IPv6 unicast routes in IS-IS based on different topologies.	
	multi-topology [**transition**]	Configures IS-IS to support IPv6 unicast topologies.

Configuration	Description and Command	
Configuring Simple Network Management Protocol (SNMP) for IS-IS	(Optional) It is used to allow the SNMP software to perform Management Information Base (MIB) operations on IS-IS instances.	
	enable mib-binding	Performs MIB operations on the instance bound with Tag 1.
	configure terminal	Enters global configuration mode.
	snmp-server enable traps isis	Enables IS-IS trap globally.
	snmp-server host { *host-addr* | **ipv6** *ipv6-addr* } [**vrf** *vrfname*] [**traps**] [**version** { **1** | **2c** | **3** { **auth** | **noauth** | **priv** }] *community-string* [**udp-port** *port-num*]	Configures an SNMP host in global configuration mode to receive IS-IS trap messages.
	router isis	Enters IS-IS routing process configuration mode.
	enable traps all	Allows the sending of all IS-IS trap messages to the host with the IP address 10.1.1.1.
Running ISIS on Super VLAN	Optional.	
	isis subvlan [all | *vid*]	Runs ISIS on Super VLAN.
Configuring IS-IS Two-way Maintenance	Optional.	
	two-way-maintain	Enables IS-IS two-way maintenance.
.Configuring Other IS-IS Parameters	Optional.	
	maximum-paths *maximum*	Configures the maximum number of IS-IS IPv4/IPv6 equal-cost paths.
	lsp-length *receive size*	Configures the maximum length allowed for received LSPs.
	lsp-length originate *size* [**level-1** | **level-2**]	Configures the maximum length allowed for sent LSPs.
	passive-interface [**default**] { *interface-type interface-number* }	Configures a passive interface.
	isis metric *metric* [**level-1** | **level-2**]	Configures the interface metric, which is valid only when **metric-style** is set to **Narrow**.
	isis wide-metric *metric* [**level-1** | **level-2**]	Configures the interface wide-metric value, which is valid only when **metric-style** is set to **Wide**.
	isis priority *value* [**level-1** | **level-2**]	Configures the priority for DIS election on an interface.
	default-information originate [**route-map** *map-name*]	Generates a Level-2 default route, which will be advertised through LSPs. When the command includes the **route-map** option, a default route is generated only if the criteria in route-map are met.

Configuration	Description and Command	
	summary-address *ip-address net-mask* [**level-1** | **level-2** | **level-1-2**] [**metric** *number*]	Configures an IPv4 summary route.
	summary-prefix *ipv6-prefix/prefix-length* [**level-1** | **level-2** | **level-1-2**]	Configures an IPv6 summary route.
	ignore-lsp-errors	Configures to ignore LSP checksum errors.
	log-adjacency-changes	Activates logging of IS-IS neighbor relationship changes.
	redistribute	Configures route redistribution.

4.4.1 Enabling IS-IS

Configuration Effect

- Before you run IS-IS, create an IS-IS routing process in global configuration mode. You can set the **tag** parameter after the **router isis** command to name the process. You can add different tags to configure different IS-IS routing processes. The setting of the **tag** parameter is optional.
- A system ID uniquely identifies an IS in a routing AS; therefore, the system ID must be unique across the AS. In IS-IS, each area may contain one or multiple area IDs. Normally, you only need to configure one area ID. You can configure multiple area IDs to realize area division. If an IS is configured with multiple area IDs, the system IDs must be the same.
- After an interface is added to the specified IS-IS routing process, the interface will establish a neighbor relationship.

Notes

- The Level-1 IS devices in an area must be configured with the same area ID.
- The core routing table does not distinguish the routing entries generated by different IS-IS routing processes.
- The IP addresses of interfaces connected between neighbors must be in the same network segment.
- If the two IP addresses are in different network segments, a neighbor relationship cannot be established.
- If you need to add an interface to the specified IS-IS routing process, set the **tag** parameter after the **ip router isis** command to indicate the process name.
- If you run the **no ip routing** command in global configuration mode, IS-IS will disable IPv4 routing on all interfaces. That is, the **no ip router isis** [*tag*] command is automatically executed on all interfaces. Other IS-IS settings remain unchanged.
- By default, CPU protection is enabled on devices. For packets mapped to the destination group addresses (AllISSystems, AllL1ISSystems, and AllL2ISSystems) in IS-IS, there is a default limit (for example, 400 pps) on the number of packets sent to the CPU. If a device has many neighbor relationships or sends Hello packets at short intervals, the IS-IS packets that the device receives may exceed the default limit, causing frequent flapping of neighbor relationships. To solve the problem, you can use the CPU protection command in global configuration mode to increase the limit.

Configuration Steps

Starting an IS-IS Routing Process

- Mandatory.

- Perform this configuration in global configuration mode on each device, unless otherwise specified.

Configuring a NET Address in IS-IS

- Mandatory.
- Perform this configuration in IS-IS routing process configuration mode on each device, unless otherwise specified.

Enabling IS-IS on an Interfaces

- Mandatory.
- Perform this configuration in interface configuration mode on each device, unless otherwise specified.

Verification

- Check whether devices send Hello packets.
- Check whether devices establish neighbor relationships.
- Check whether devices exchange LSPs.

Related Commands

Starting an IS-IS Routing Process

Command	**router isis** [*tag*]
Parameter Description	*tag*: Indicates the name of an IS-IS instance.
Command Mode	Global configuration mode
Usage Guide	Use this command to initialize an IS-IS instance and enter IS-IS routing process configuration mode. An IS-IS instance will start running after a NET address is configured. If you set the **tag** parameter when you start an IS-IS routing process, you need to add the **tag** parameter when closing the IS-IS routing process. By default, CPU protection is enabled on devices. For packets mapped to the destination group addresses (AllISSystems, AllL1ISSystems, and AllL2ISSystems) in IS-IS, there is a default limit (for example, 400 pps) on the number of packets sent to the CPU. If a device has many neighbor relationships or sends Hello packets at short intervals, the IS-IS packets that the device receives may exceed the default limit, causing frequent flapping of neighbor relationships. To solve the problem, you can use the CPU protection command in global configuration mode to increase the limit.

Configuring a NET Address in IS-IS

Command	**net** *net-address*
Parameter Description	*net-address:* The NET address is in the format of XX.XXXX.YYYY.YYYY.YYYY.00. XX.XXXX indicates the area ID, and YYYY.YYYY.YYYY indicates the system ID.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to configure an area ID and a system ID in IS-IS.

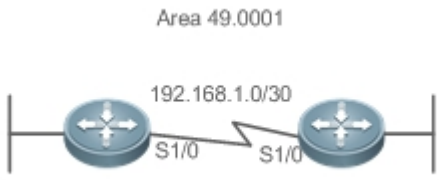
	Different NET addresses must have the same system ID.

Enabling IS-IS on an Interface

Command	**ip router isis** [*tag*]
Parameter Description	*tag*: Indicates the name of an IS-IS instance.
Command Mode	Interface configuration mode
Usage Guide	Use this command to enable an interface to participate in IS-IS IPv4 routing. Use the **no** form of this command to disable the IS-IS routing process on the interface. If you run the **no ip routing** command in global configuration mode, IS-IS will disable IPv4 routing on all interfaces. That is, the **no ip router isis** [*tag*] command is automatically executed on all interfaces. Other IS-IS settings remain unchanged.

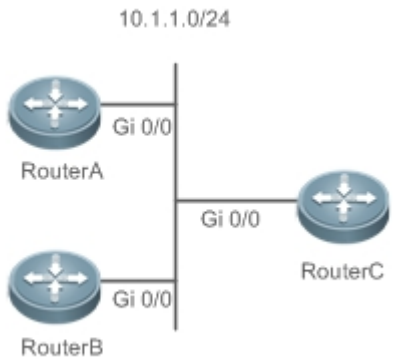
Configuration Example

Establishing a Neighbor Relationship on an IS-IS P2P Link

Scenario	Router A and Router B are connected in P2P mode.
Figure 4- 9 P2P Link Topology	
Configuration Steps	● Configure IS-IS. ● Configure Wide Area Network (WAN) interfaces.
A	A(config)# router isis A(config-router)# net 49.0001.0000.0000.0001.00 A(config)# interface Serial 1/0 A(config-if)# ip address 192.168.1.1 255.255.255.252 A(config-if)# ip router isis
B	B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00 B(config)# interface Serial 1/0 B(config-if)# ip address 192.168.1.2 255.255.255.252 B(config-if)# ip router isis

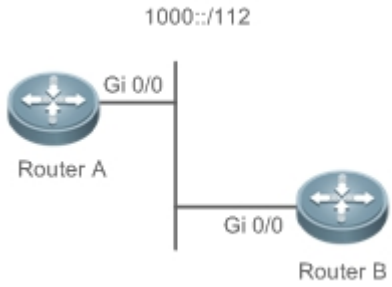
Verification	● Enable sending of Hello packets from the interface 192.168.1.1 on Router A to the interface 192.168.1.2 on Router B. ● Establish an IS-IS neighbor relationship between Router A and Router B, with the neighbor state being Up. ● Check the LSPs on Router A and Router B. The system IDs 0000.0000.0001 and 0000.0000.0002 should exist.
A	A# show isis neighbors A# show isis database detail
B	B# show isis neighbors

Establishing a Neighbor Relationship on an IS-IS Broadcast Link

Scenario	Router A, Router B, and Router C are interconnected through the Ethernet.
Figure 4-10 IS-IS Broadcast Link Topology	
Configuration Steps	● Configure IS-IS. ● Configure Ethernet interfaces.
A	A(config)# router isis A(config-router)# net 49.0001.0000.0000.0001.00 A(config)# interface GigabitEthernet 0/0 A(config-if)# ip address 10.1.1.1 255.255.255.0 A(config-if)# ip router isis
B	B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00 B(config)# interface GigabitEthernet 0/0 B(config-if)# ip address 10.1.1.2 255.255.255.0 B(config-if)# ip router isis
C	C(config)# router isis C(config-router)# net 49.0001.0000.0000.0003.00 C(config)# interface GigabitEthernet 0/0

	C(config-if)# ip address 10.1.1.3 255.255.255.0 C(config-if)# ip router isis
Verification	Enable sending of Hello packets from the interface 10.1.1.1 on Router A to the interface 10.1.1.2 on Router B and the interface 10.1.1.3 on Router C. ● Establish IS-IS neighbor relationships between Router A and Router B and between Router A and Router C, with the neighbor state being Up. ● Check the LSPs on Router A, Router B, and Router C. The system IDs 0000.0000.0001, 0000.0000.0002, and 0000.0000.0003 should exist.
A	A# show isis neighbors A# show isis database detail
B	B# show isis neighbors
C	C# show isis neighbors

Performing Simple IS-ISv6 Configuration

Scenario	Router A and Router B are connected through the Ethernet.
Figure 4- 11 IS-ISv6 Broadcast Link Topology	
Configuration Steps	● Configure IS-IS. ● Configure Ethernet interfaces.
A	A(config)# router isis A(config-router)# net 49.0001.0000.0000.0001.00 A(config)# interface GigabitEthernet 0/0 A(config-if)# ipv6 address 1000 ::1/112 A(config-if)# ipv6 router isis
B	B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00 B(config)# interface GigabitEthernet 0/0 B(config-if)# ipv6 address 1000 ::2/112

	B(config-if)# ipv6 router isis
Verification	Enable sending of Hello packets from the interface 1000 ::1 on Router A to the interface 1000 ::2 on Router B. Establish an IS-IS neighbor relationship between Router A and Router B, with the neighbor state being Up. Check the LSPs on Router A and Router B. The system IDs 0000.0000.0001 and 0000.0000.0002 should exist.
A	A# show isis neighbors A# show isis database detail
B	B# show isis neighbors

Common Errors

- The IP addresses of the interfaces connected between neighbors are not in the same network segment.
- The **ip router isis** command is not executed on interfaces.
- No NET address is configured, or different NET addresses exist at Level-1.
- **max-area-addresses** is configured differently on both sides.
- **metric-style** is configured differently on both sides.
- The interface Levels on both sides are different. One side is Level-1, whereas the other side is Level-2.
- One side is configured with the P2P mode, whereas the other side is configured with the broadcast mode.
- One side is enabled with authentication, whereas the other side is not.

4.4.2 Configuring IS-IS Hello Packets

Configuration Effect

- Configure the Hello packet interval on an interface. The value range is 1 to 65,535, in the unit of seconds.
- Configure the Hello packet holdtime multiplier on an IS-IS interface.

Notes

- You can change the Hello packet holdtime by using the **isis hello-multiplier** command or **isis hello-interval** command or both.
- By default, CPU protection is enabled on devices. For packets mapped to the destination group addresses (AllISSystems, AllL1ISSystems, and AllL2ISSystems) in IS-IS, there is a default limit (for example, 400 pps) on the number of packets sent to the CPU. If a device has many neighbor relationships or sends Hello packets at short interval, the IS-IS packets that the device receives may exceed the default limit, causing frequent flapping of neighbor relationships. To solve the problem, you can use the CPU protection command in global mode to increase the limit.

Configuration Steps

Configuring the Hello Packet Interval on an Interface

- Perform this configuration based on requirements.
- Run the **isis hello-interval** command in interface configuration mode on the desired device, unless otherwise specified.

Configuring the Hello Packet Holdtime Multiplier on an Interface

- Perform this configuration based on requirements.
- Run the **isis hello-multiplier** command in interface configuration mode on the desired device, unless otherwise specified.

Verification

- Enable Router A to send Hello packets to Router B and Router C, and capture packets to check the packet interval.
- Make Router B or Router C down. After the holdtime has elapsed, check whether the corresponding neighbor relationship on Router A is invalid.

Related Commands

Configuring the Hello Packet Interval on an Interface

Command	**isis hello-interval** { *interval* | **minimal** } [**level-1 | level-2**]
Parameter Description	*interval*: Indicates the Hello packet interval. The value range is 1 to 65,535, in the unit of seconds. The default value is 10. **minimal**: Indicates the minimum value of the holdtime, which is 1. **level-1:** Applies the setting to Level-1 Hello packets. **level-2:** Applies the setting to Level-2 Hello packets.
Command Mode	Interface configuration mode
Usage Guide	Use this command to change the Hello packet interval. The default interval is 10s. A DIS sends Hello packets at a frequency three times that by non-DIS devices in a broadcast network. If an IS is elected as the DIS on the interface, by default, the interface sends a Hello packet every 3.3s. If the keyword **minimal** is used, the Hello packet holdtime is set to 1. The Hello packet interval will be calculated based on the holdtime multiplier. If the holdtime multiplier is set to 4 and the **isis hello-interval minimal** command is executed, the Hello packet interval is equal to 1s divided by 4. The default Hello packet holdtime multiplier on an IS-IS interface is 3. The holdtime is equal to the holdtime multiplier multiplied by the packet interval. If the keyword **minimal** is used, the holdtime is set to 1. The packet interval is equal to 1 divided by the holdtime multiplier. If the holdtime multiplier is set to 4 and the **isis hello-interval minimal** command is executed, the packet interval is equal to 1 divided by 4s, which is 250 ms.

Configuring Hello Packet Holdtime Multiplier on an Interface

Command	**isis hello-multiplier** *multiplier-number* [**level-1 | level-2**]
Parameter Description	*multiplier-number*: Indicates the Hello packet holdtime multiplier. The value range is 2 to 100. The default value is 3.
Command Mode	Interface configuration mode
Usage Guide	The Hello packet holdtime is equal to the Hello packet interval multiplied by the holdtime multiplier.

Configuration Example

Configuring the Hello Packet Interval and Holdtime on an IS-IS Interface

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the Hello packet interval on an IS-IS interface. ● Configure the Hello packet holdtime multiplier on an IS-IS interface.
	A(config)# interface GigabitEthernet 0/0 A(config-if)# isis hello-interval 5 A(config-if)# isis hello-multiplier 5
Verification	Enable Router A to send Hello packets to Router B and Router C, and capture packets to check the packet interval. Make Router B or Router C down. After the holdtime has elapsed, check whether the corresponding neighbor relationship on Router A is invalid.
	A# show isis neighbor

4.4.3 Configuring IS-IS LSPs

Configuration Effect

- **isis lsp-interval**: Configures the LSP interval on an IS-IS interface.
- **isis retransmit-interval**: After a device at one end of a P2P link sends an LSP packet, if the device receives no response within a period of time, it determines that the LSP packet is lost or dropped due to an error. The device will resend the LSP packet.
- **lsp-gen-interval**: Indicates the exponential backoff algorithm of LSP packet generation. Any update to related information forming the LSP packet leads to LSP packet generation. During network flapping, LSP packets are frequently generated, which increases system resource consumption. An appropriate value can be set by running the **lsp-gen-interval** command. In this way, LSP packets are generated and advertised in time when the network is stable. When the network becomes unstable, less LSP packets are generated as the flapping continues, reducing the device consumption.
- **lsp-refresh-interval**: All current LSPs are periodically retransmitted to enable each network node to maintain the latest LSPs. The retransmission period is called the LSP refresh interval, which aims to update and synchronize LSPs in the entire area.
- **max-lsp-lifetime**: An LSP contains a field to indicate its lifetime. When a device generates an LSP, the field is set to the maximum lifetime of the LSP. After the LSP is received by the peer device, its lifetime will decrease with time. The peer device will replace the old LSP with the newly received one. If the device receives no new LSP until the existing LSP's lifetime decreases to 0, the existing LSP is still maintained in the LSDB for another 60s. If the device still receives no new LSP during this period, the existing LSP will be deleted from the LSDB. This mechanism updates and synchronizes LSPs in the entire area.
- **ignore-lsp-errors:** After receiving an LSP, the local IS-IS neighbor calculates its checksum and compares it with the checksum contained in the LSP. By default, if the two checksums are inconsistent, the LSP will be discarded. If you run the **ignore-lsp- errors** command to configure to ignore checksum errors, the LSP will be processed normally despite checksum inconsistency.
- **lsp-fragments-extend**: Enables LSP fragment extension, which is used to generate an extended LSP when the 256 fragments of the original LSP are used up.

Notes

- The LSP refresh interval must be smaller than the maximum LSP lifetime.

- The maximum LSP lifetime must be greater than the LSP refresh interval.
- The value of **initial-interval** cannot be greater than that of **maximum-interval**. Otherwise, the value of **initial-interval** will be used as the value of **maximum-interval**.
- The value of **hold-interval** cannot be greater than that of **maximum-interval**. Otherwise, the value of **hold-interval** will be used as the value of **maximum-interval**.
- The value of **initial-interval** cannot be greater than that of **hold-interval**. Otherwise, the value of **initial-interval** will be used as the value of **hold-interval.**

Configuration Steps

Configuring the Minimum LSP Interval

- Perform this configuration based on requirements.
- Run the **isis lsp-interval** command in interface configuration mode on the desired device, unless otherwise specified.

Configuring the LSP Retransmission Interval

- Perform this configuration based on requirements.
- Run the **isis retransmit-interval** command in interface configuration mode on the desired device, unless otherwise specified.

Configuring LSP Packet Generation Cycle

- Perform this configuration based on requirements.
- Run the **lsp-gen-interval** command in interface configuration mode on the desired device, unless otherwise specified.

Configuring the LSP Refresh Interval

- Perform this configuration based on requirements.
- Run the **lsp-refresh-interval** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring the LSP Lifetime

- Perform this configuration based on requirements.
- Run the **max-lsp-lifetime** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring to Ignore LSP Checksum Errors

- Perform this configuration based on requirements.
- Run the **ignore-lsp-errors** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring LSP Fragment Extension

- Perform this configuration based on requirements.
- Run the **lsp-fragment-extend** and **virtual-system** commands in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- Update LSPs continuously and capture LSPs to check the minimum LSP interval.
- Disable neighboring routes and capture LSPs to check the LSP retransmission interval.
- Capture LSPs to check the refresh interval.
- Check the LSP lifetime.
- Send an LSP with an incorrect checksum and check whether the LSP is discarded.
- Reduce the **lsp-length originate** command value, add routing information, and capture LSPs to check whether more than 256 LSP fragments are generated.

Related Commands

Configuring the Minimum LSP Interval

Command	**isis lsp-interval** *interval* [**level-1** | **level-2**]
Parameter Description	*milliseconds*: Indicates the LSP interval. The value range is 1 to 4,294,967,295, in the unit of milliseconds. **level-1**: Applies the setting only to Level-1 LSPs. **level-2:** Applies the setting only to Level-2 LSPs.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the LSP Retransmission Interval

Command	**isis retransmit-interval** *interval* [**level-1** | **level-2**]
Parameter Description	***seconds***: Indicates the LSP retransmission interval. The value range is 0 to 65,535, in the unit of seconds. **level-1**: Applies the setting only to Level-1 LSPs. **level-2**: Applies the setting only to Level-2 LSPs.
Command Mode	Interface configuration mode
Usage Guide	Use this command to configure the LSP retransmission interval. In a P2P network, after a device sends an LSP, if the device receives no PSNP response within the time specified by this command, it will resend the LSP. If the retransmission interval is set to 0, the LSP will not be resent,

Configuring LSP Packet Generation Cycle

Command	**lsp-gen-interval** [**level-1 | level-2**] *maximum-interval* [*initial-interval hold-interval*]
Parameter Description	**level-1**: Applies the configuration only to Level-1. **level-2**: Applies the configuration only to Level-2. *maximum-interval*: Indicates the maximum interval for generating two consecutive LSP packets. The value range is **1** to **65535** (in seconds). The default value is **5**. *initial-interval*: Indicates the waiting time for generating an LSP packet for the first time. The value range is **0** to **60000** (in milliseconds). The default value is **50**. *hold-interval*: Indicates the minimum interval for generating an LSP packet for the second time. The value range is **10** to

	60000 (in milliseconds). The default value is **200**.
Configuration Mode	IS-IS routing process configuration mode
Usage Guide	The LSP packet generation interval refers to the interval for generating two different LSP packets. A smaller generation interval indicates faster network convergence, which, however, will be accompanied by frequent flooding on the network. The waiting time for generating an LSP packet for the first time is the initial interval. If the network becomes unstable, the LSP packet regeneration interval is changed to be less than the maximum interval, and the interval for generating an LSP packet for the second time becomes the hold interval. A corresponding penalty will be added to this interval: The next interval for regenerating a LSP packet doubles the previous interval for generating the same LSP packet, until the regeneration interval reaches the maximum interval. Subsequent LSP packets will be generated at the maximum interval. When the network becomes stable, the LSP packet regeneration interval becomes greater than the maximum interval, and the waiting time for LSP packet generation is restored to the initial interval. Link changes have high requirements for convergence. The initial interval can be set to a small value. The preceding parameters can also be adjusted to larger values to reduce CPU consumption. The value of **initial-interval** cannot be greater than that of **maximum-interval**. Otherwise, the value of **initial-interval** will be used as the value of **maximum-interval**. The value of **hold-interval** cannot be greater than that of **maximum-interval**. Otherwise, the value of **hold-interval** will be used as the value of **maximum-interval**. The value of **initial-interval** cannot be greater than that of **hold-interval**. Otherwise, the value of **initial-interval** will be used as the value of **hold-interval**.

Configuring the LSP Refresh Interval

Command	**lsp-refresh-interval** *interval*
Parameter Description	*interval*: Indicates the LSP refresh interval. The value range is 1 to 65,535, in the unit of seconds. The default value is 900.
Command Mode	IS-IS routing process configuration mode
Usage Guide	After an LSP has remained stable for a period specified by this command, it will be refreshed and updated before being published. The LSP refresh interval must be smaller than the maximum LSP lifetime.

Configuring the LSP Lifetime

Command	**max-lsp-lifetime** *value*
Parameter Description	*value*: Indicates the maximum time that LSPs keep alive. The value range is 1 to 65,535, in the unit of seconds. The default value is 1,200.
Command Mode	IS-IS routing process configuration mode
Usage Guide	The maximum LSP lifetime must be greater than LSP refresh interval.

Configuring to Ignore LSP Checksum Errors

Command	**ignore-lsp-errors**

Parameter Description	N/A
Command Mode	IS-IS routing process configuration mode
Usage Guide	After receiving an LSP, the local IS-IS neighbor calculates its checksum and compares it with the checksum contained in the LSP. By default, if the two checksums are inconsistent, the LSP will be discarded. If you run the **ignore-lsp- errors** command to configure to ignore checksum errors, the LSP will be processed normally despite checksum inconsistency.

Configuring LSP Fragment Extension

Command	**lsp-fragments-extend** [**level-1 | level-2**] [**compatible rfc3786**]
Parameter Description	**level-1**: Applies the setting only to Level-1 LSPs. **level-2**: Applies the setting only to Level-2 LSPs. **compatible**: Indicates compatibility with the RFC version of extended LSPs. **rfc3786**: Extends the LSP old version.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to enable LSP fragment extension.

Configuring an Additional System ID

Command	**virtual-system** *system-id*
Parameter Description	*system-id*: Indicates an additional system ID (6-byte).
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to configure the additional system ID of an IS-IS routing process, which is used by the extended LSP that is generated after the 256 fragments of the original LSP are used up. To enable fragment extension, run the **lsp-fragments-extend** command.

Configuration Example

Configuring the Minimum LSP Interval

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the minimum LSP interval.
	A(config)# interface GigabitEthernet 0/1 A(config-if)# isis lsp-interval 100 level-2
Verification	Run the **clear isis *** command to update LSPs continuously and capture LSPs to check the minimum LSP interval.

Configuring the LSP Retransmission Interval

Configuration Steps	● Configure IS-IS neighbors in P2P mode. (Omitted) ● Configure the LSP retransmission interval.
	A(config)# interface serial 0/1 A(config-if)# isis retransmit-interval 10 level-2
Verification	Disable neighboring routes and capture LSPs to check the LSP retransmission interval.

Configuring LSP Packet Generation Cycle

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the LSP packet generation cycle.
	A(config)# router isis A(config-router)# lsp-gen-interval 5 50 100
Verification	Generate LSP packet frequently to see whether the exponential backoff algorithm is used.

Configuring the LSP Refresh Interval

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the LSP refresh interval.
	A(config)# router isis A(config-router)# lsp-refresh-interval 600
Verification	Capture LSPs to check the refresh interval.

Configuring the LSP Lifetime

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the LSP lifetime.
	A(config)# router isis A(config-router)# max-lsp-lifetime 1500
Verification	Check the LSP lifetime (**LSP Holdtime** field).
	A# show isis database

Configuring to Ignore LSP Checksum Errors

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure to ignore LSP checksum errors.
	A(config)# router isis A(config-router)# ignore-lsp-errors
Verification	Send an LSP with an incorrect checksum and check whether the LSP is discarded.

Configuring LSP Fragment Extension

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure LSP fragment extension. ● Configure the additional system ID of the IS-IS routing process.
	A(config)# router isis A(config-router)# lsp-fragments-extend A(config-router)# virtual-system 0000.0000.0034
Verification	Reduce the **lsp-length originate** command value, add routing information, and capture LSPs to check whether more than 256 LSP fragments are generated.

4.4.4 Configuring IS-IS SNPs

Configuration Effect

● CSNPs are periodically broadcast by the DIS in a broadcast network for LSDB synchronization. In a P2P network, a CSNP is sent only after a neighbor relationship is established. An interface set to **mesh-groups** can be configured to periodically send CSNPs.

● When you need to set **mesh-group** on an IS-IS interface, run the **isis csnp-interval** command to configure the non-0 CSNP interval to ensure complete LSP synchronization between neighbors in the network. After that, CNSPs will be periodically sent to synchronize LSPs.

Configuration Steps

● Perform this configuration based on requirements.

● Run the **isis csnp-interval** *interval* [**level-1** | **level-2**] command in interface configuration mode on the desired device, unless otherwise specified.

Verification

Capture CSNPs in the broadcast network to check the CSNP interval.

Related Commands

Configuring Source Registration Filter

Command	**isis csnp-interval** *interval* [**level-1** | **level-2**]
Parameter Description	*interval*: Indicates the CSNP interval. The value range is 0 to 65,535, in the unit of seconds. **level-1**: Applies the setting only to Level-1 CSNPs. **level-2**: Applies the setting only to Level-2 CSNPs.
Command Mode	Interface configuration mode
Usage Guide	Use this command to change the CSNP interval. By default, a DIS sends a CSNP every 10s in a broadcast network. In a P2P network, a CSNP is sent only after a neighbor relationship is established. An interface set to **mesh-groups** can be configured to periodically send CSNPs. No CSNPs are sent if the CSNP interval is set to 0.

Configuration Example

Configuring the CSNP Broadcast Interval

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the CSNP broadcast interval.
	A(config)# interface GigabitEthernet 0/1 A(config-if)# isis csnp-interval 20
Verification	Capture packets to check the CSNP interval.

4.4.5 Configuring the IS-IS Level Type

Configuration Effect

- IS-IS supports a two-Level system to realize routing management and extensible route selection in a large network. Each Level is only concerned about maintaining the topology of the corresponding area.
- You can run the **is-type** command in IS-IS routing process configuration mode to configure an IS-IS Level, or run the **isis circuit-type** command in interface configuration mode to configure the IS-IS Level of an interface. The default Levels specified by the **is-type** and **isis circuit-type** commands are Level-1/Level-2. If you run both commands, the interface only sends the PDUs of the same Level specified by the two commands.

Notes

- If Level-1 or Level-2-only is configured using the **circuit-type** command, IS-IS will only send PDUs of the corresponding Level.
- If an interface is set to **external**, the interface will work as an external domain interface and IS-IS will not send PDUs of the corresponding Level.
- A device can have only one instance running at Level-2 (including Level-1/Level-2).

Configuration Steps

Configuring the System Type

- Perform this configuration based on requirements.
- Run the **is-type** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring the Interface Circuit Type

- Perform this configuration based on requirements.
- Run the **isis circuit-type** command in interface configuration mode on the desired device, unless otherwise specified.

Verification

- Check whether only the instances of the Level specified by the **is-type** command are processed, and neighbors of the corresponding Level are created.
- Check whether the interface only sends the PDUs of the same Level specified by the **is-type** and **circuit-type** commands.

Related Commands

Configuring the System Type

Command	**is-type** { **level-1 | level-1-2 | level-2-only** }
Parameter Description	**level-1**: Indicates that IS-IS only runs at Level-1. **level-1-2**: Indicates that IS-IS runs at Level-1 and Level-2. **level-2-only**: Indicates that IS-IS only runs at Level-2.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Changing the **is-type** value will enable or disable the routes of the corresponding level.

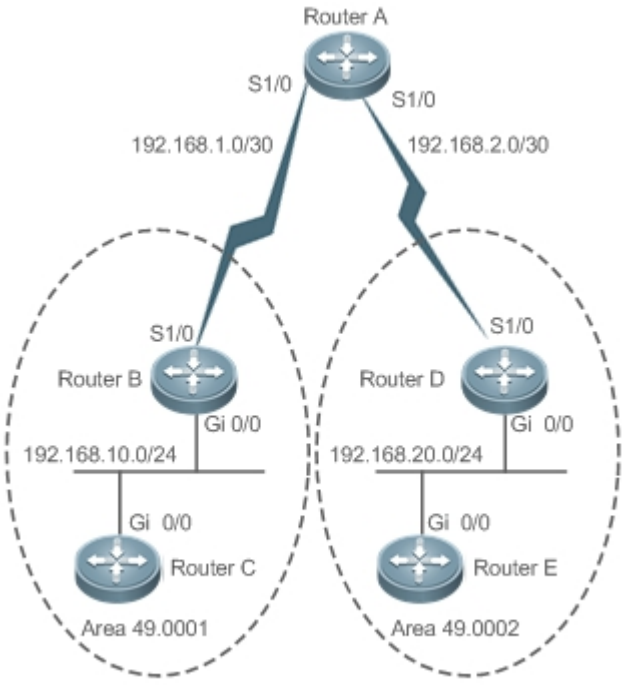
Configuring the Interface Circuit Type

Command	**isis circuit-type** { **level-1 | level-1-2 | level-2-only** [**external**] }
Parameter Description	**level-1**: Establishes a Level-1 neighbor relationship. **level-2-only**: Establishes a Level-2 neighbor relationship. **level-1-2**: Establishes a Level-1/Level-2 neighbor relationship. **external**: Uses the interface as an external domain interface.
Command Mode	Interface configuration mode
Usage Guide	If the circuit type is set to Level-1 or Level-2-only, IS-IS will only send PDUs of the corresponding Level. If the system type is set to Level-1 or Level-2-only, IS-IS only processes the instances of the corresponding Level, and the interface only sends the PDUs of the same Level specified by the **is-type** and **circuit-type** commands. If the interface is set to **external**, the interface will work as an external domain interface and IS-IS will not send PDUs of the corresponding Level.

Configuration Example

Configuring IS-IS Levels

Configuration Requirements	Router A is connected to Router B and Router C by P2P serial links. Router B and Router C are connected by the Ethernet, and Router D and Router E are also connected by the Ethernet. On Router A, configure IS-IS area route summarization.

	Note that area route summarization can be configured only on border devices.
Figure 4- 12 IS-IS Level Configuration	
Configuration Steps	● Configure IS-IS. ● Configure Ethernet interfaces. ● Configure the IS-IS Level structure.
A	Configure IS-IS. A(config)# router isis A(config-router)# net 50.0001.0000.0000.0001.00 A(config-router)# is-type level-2-only
	Configure two serial link ports. A(config)# interface Serial 1/0 A(config-if)# ip address 192.168.1.1 255.255.255.252 A(config-if)# ip router isis A(config)# interface Serial 1/1 A(config-if)# ip address 192.168.2.1 255.255.255.252 A(config-if)# ip router isis
B	Configure IS-IS.
	B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00

	Configure an Ethernet interface.
	B(config)# interface GigabitEthernet 0/0 B(config-if)# ip address 192.168.10.1 255.255.255.0 B(config-if)# ip router isis
	Configure a serial link port.
	B(config)# interface Serial 1/0 B(config-if)# ip address 192.168.1.2 255.255.255.252 B(config-if)# ip router isis
C	Configure IS-IS.
	C(config)# router isis C(config-router)# net 49.0001.0000.0000.0003.00 C(config-router)# is-type level-1
	Configure an Ethernet interface.
	C(config)# interface GigabitEthernet 0/0 C(config-if)# ip address 192.168.10.2 255.255.255.0 C(config-if)# ip router isis
D	Configure IS-IS.
	D(config)# router isis D(config-router)# net 49.0002.0000.0000.0004.00
	Configure an Ethernet interface.
	D(config)# interface GigabitEthernet 0/0 D(config-if)# ip address 192.168.20.1 255.255.255.0 D(config-if)# ip router isis
	Configure a serial link port.
	D(config)# interface Serial 1/0 D(config-if)# ip address 192.168.2.2 255.255.255.252 D(config-if)# ip router isis
E	Configure IS-IS.
	E(config)# router isis E(config-router)# net 49.0002.0000.0000.0005.00 E(config-router)# is-type level-1
	Configure an Ethernet interface.

	E(config)# interface GigabitEthernet 0/0 E(config-if)# ip address 192.168.20.2 255.255.255.0 E(config-if)# ip router isis
Verification	● Check whether neighbor relationships are established normally. ● Capture packets to check whether Router A only sends and receives Level-2 packets. ● Capture packets to check whether Router B and Route D only send and receive Level-1 and Level-2 packets. ● Capture packets to check whether Router C and Router E only send and receive Level-1 packets.
A	A# show isis neighbors A# show isis database detail
B	B# show isis neighbors B# show isis database detail
C	C# show isis neighbors C# show isis database detail
D	D# show isis neighbors D# show isis database detail
E	E# show isis neighbors E# show isis database detail

4.4.6 Configuring IS-IS Authentication

Configuration Effect

- Interface authentication is intended for establishing and maintaining neighbor relationships. A neighbor relationship cannot be established between two IS-IS devices with different interface authentication passwords. This prevents unauthorized or unauthenticated IS-IS devices from joining an IS-IS network that requires authentication. Interface authentication passwords are encapsulated in Hello packets before being sent.

- Area authentication and RD authentication in IS-IS are performed to verify LSPs, CSNPs, and PSNPs to prevent unauthorized or unauthenticated routing information from being injected into the LSDB. Authentication passwords are encapsulated in LSPs, CSNPs, and PSNPs before being sent.

Notes

- An interface authentication password is encapsulated in a Hello packet before being sent by an interface. When an interface receives a Hello packet, it checks the password in the packet against the existing one.

- Area authentication passwords are encapsulated in Level-1 LSPs, CSNPs, and PSNPs. When an interface receives an LSP, CSNP, or PSNP, it checks the password in the packet against the existing one.

- RD authentication passwords are encapsulated in Level-2 LSPs, CSNPs, and PSNPs. When an interface receives an LSP, CSNP, or PSNP, it checks the password in the packet against the existing one.

Configuration Steps

Configuring Interface Authentication

- Perform this configuration based on requirements.
- Configure **isis password** in interface configuration mode on the desired device, unless otherwise specified.

Configuring Area Authentication

- Perform this configuration based on requirements.
- Run the **area-password** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring RD Authentication

- Perform this configuration based on requirements.
- Run the **domain-password** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- IS-IS plaintext authentication provides only limited security because the password transferred through a packet is visible.
- IS-IS MD5 authentication provides higher security because the password transferred through a packet is encrypted using the MD5 algorithm.

Related Commands

Configuring the Password for Plaintext Authentication of Hello Packets on an Interface

Command	**isis password** [**0** | **7**] *password* [**send-only**] [**level-1** | **level-2**]
Parameter Description	**0**: Indicates that the key is displayed in plaintext. **7**: Indicates that the key is displayed in ciphertext. ***password-string***: Indicates the password string for plaintext authentication. The string can contain up to 126 characters. **send-only**: Indicates that the plaintext authentication password is only used to authenticate sent packets. Received packets are not authenticated. **level-1**: Applies the setting to the Level-1 circuit type. **level-2**: Applies the setting to the Level-2 circuit type.
Command Mode	Interface configuration mode
Usage Guide	Use this command to configure the password for Hello packet authentication on an interface. Use the **no** form of this command to clear the password. If no Level is specified, by default, the password takes effect for Level-1 and Level-2 circuit types. This command does not take effect if the **isis authentication mode** command is executed. You need to first delete the previous command configuration. If you include the **send-only** parameter when deleting the **isis authentication mode** command configuration, only the

	parameter setting is canceled.

Specifying Interface Authentication as Plaintext or MD5

Command	**isis authentication mode** { **md5** | **text** } [**level-1** | **level-2**]
Parameter Description	**md5**: Uses MD5 authentication. **text**: Uses plaintext authentication. **level-1**: Applies the setting to the Level-1 circuit type. **level-2**: Applies the setting to the Level-2 circuit type.
Command Mode	Interface configuration mode
Usage Guide	Use this command to specify the authentication mode before you can make the key chain configured using the **isis authentication key-chain** command take effect. If no Level is specified, the authentication mode will take effect for Level-1 and Level-2 circuit types. If you use the **isis authentication mode** command after the **isis password** command is executed to configure plaintext authentication, the previous command configuration will be overwritten. The **isis password** command does not take effect if the **isis authentication mode** command is executed. To run the **isis password** command, delete the **isis authentication mode** command configuration first.

Configuring the Password for Interface Authentication

Command	**isis authentication key-chain** *name-of-chain* [**level-1** | **level-2**]
Parameter Description	*name-of-chain*: Indicates the name of a key chain. The maximum length is 255. **level-1**: Indicates that the authentication key chain takes effect for Level-1. **level-2**: Indicates that the authentication key chain takes effect for Level-2.
Command Mode	Interface configuration mode
Usage Guide	Authentication is not performed if no key chain is configured using the **key chain** command. In addition to the **key chain** command, you also need to run the **isis authentication mode** command to make IS-IS key chain authentication take effect. The key chain is applicable to plaintext authentication and MD5 authentication. Which authentication mode to use can be determined using the **isis authentication mode** command. For plaintext authentication, the key-string in the key chain cannot exceed 80 characters; otherwise, the key chain will be invalid. Only one key chain can be used at a time. After you configure a new key chain, it will replace the original one. If no Level is specified, the key chain takes effect for Level-1 and Level-2. The key chain is applicable to Hello packets. IS-IS will send or receive passwords that belong to the key chain. A key chain may contain multiple passwords. A password with a smaller SN is preferentially used for sending a packet. When the packet arrives at the peer device, the device will receive the packet if the packet-carried password is consistent with a password in the key chain. The authentication commands (for example, **authentication key-chain**) executed in IS-IS routing process configuration mode are intended for LSPs and SNPs. They do not take effect for IS-IS interfaces.

(Optional) Applying Interface Authentication Only to Sent Packets (Received Packets Are Not Authenticated)

Command	**isis authentication send-only** [**level-1** | **level-2**]
Parameter Description	**level-1**: Sets **send-only** for Level-1 on an interface. **level-2**: Sets **send-only** for Level-2 on an interface.
Command Mode	Interface configuration mode
Usage Guide	Use this command to enable IS-IS to set an authentication password in the Hello packet sent by an interface. However, IS-IS does not authenticate the Hello packet received by the interface. You can use this command before you deploy IS-IS interface authentication on all devices in the network or before you change the authentication password or authentication mode. After you run the **isis authentication send-only** command, the devices will not authenticate received Hello packets to avoid network flapping when IS-IS interface authentication is deployed. After authentication is deployed in the entire network, run the **no isis authentication send-only** command to cancel the **send-only** setting. The **isis authentication send-only** command is applicable to plaintext authentication and MD5 authentication. You can run the **isis authentication mode** command to specify the authentication mode for an IS-IS interface. If no Level is specified, the authentication mode will take effect for Level-1 and Level-2 on the interface.

Configuring the Password for Area (Level-1) Plaintext Authentication

Command	**area-password** [**0** | **7**] *password* [**send-only**]
Parameter Description	**0**: Indicates that the key is displayed in plaintext. **7**: Indicates that the key is displayed in ciphertext. ***password-string***: Indicates the password string for plaintext authentication. The string can contain up to 126 characters. **send-only**: Indicates that the plaintext authentication password is only used to authenticate sent Hello packets in Level-1 areas. Received Hello packets are not authenticate.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Run this command to enable authentication of received LSPs, CSNPs, and PSNPs in Level-1 areas and include authentication information in these packets before they are sent. All IS-IS devices in an area must be configured with the same password. This command does not take effect if the **authentication mode** command is executed. You need to first delete the previous command configuration. To delete the password, run the **no area-password** command. If you run the **no area-password send-only** command, only the **send-only** setting is canceled. If you run the **area-password** *psw* **send-only** and **no area-password send-only** commands in sequence, the configuration is changed to **area-password** *psw*.

Configuring the Password for RD (Level-2) Plaintext Authentication

Command	**domain-password** [**0** | **7**] *password* [**send-only**]
Parameter Description	**0**: Indicates that the key is displayed in plaintext. **7**: Indicates that the key is displayed in ciphertext. ***password-string***: Indicates the password string for plaintext authentication. The string can contain up to 126 characters. **send-only**: Indicates that the plaintext authentication password is only used to authenticate sent Hello packets in Level-1 areas. Received Hello packets are not authenticated.
Command Mode	IS-IS routing process configuration mode

Usage Guide	Run this command to enable authentication of received LSPs, CSNPs, and PSNPs in Level-2 domains and include authentication information in these packets before they are sent. All IS-IS devices in a Level-2 domain must be configured with the same password. This command does not take effect if the **authentication mode** command is executed. You need to first delete the previous command configuration. To delete the password, run the **no domain-password** command. If you run the **no domain-password send-only** command, only the **send-only** setting is canceled. If you run the **domain-password** *psw* **send-only** and **no domain-password send-only** commands in sequence, the configuration is changed to **domain-password** *psw*.

Specifying the IS-IS RD Authentication Mode

Command	**authentication mode** { **md5** | **text** } [**level-1** | **level-2**]
Parameter Description	**md5**: Uses MD5 authentication. **text**: Uses plaintext authentication. **level-1**: Indicates that the authentication mode takes effect for Level-1. **level-2**: Indicates that the authentication mode takes effect for Level-2.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to specify the authentication mode before you can make the key chain configured using the **authentication key-chain** command take effect. If no Level is specified, the authentication mode will take effect for Level-1 and Level-2. If you use the **authentication mode** command after the **area-password** or **domain-password** command is executed to configure plaintext authentication, the previous command configuration will be overwritten. The **area-password** or **domain-password** command does not take effect if the **authentication mode** command is executed. To run the **area-password** or **domain-password** command, delete the **authentication mode** command configuration first.

Specifying the Key Chain for IS-IS Authentication

Command	**authentication key-chain** *name-of-chain* [**level-1 | level-2**]
Parameter Description	*name-of-chain*: Indicates the name of a key chain. The maximum length is 255. **level-1**: Indicates that the authentication key chain takes effect for Level-1. **level-2**: Indicates that the authentication key chain takes effect for Level-2.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Authentication is not performed if no key chain is configured using the **key chain** command. In addition to the **key chain** command, you also need to run the **authentication mode** command to make IS-IS key chain authentication take effect. The key chain is applicable to plaintext authentication and MD5 authentication. Which authentication mode to use can be determined using the **authentication mode** command. For plaintext authentication, the key-string in the key chain cannot exceed 80 characters; otherwise, the key chain will be invalid. Only one key chain can be used at a time. After you configure a new key chain, it will replace the original one. If no Level is specified, the key chain takes effect for Level-1 and Level-2. The key chain is applicable to LSPs, CSNPs, and PSNPs. IS-IS will send or receive passwords that belong to the key chain. A key chain may contain multiple passwords. A password with a SN is preferentially used for sending a packet. When the

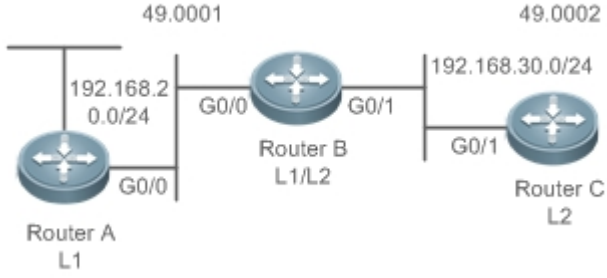
	packet arrives at the peer device, the device will receive the packet if the packet-carried password is consistent with a password in the key chain.

Applying IS-IS Authentication Only to Sent Packets

Command	**authentication send-only** [**level-1** | **level-2**]
Parameter Description	**level-1**: Applies the **send-only** setting to Level-1. **level-2**: Applies the **send-only** setting to Level-2.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to enable IS-IS to set an authentication password in the Hello packet to be sent. However, IS-IS does not authenticate received Hello packets. You can use this command before you deploy IS-IS authentication on all devices in the network or before you change the authentication password or authentication mode. After you run the **authentication send-only** command, the devices will not authenticate received packets to avoid network flapping when authentication passwords are deployed. After authentication is deployed in the entire network, run the **no isis authentication send-only** command to cancel the **send-only** setting. The **authentication send-only** command is applicable to plaintext authentication and MD5 authentication. You can run the **authentication mode** command to specify the authentication mode. If no Level is specified, the authentication mode will take effect for Level-1 and Level-2.

Configuration Example

Configuring IS-IS Authentication

Configuration Requirements	Router A, Router B, and Router C are connected through the Ethernet and run IS-IS. Router A is a Level-1 device, Route B is a Level-1/Level-2 device, and Router C is a Level-2 device. The following configuration requirements exist: Apply plaintext authentication to the Hello packets between Router A and Router B, as well as Level-1 LSPs and SNPs. Apply MD5 authentication to the Hello packets between Router B and Router C, as well as Level-2 LSPs and SNPs.
Figure 4- 13 IS-IS Authentication Topology	
Configuration Steps	● Configure IS-IS. ● Configure Ethernet interfaces. ● Configure the password for IS-IS authentication.
A	Configure IS-IS.
	A(config)# router isis A(config-router)# net 49.0001.0000.0000.0001.00

	A(config-router)# is-type level-1 A(config-router)# area-password aa
	Configure an Ethernet interface. A(config)# interface GigabitEthernet 0/0 A(config-if)# ip address 192.168.20.1 255.255.255.0 A(config-if)# ip router isis A(config-if)# isis password cc
B	Configure the password for IS-IS authentication.
	B(config)# key chain kc1 B(config-keychain)# key 1 B(config-keychain-key)# key-string aa B(config)# key chain kc2 B(config-keychain)# key 1 B(config-keychain-key)# key-string bb B(config)# key chain kc3 B(config-keychain)# key 1 B(config-keychain-key)# key-string cc
	Configure IS-IS.
	B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00 B(config-router)# authentication mode text level-1 B(config-router)# authentication key-chain kc1 B(config-router)# authentication mode md5 level-2 B(config-router)# authentication key-chain kc2
	Configure two Ethernet interfaces.
C	B(config)# interface GigabitEthernet 0/0 B(config-if)# ip address 192.168.20.2 255.255.255.0 B(config-if)# ip router isis B(config-if)# isis authentication mode text B(config-if)# isis authentication key-chain kc3 B(config)# interface GigabitEthernet 0/1 B(config-if)# ip address 192.168.30.2 255.255.255.0 B(config-if)# ip router isis

	B(config-if)# isis authentication mode md5 B(config-if)# isis authentication key-chain kc3 Configure the password for IS-IS authentication.
	C(config)# key chain kc2 C(config-keychain)# key 1 C(config-keychain-key)# key-string bb C(config)# key chain kc3 C(config-keychain)# key 1 C(config-keychain-key)# key-string cc
	Configure IS-IS.
	C(config)# router isis C(config-router)# net 49.0002.0000.0000.0002.00 C(config-router)# is-type level-2 C(config-router)# authentication mode md5 level-2 C(config-router)# authentication key-chain kc2 Configure an Ethernet interface.
	C(config)# interface GigabitEthernet 0/1 C(config-if)# ip address 192.168.30.3 255.255.255.0 C(config-if)# ip router isis C(config-if)# isis authentication mode md5 C(config-if)# isis authentication key-chain kc3
Verification	Check whether neighbor relationships are established normally.
A	A# show isis neighbors A# show isis database detail
B	B# show isis neighbors
C	C# show isis neighbors

Common Errors

- Different authentication passwords are configured between neighbors.
- Different authentication modes are configured between neighbors.

4.4.7 Configuring IS-IS GR

Configuration Effect

- IS-IS GR helps improve system reliability. On devices that separate the control plane from the forwarding plane, GR ensures that data forwarding is not interrupted during routing protocol restart.

IS-IS GR Working Mechanism

For GR to be successful, the following two conditions must be met: (1) The network topology is stable; (2) The device can ensure uninterrupted forwarding when it restarts IS-IS.

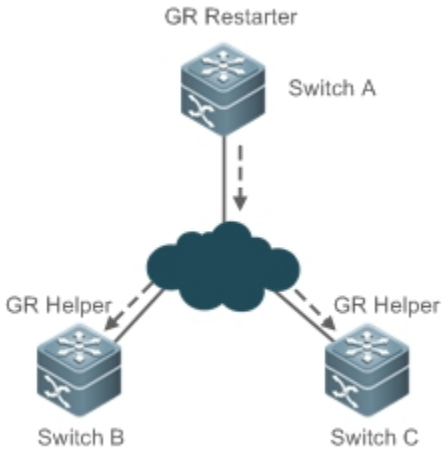
Two roles exist during the GR process: Restarter and Helper. Accordingly, IS-IS GR is divided into the IS-IS GR Restart capability and IS-IS GR Help capability. A device with the GR Restart capability can send a GR request and execute GR. A device with the GR Help capability can receive a GR request and help its neighbor with GR implementation. The GR process starts when the Restarter sends a GR request. After receiving the GR request, the neighboring device enters Help mode to help the Restarter reestablish its LSDB while maintaining the neighbor relationship with the Restarter. The main GR working mechanism is as follows:

When an IS-IS device needs to perform GR, it instructs its neighbor to maintain their neighbor relationship so that other devices in the network cannot sense the change in the topological relationship and the neighbor will not recalculate the route and update its forwarding table. The IS-IS device synchronizes and restores the LSDB to its pre-GR state with the help of the neighbor to ensure that the route and forwarding table remain unchanged before and after GR implementation and data forwarding is not interrupted.

The Restarter performs the following operations during the GR process:

1. The GR Restarter notifies the GR Helpers that it will be restarted.

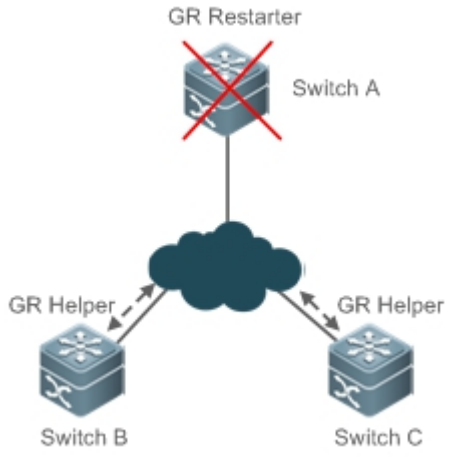
Figure 4- 14 Restart Notification by the GR Restarter



Switch A is a GR Restarter, and Switch B and Switch C are the GR Helpers for Switch A. Switch A sends a GR request instructing all its neighbors not to delete the neighbor relationships with Switch A when it is restarted. After receiving the GR request, the neighbors send GR responses to the GR Restarter, and will maintain their neighbor relationships with the GR Restarter during the GR time (specified by **GR grace-period**) notified by the GR Restarter.

2. The GR Restarter is restarted.

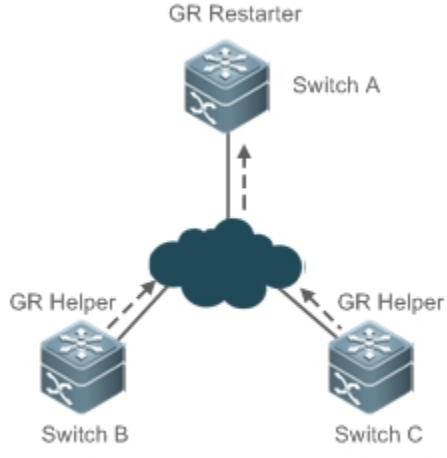
Figure 4- 15 Restart Performed by the GR Restarter



When the GR Restarter is restarted, its IS-IS interface goes from Down to Up. Because the GR Helpers know that the GR Restarter is in IS-IS restart state, they maintain their neighbor relationships with the GR Restarter during the GR time and retain the routes from the GR Restarter.

3. The GR Restarter synchronizes topology and routing information from the GR Helpers.

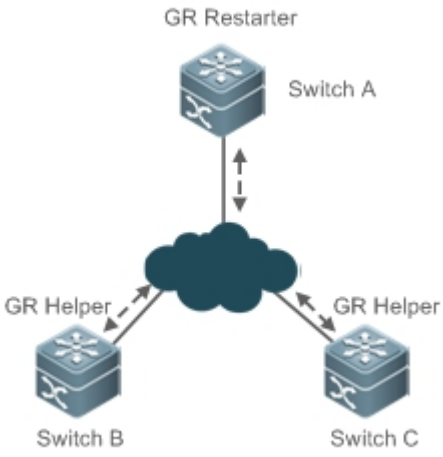
Figure 4- 16 LSDB Synchronization



After IS-IS restart, the GR Restarter synchronizes topology or routing information from the GR Helpers and recalculates its routing table. During this process, any change in the routing table is not updated to the forwarding table.

4. GR is completed when the GR Restarter finishes LSDB synchronization. Then all devices enter IS-IS interaction state.

Figure 4- 17 GR Completion



After the GR Restarter synchronizes all required data, all devices enter IS-IS interaction state. The GR Restarter's routing table is updated to the forwarding table and invalid entries are cleared. Because the GR Restarter is completely restored to the pre-restart state under stable network conditions, its routing table and forwarding table remain unchanged before and after GR.

Notes

- IS-IS GR is implemented based on RFC5306: Restart Signaling for IS-IS.
- All products support the IS-IS GR Helper capability.

Configuration Steps

Enabling the IS-IS GR Restart Capability

- Perform this configuration based on requirements.
- Run the **graceful-restart** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring the Maximum GR Time

- Perform this configuration based on requirements.
- Run the **graceful-restart grace-period** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Enabling the IS-IS GR Help Capability

- Perform this configuration based on requirements.
- Run the **graceful-restart helper** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- Check whether the routing table and forwarding table remain unchanged before and after GR.

Related Commands

Enabling the IS-IS GR Restart Capability

Command	**graceful-restart**

Parameter Description	N/A
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to enable the IS-IS GR Restart capability. As long as the network conditions remain unchanged, IS-IS can be restarted and restored to the pre-restart state without impact on data forwarding.

Configuring the Maximum GR Time

Command	**graceful-restart grace-period** *seconds*
Parameter Description	*seconds*: Indicates the GR time. The value range is 1s to 65,535s. The default value is 300s.
Command Mode	IS-IS routing process configuration mode
Usage Guide	N/A

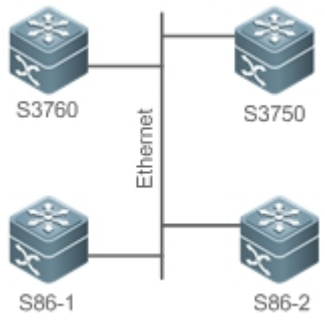
Enabling the IS-IS GR Help Capability

Command	**graceful-restart helper disable**
Parameter Description	N/A
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use the **graceful-restart helper disable** command to disable the IS-IS GR Help capability. The command enables IS-IS to ignore the GR request sent by the device to be restarted.

Configuration Example

Configuring IS-IS GR

Configuration Requirements	Two S8600 series high-end devices have the IS-IS GR Restart capability and are equipped with master/slave management boards for redundant backup at the control plane. IS-IS neighbor relationships are established between S86-1 and S3750/S3760 and between S86-2 and S3750/S3760. The system software of all devices supports the IS-IS GR Help capability. The following configuration requirements exist: Enable the IS-IS GR Restart capability with proper GR Time setting on S86-1 and S86-2 to realize uninterrupted forwarding and improve core device reliability. Disable the IS-IS GR Help capability on S3750 to exclude it from the Help process. By default, other device supports the IS-IS GR Help capability and require no additional configuration.

Figure 4- 18 IS-IS GR Topology	
Configuration Steps	Configure IS-IS. (Omitted) Configure Ethernet interfaces. (Omitted)
S86-1	Configure IS-IS GR.
	S86-1 (config)# router isis CS86-1(config-router)# graceful-restart CS86-1(config-router)# graceful-restart grace-period 60
S86-2	Configure IS-IS GR.
	CS86-2(config)# router isis CS86-2(config-router)# graceful-restart CS86-2(config-router)# graceful-restart grace-period 80
S3750	Disable the IS-IS Help capability.
	S3750(config)# router isis S3750(config-router)# graceful-restart helper disable
Verification	Check whether the routing table and forwarding table remain unchanged before and after GR. Check whether S86-1 and S86-2 synchronize topology and routing information from S3760.
S86-1	S86-1# show isis neighbors S86-1# show isis database detail
S86-2	S86-2# show isis neighbors
S3760	S3760# show isis neighbors

4.4.8 Configuring BFD Support for IS-IS

Configuration Effect

- IS-IS dynamically discovers neighbors through Hello packets. After IS-IS enables the BFD function, a BFD session will be set up with the neighbor in Up state. The BFD mechanism is used to detect the neighbor state. Once a neighbor failure is detected through BFD, IS-IS performs network convergence immediately. The convergence time can be reduced from 30s to less than 1s. By default, IS-IS Hello

packets are sent at an interval of 10s in a P2P network, and the time required to detect a neighbor failure is three times the packet interval, that is 30s.

Notes

- You must set BFD session parameters before you enable BFD support for IS-IS.
- When you run the **bfd up-dampening** command on an interface with BFD support for IS-IS, you need to run the **bfd all-interfaces** command with the [*anti-congestion*] option selected.
- When you run the **bfd all-interfaces** command with the [*anti-congestion*] option selected, run the **bfd up-dampening** command on the interface.
- IP routing may cause a neighbor's interface for BFD session setup to be inconsistent with the interface for outgoing BFD packets. If this happens, the BFD session cannot be set up.
- If a neighbor's interface for BFD session setup is inconsistent with the interface for outgoing BFD packets, the BFD session cannot be set up.

Configuration Steps

Enabling BFD Support for IS-IS on All Interfaces

- Perform this configuration based on requirements.
- Run the **bfd ll-interfaces** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Enabling BFD Support for IS-IS on the Current Interface

- Perform this configuration based on requirements.
- Run the **isis bfd** command in interface configuration mode on the desired device, unless otherwise specified.

Verification

- Build a topology with two parallel lines. Typically, IS-IS selects one line as the master line and the other as the backup line. Enable BFD on the master line.
- Make the master line fail. Check whether IS-IS performs route convergence based on the BFD monitoring state and starts the backup line.

Related Commands

Enabling BFD Support for IS-IS on the Current Interface

Command	**bfd all-interfaces** [*anti-congestion*]
Parameter Description	*anti-congestion*: Indicates the IS-IS BFD anti-congestion option.
Command Mode	IS-IS routing process configuration mode
Usage Guide	You can enable or disable BFD on an IS-IS interface by using any of the following two methods: Method 1: Run the **bfd all-interfaces** command in IS-IS routing process configuration mode to enable BFD on all IS-IS interfaces, and then run the **no bfd all-interfaces** command to disable BFD on all IS-IS interfaces. Method 2: Run the **isis bfd** [**disable**] command in interface configuration mode to enable BFD on the specified IS-IS

	interface, and then run the **isis bfd disable** command to disable BFD on the interface.

Enabling BFD Support for IS-IS on the Current Interface

Command	**isis bfd** [*disable* **|** *anti-congestion*]
Parameter Description	*disable*: Disables BFD support for IS-IS on the current interface. *anti-congestion*: Indicates the IS-IS BFD anti-congestion option.
Command Mode	Interface configuration mode
Usage Guide	You can enable or disable BFD on an IS-IS interface by using any of the following two methods: Method 1: Run the [**no**] **bfd all-interfaces** [**anti-congestion**] command in IS-IS routing process configuration mode to enable or disable BFD on all IS-IS interfaces. Method 2: Run the **isis bfd** [**disable** | **anti-congestion**] command in interface configuration mode to enable or disable BFD on the specified interface. Normally, BFD sends detection packets at millisecond intervals to detect the link state. When a link exception (such as a disconnected link) occurs, BFD can quickly detect it and instruct IS-IS to delete the neighbor relationship and the neighbor reachability information in LSPs. Then IS-IS recalculates and generates a new route to bypass the abnormal link, thus realizing fast convergence. With the introduction of new techniques such as the Multi-Service Transport Platform (MSTP), link congestion tends to occur during peak hours of data communication. BFD quickly detects the link exception and instructs IS-IS to delete the neighbor relationship and the neighbor reachability information in LSPs. Link switch is performed to bypass the congested link. A Hello packet for IS-IS neighbor detection is sent every 10s and its expiration time is 30s. The Hello packet can still be received normally when BFD detects an exception, and therefore an IS-IS neighbor relationship is reestablished quickly, causing the route to be restored to the congested link. Then BFD detects the abnormal link and link switch is performed again. This process is repeated, which makes the route be switched between the congested link and other links, causing repetitive flapping. The **anti-congestion** option is used to avoid routing flapping in case of link congestion. After the option is configured, the IS-IS neighbor state is still kept alive when link congestion occurs, but the neighbor reachability information in LSPs is deleted. The route is switched to a normal link. When the congested link is restored, the neighbor reachability information in LSPs is recovered and the route is switched back, which avoids route flapping. When you run the **bfd all-interfaces** [*anti-congestion*] command, run the **bfd up-dampening** command on the interface. The two commands must be used together. If you run only one command, the route flap dampening feature may not take effect or other network exceptions may occur.

Configuration Example

Enabling BFD Support for IS-IS on the Current Interface

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Set BFD session parameters. (Omitted) ● Enable BFD support for IS-IS on the current interface.
	A(config)# interface GigabitEthernet 0/1 A(config-if)# isis bfd

Verification	Enable S1 (192.168.1.10) and S2 (192.168.2.10) to send packets to G1 (229.1.1.1) and G2 (229.1.2.1). Add User to the G1 and G2 groups. ● Check the multicast packet that User receives. User should only receive the (S1, G1) packet. ● Check that the PIM-SM routing table does not have the (S1, G2), (S2, G1), and (S2, G2) entries.
	A# show bfd neighbors detail

Common Errors

- BFD support for IS-IS is not enabled on neighbors.

4.4.9 Setting the IS-IS Overload Bit

Configuration Effect

The overload bit is used in the following three situations:

- Device overload

The local IS-IS node has overload issues, such as insufficient memory or full CPU load; as a result, its routing table has incomplete routes or does not have resource forwarding data. You can set the overload bit in an LSP to instruct the neighbor not to use the local node as a forwarding device.

To set the overload bit, run the **set-overload-bit** command without the **on-startup** keyword. The overload bit can be configured or canceled manually. When the local IS-IS node is restored, manually cancel the command configuration; otherwise, the node is always in overload state.

- Instantaneous black hole

In the scenario described by RFC3277, the IS-IS convergence speed is faster than the BGP speed; as a result, after an IS-IS node is restarted, a route may be instantaneously unreachable, which is called an instantaneous black hole. You can set the overload bit in an LSP to instruct the neighbor not to use the local node as a forwarding device until the specified time has elapsed.

To set the overload bit, run the **set-overload-bit** command with the **on-startup** keyword. The overload bit can be configured or canceled automatically by the IS-IS node based on the configuration. If the **on-startup** keyword is selected, the IS-IS node automatically enters instantaneous black hole state after restart. When a neighbor relationship is established, the IS-IS node sends an LSP with the overload bit to notify the neighbor that the local node enters instantaneous black hole (or overload) state and instruct the neighbor not to use the local node as a forwarding device. After the specified time has elapsed, the IS-IS node immediately sends an LSP with the overload bit canceled to notify the neighbor that the local node has exited instantaneous black hole (or overload) state and can work as a forwarding device.

- Disabling real data forwarding on the local IS-IS node

If you only need to connect the local IS-IS node to a production network for testing or to meet other functional requirements, but does not require the node to forward real data in the network, you can set the overload bit in an LSP to instruct the neighbor not to use the local node as a forwarding device.

To set the overload bit, run the **set-overload-bit** command without the **on-startup** keyword. The overload bit can be configured or canceled manually. You can set the **suppress** keyword based on requirements to limit the routing information carried in an LSP in case of overload. For example, internal and external routes can be suppressed, and only the local direct route is advertised.

Notes

- At the same Level, the configuration with the **on-startup** keyword is mutually exclusive with the configuration without the **on-startup** keyword.

Configuration Steps

- Perform this configuration based on requirements.
- Run the **set-overload-bit** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- Capture packets and check that the neighbor does not forward LSPs from the local node.

Related Commands

Command	**set-overload-bit** [**on-startup** *seconds*] [**suppress** { [**interlevel**] [**external**] }] [**level-1** | **level-2**]
Parameter Description	***on-startup*** *seconds*: Indicates the duration when an IS-IS node remains in overload state after restart. The value range is 5s to 86,400s. ***suppress***: Indicates not to advertise internal routes (intra-area and inter-area routes) or external routes to neighbors when the IS-IS node is in overload state. ***interlevel***: Indicates not to advertise intra-area and inter-area routes to neighbors when the IS-IS node is in overload state. It is used with the **suppress** keyword. ***external***: Indicates not to advertise external routes to neighbors when the IS-IS node is in overload state. It is used with the **suppress** keyword. ***level-1***: Sends LSPs with the overload bit only to Level-1 neighbors. ***level-2***: Sends LSPs with the overload bit only to Level-2 neighbors.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to force an IS-IS node to set the overload bit in a non-virtual LSP to instruct its IS-IS neighbors not to use the local node as a forwarding device. If you select the **on-startup** keyword, the IS-IS node automatically enters overload state after restart. If you do not select the **on-startup** keyword, the IS-IS node enters overload state immediately after restart.

Configuration Example

Configuring the Overload Bit in Case of an Instantaneous Black Hole

Configuration Steps	- Configure IS-IS neighbors. (Omitted) - Verify that the IS-IS node enters instantaneous black hole state immediately after restart and remains in this state until the specified time (300s) has elapsed, and the IS-IS node only advertises local direct links to its neighbors during the specified time.
	A(config)# router isis A(config-router)#set-overload-bit on-startup 300 suppress interlevel external

Verification	Capture packets to check LSPs. ● Verify that the IS-IS node automatically enters instantaneous black hole state after restart. Once a neighbor relationship is established, the IS-IS node sends an LSP with the overload bit. ● After the specified time has elapsed, the IS-IS node immediately sends an LSP with the overload bit canceled to notify its neighbors that the local node has exited instantaneous black hole (or overload) state.
	A# show isis neighbors

Disabling Real Data Forwarding on the Local IS-IS Node

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Connect the local IS-IS node as a test device to a production network. The node is not required to forward real data in the network to avoid impact on production.
	A(config)# router isis A(config-router)#set-overload-bit suppress interlevel external
Verification	Capture packets to check LSPs. Verify that the LSPs carry the overload bit and only advertise local direct routes.
	A# show isis neighbors

4.4.10 Configuring IS-IS VRF

Configuration Effect

- Each VRF table can be seen as a virtual device or a dedicated PE device.
- The virtual device contains the following elements: an independent routing table, as well as an independent address space; a set of interfaces that belong to the VRF table; a set of routing protocols applicable only to the VRF table.
- Each device can maintain one or more VRF tables and a public-network routing table (also called a global routing table). Multiple VRF instances are separated from each other.

Notes

- Note the following constraints or conventions when you bind IS-IS instances and VRF tables:
- The IS-IS instances bound with the same VRF table must be configured with different system IDs. The IS-IS instances bound with different VRF tables can be configured with the same system ID.
- One IS-IS instance can be bound with only one VRF table, but one VRF table can be bound to multiple IS-IS instances.
- When the VRF table bound to an IS-IS instance is changed, all IS-IS interfaces associated with the instance will be deleted. That is, the **ip router isis** *[tag]* interface configuration and the redistribution configuration in routing process configuration mode will be deleted.

Configuration Steps

- Perform this configuration based on requirements.

- Run the **vrf** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- Check whether the local device establishes neighbor relationships with other devices specified in the VRF table.

Related Commands

Configuring IS-IS VRF

Command	**vrf** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of an existing VRF table.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Before you bind an IS-IS instance to a VRF table, ensure that the VRF table has been configured. If you need to establish an IS-ISv6 neighbor relationship, enable IPv6 and ensure that the table to be bound is a multiprotocol VRF table. Note the following constraints or conventions when you bind IS-IS instances and VRF tables: ● The IS-IS instances bound with the same non-default VRF table must be configured with different system IDs. The IS-IS instances bound with different VRF tables can be configured with the same system ID. ● One IS-IS instance can be bound with only one VRF table, but one VRF table can be bound to multiple IS-IS instances. ● When the VRF table bound to an IS-IS instance is changed, all IS-IS interfaces associated with the instance will be deleted. That is, the **ip** (or **ipv6**) **router isis** [*tag*] interface configuration and the redistribution configuration in routing process configuration mode will be deleted.

Configuration Example

Configuring IS-IS VRF

Configuration Steps	● Bind an IS-IS instance to a VRF table. ● Add interfaces to the VRF table and IS-IS instance. (Omitted)
	A(config)#vrf definition vrf_1 A(config-vrf)#address-family ipv4 A(config-vrf-af)#exit-address-family A(config)# router isis A(config-router)# vrf vrf_1
Verification	Check whether the local device establishes neighbor relationships with other devices specified in the VRF table.
	A# show isis neighbors

Common Errors

- Interfaces are not added to the VRF table.

- The IP addresses of the interfaces connected between neighbors are not in the same network segment.
- The **ip router isis** command is not executed on interfaces.
- No NET address is configured, or different NET addresses exist at Level-1.
- **max-area-addresses** is configured differently on both sides.
- **metric-style** is configured differently on both sides.
- The interface Levels on both sides are different. One side is Level-1, whereas the other side is Level-2.
- One side is configured with the P2P mode, whereas the other side is configured with the broadcast mode.
- One side is enabled with authentication, whereas the other side is not.

4.4.11 Configuring IS-IS MTR

Configuration Effect

- If the **multi-topology** command is not executed, IPv4 and IPv6 share one IS-IS physical topology, also called the default topology. If the **multi-topology** command is executed without the **transition** parameter, routing devices run in MT mode. IS-ISv4 runs in the default topology, and IS-ISv6 runs in the IPv6 unicast topology. If the **multi-topology** command is executed with the **transition** parameter, routing devices run in MTT mode. IS-ISv6 runs in the default topology and IPv6 unicast topology. The three configurations are mutually exclusive. The routing devices in MTT mode can transfer the MT TLV or the default topology TLV. The MTT mode is applicable to incremental deployment to ensure smooth network migration. The MTT mode can cause route leaking between the default topology and IPv6 unicast topology. If the MTT mode is configured improperly, network failures such as routing black holes and loops may occur.

Notes

Note the following constraints or conventions when you configure the IS-IS MTR feature:

- Set **metric-style** to **Wide** or **Transition** before you run the **multi-topology** command.
- The MTR feature will be disabled if **metric-style** is set to **Narrow** or only one Level is configured to support the Wide or Transition mode.

Configuration Steps

- Perform this configuration based on requirements.
- Configure the MTR feature in IS-IS address-family ipv6 configuration mode on the desired device, unless otherwise specified.

Verification

- Check whether the local device establishes neighbor relationships with other devices.

Related Commands

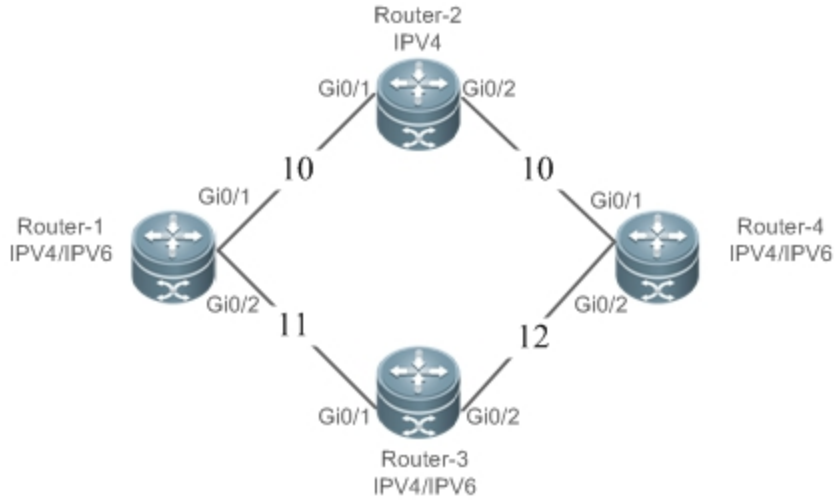
Configuring IS-IS MTR

Command	**multi-topology** [*transition*]
Parameter Description	*transition*: Configures the MTT mode, which supports smooth migration from an IPv4-IPv6 hybrid topology to separate IPv4 and IPv6 topologies.
Command Mode	IS-IS address-family ipv6 configuration mode
Usage Guide	If the **multi-topology** command is not executed, IPv4 and IPv6 share one IS-IS physical topology, also called the default topology. If the **multi-topology** command is executed without the **transition** parameter, routing devices run in MT

	mode. IS-ISv4 runs in the default topology, and IS-ISv6 runs in the IPv6 unicast topology. If the **multi-topology** command is executed with the **transition** parameter, routing devices run in MTT mode. IS-ISv6 runs in the default topology and IPv6 unicast topology. The three configurations are mutually exclusive. The routing devices in MTT mode can transfer the MT TLV or the default topology TLV. The MTT mode is applicable to incremental deployment to ensure smooth network migration. The MTT mode can cause route leaking between the default topology and IPv6 unicast topology. If the MTT mode is configured improperly, network failures such as routing black holes and loops may occur. Set **metric-style** to **Wide** or **Transition** before you run the command. The MTR feature will be disabled if **metric-style** is set to **Narrow** or only one Level is configured to support the Wide or Transition mode.

Configuration Example

Configuring IS-IS MTR

Configuration Requirements	The typical application scenario of MTR is to retain devices that only support IPv4 services in a network where IPv6 service extension will be performed. In Figure 4- 19, Router 2 only supports the IPv4 protocol stack but does not support the MTR feature; therefore, it can only run IPv4 services. The network capacity needs to be scaled to support IPv6 services in order to meet service extension requirements. (Router 1, Router 3, and Router 4 that support the MTR feature will be added.) The device (Router 2) that supports only one protocol stack must be replaced to maintain the stability of the network running IPv4 and IPv6 dual protocol stacks; otherwise, IPv6 routing black holes may occur. If you need to retain Router 2, you can configure the MTR feature on Router 1, Router 3, and Router 4. The MTR feature enables Router 2 to continue to run IPv4 services without interference on the IPv4 and IPv6 services on Router 1, Router 3, and Router 4. The MTR feature improves networking flexibility, indirectly prolongs the service life of old devices, and meets service extension requirements while maximizing the values of old devices. The configuration requirements are as follows: ● Retain Router 2, which only supports IPv4 services. ● Add devices that support IPv4 and IPv6 dual topologies, and separate IPv4 route calculation and IPv6 route calculation based on different topologies.
Figure 4- 19 IS-IS MTR Topology	

Router 1	Configure IS-IS and Ethernet interfaces.
	Configure IS-IS: FS(config)# router isis FS(config-router)# net 49.0001. 0000.0000.0001.00 FS(config-router)# is-type level-1 FS(config-router)# metric-style wide FS(config-router)# address-family ipv6 FS(config-router-af)# multi-topology Configure Ethernet interfaces: FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ipv6 enable FS(config-if-GigabitEthernet 0/1)# ipv6 address 1002: : 1/112 FS(config-if-GigabitEthernet 0/1)# ipv6 router isis FS(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0 FS(config-if-GigabitEthernet 0/1)# ip router isis FS(config-if-GigabitEthernet 0/1)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)# ipv6 enable FS(config-if-GigabitEthernet 0/2)# ipv6 address 1003: : 1/112 FS(config-if-GigabitEthernet 0/2)# ipv6 router isis FS(config-if-GigabitEthernet 0/2)# ip address 192.168.2.1 255.255.255.0 FS(config-if-GigabitEthernet 0/2)# ip router isis FS(config-if-GigabitEthernet 0/2)#isis wide-metric 11
Router 2	Configure IS-IS and Ethernet interfaces.
	Configure IS-IS: FS(config)# router isis FS(config-router)# net 49.0001. 0000.0000.0002.00 FS(config-router)# is-type level-1 FS(config-router)# metric-style wide FS(config-router)#address-family ipv6 FS(config-router-af)#no adjacency-check Configure Ethernet interfaces: FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ip address 192.168.1.2 255.255.255.0

	FS(config-if-GigabitEthernet 0/1)# ip router isis FS(config-if-GigabitEthernet 0/1)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)# ip address 192.168.3.2 255.255.255.0 FS(config-if-GigabitEthernet 0/2)# ip router isis
Router 3	Configure IS-IS and Ethernet interfaces.
	Configure IS-IS: FS(config)# router isis FS(config-router)# net 49.0001. 0000.0000.0003.00 FS(config-router)# is-type level-1 FS(config-router)# metric-style wide FS(config-router)# address-family ipv6 FS(config-router-af)# multi-topology Configure Ethernet interfaces: FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ipv6 enable FS(config-if-GigabitEthernet 0/1)# ipv6 address 3001: : 1/112 FS(config-if-GigabitEthernet 0/1)# ipv6 router isis FS(config-if-GigabitEthernet 0/1)# ip address 192.168.2.3 255.255.255.0 FS(config-if-GigabitEthernet 0/1)# ip router isis FS(config-if-GigabitEthernet 0/1)#isis wide-metric 11 FS(config-if-GigabitEthernet 0/1)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)# ipv6 enable FS(config-if-GigabitEthernet 0/2)# ipv6 address 3004: : 1/112 FS(config-if-GigabitEthernet 0/2)# ipv6 router isis FS(config-if-GigabitEthernet 0/2)# ip address 192.168.4.3 255.255.255.0 FS(config-if-GigabitEthernet 0/2)# ip router isis FS(config-if-GigabitEthernet 0/2)#isis wide-metric 12
Router 4	Configure IS-IS and Ethernet interfaces.
	Configure IS-IS: FS(config)# router isis FS(config-router)# net 49.0001.0000.0000.0004.00 FS(config-router)# is-type level-1

	FS(config-router)# metric-style wide FS(config-router)# address-family ipv6 FS(config-router-af)# multi-topology Configure Ethernet interfaces: FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ipv6 enable FS(config-if-GigabitEthernet 0/1)# ipv6 address 4002: : 1/112 FS(config-if-GigabitEthernet 0/1)# ipv6 router isis FS(config-if-GigabitEthernet 0/1)# ip address 192.168.3.4 255.255.255.0 FS(config-if-GigabitEthernet 0/1)# ip router isis FS(config-if-GigabitEthernet 0/1)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)# ipv6 enable FS(config-if-GigabitEthernet 0/2)# ipv6 address 4003: : 1/112 FS(config-if-GigabitEthernet 0/2)# ipv6 router isis FS(config-if-GigabitEthernet 0/2)# ip address 192.168.4.4 255.255.255.0 FS(config-if-GigabitEthernet 0/2)# ip router isis
Verification	● Run the **show** command on Router 1 to check whether the next hop of the IPv4 route destined for Router 4 is Router 2. ● Run the show command on Router 1 to check whether the next hop of the IPv6 route destined for Router 4 is Router 3.
Checking the IPv4 route	FS#show ip route Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.1.0/24 is directly connected, GigabitEthernet 0/1 C 192.168.1.1/32 is local host. C 192.168.2.0/24 is directly connected, GigabitEthernet 0/2 C 192.168.2.1/32 is local host. i L1 192.168.3.0/24 [115/20] via 192.168.1.2, 00:13:14, GigabitEthernet 0/1 i L1 192.168.4.0/24 [115/23] via 192.168.2.3, 00:02:40, GigabitEthernet 0/2

Checking the IPv6 route	FS#show ipv6 route IPv6 routing table name is - Default - 16 entries Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary O - OSPF intra area, OI - OSPF inter area, OE1 - OSPF external type 1, OE2 - OSPF external type 2 ON1 - OSPF NSSA external type 1, ON2 - OSPF NSSA external type 2 L ::1/128 via Loopback, local host C 1002::/112 via GigabitEthernet 0/1, directly connected L 1002::1/128 via GigabitEthernet 0/1, local host C 1003::/112 via GigabitEthernet 0/2, directly connected L 1003::1/128 via GigabitEthernet 0/2, local host I1 3001::/112 [115/21] via FE80::C806:5FF:FEE8:38, GigabitEthernet 0/2 I1 3004::/112 [115/21] via FE80::C806:5FF:FEE8:38, GigabitEthernet 0/2 I1 4002::/112 [115/31] via FE80::C806:5FF:FEE8:38, GigabitEthernet 0/2 I1 4003::/112 [115/31] via FE80::C806:5FF:FEE8:38, GigabitEthernet 0/2 L FE80::/10 via ::1, Null0 C FE80::/64 via GigabitEthernet 0/2, directly connected L FE80::1614:4BFF:FE12:ADFC/128 via GigabitEthernet 0/2, local host C FE80::/64 via GigabitEthernet 0/1, directly connected L FE80::1614:4BFF:FE12:ADFD/128 via GigabitEthernet 0/1, local host C FE80::/64 via Local 0, directly connected L FE80::1614:4BFF:FE12:ADFC/128 via Local 0, local host

Common Errors

- **metric-style** is not set to **Wide** or **Transition**.
- The protocol types used by two neighbors do not match; therefore, a neighbor relationship cannot be established.
- The IP addresses of the interfaces connected between neighbors are not in the same network segment.
- The **ip router isis** command is not executed on interfaces.
- No NET address is configured, or different NET addresses exist at Level 1.
- **max-area-addresses** is configured differently on both sides.
- **metric-style** is configured differently on both sides.
- The interface Levels on both sides are different. One side is Level-1, whereas the other side is Level-2.
- One side is configured with the P2P mode, whereas the other side is configured with the broadcast mode.
- One side is enabled with authentication, whereas the other side is not.

4.4.12 Configuring SNMP for IS-IS

Configuration Effect

- By default, the SNMP software can perform the MIB operation on the first IS-IS instance. To perform the MIB operation on other instances, you need to manually specify these instances.

Notes

- By default, the SNMP software can perform the MIB operation on the first displayed IS-IS instance.

Configuration Steps

Binding the Instances on Which the IS-IS MIB Operation Will Be Performed

- Perform this configuration based on requirements.
- Run the **enable mib-binding** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Enabling IS-IS Trap Globally

- Perform this configuration based on requirements.
- Run the **snmp-server enable traps isis** command in global configuration mode on the desired device, unless otherwise specified.

Configuring an SNMP Host Globally

- Perform this configuration based on requirements.
- Run the **snmp-server host** command in global configuration mode on the desired device, unless otherwise specified.

Allowing the Sending of all IS-IS Trap Messages to the SNMP Host

- Perform this configuration based on requirements.
- Run the **enable traps all** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- Use the MIB tool to read and write IS-IS settings.

Related Commands

Binding the Instances on Which the IS-IS MIB Operation Will Be Performed

Command	**enable mib-binding**
Parameter Description	N/A
Command Mode	IS-IS routing process configuration mode
Usage Guide	The latest standards stipulate that the MIB operation can be performed on a single instance. By default, the MIB operation is performed on the first displayed IS-IS instance. Because multiple IS-IS instances can be configured, the administrator can use this command to specify the instances on which the MIB operation will be performed.

Enabling IS-IS Trap Globally

Command	**snmp-server enable traps** [*isis*]

Parameter Description	*isis*: Enables IS-IS event trap.
Command Mode	Global configuration mode
Usage Guide	This command must be used with the **snmp-server host** command in global configuration mode so that trap messages can be sent.

Configuring an SNMP Host Globally

Command	**snmp-server host** { *host-addr* | **ipv6** *ipv6-addr* } [**vrf** *vrfname*] [**traps**] [**version** { **1** | **2c** | **3** { **auth** | **noauth** | **priv** }] *community-string* [**udp-port** *port-num*] [*notification-type*]
Parameter Description	*host-addr*: Indicates the address of the SNMP host. *ipv6-addr*: Indicates the IPv6 address of the SNMP host. *vrfname*: Indicates the name of a VRF table. **version**: Indicates the SNMP version, which can be set to **V1**, **V2C**, or **V3** **auth** | **noauth** | **priv**: Indicates the security level of V3 users. *community-string*: Indicates the community string or user name (V3 version). *port-num*: Indicates the port number of the SNMP host. *notification-type*: Indicates the type of trap messages that are actively sent, for example, snmp.
Command Mode	Global configuration mode
Usage Guide	This command is used with the **snmp-server enable traps** command to actively send trap messages to a Network Management System (NMS). You can configure different SNMP hosts to receive trap messages. A host supports different trap types, ports, and VRF tables. For the same host (with the same port configuration and VRF configuration), the last configuration is combined with the previous configurations. That is, to send different trap messages to the same host, configure a type of trap messages each time. These configurations are finally combined.

Allowing the Sending of Trap Messages

Command	**enable traps** { **all** | *traps set* }
Parameter Description	**all**: Indicates all trap messages. *traps set*: Indicates a trap message type in any set.
Command Mode	IS-IS routing process configuration mode
Usage Guide	IS-IS packets are classified into 18 types of trap messages, which are grouped into several sets, with each set containing several trap message types. To enable the sending of IS-IS trap messages, run the **snmp-server enable traps isis** command in global configuration mode and specify the recipient host and the type of trap messages that can be sent.

Configuration Example

Configuring IS-IS SNMP

Configuration Steps	● Bind the instances on which the IS-IS MIB operation will be performed. ● Complete trap message-related settings.
	A(config)# router isis A(config-router)# enable mib-binding A# configure terminal A(config)#snmp-server enable traps isis A(config)#snmp-server host 10.1.1.1 traps version 2c public A(config)#router isis A(config-router)# enable traps all
Verification	Run the MIB tool to read and write IS-IS settings.
	A# show running-config

4.4.13 Configuring IS-IS to Enable Super VLAN

Configuration Effect

- Run the ISIS protocol on super VLANs.

Notes

- The ISIS basic functions must be configured.
- The designated sub VLAN is connected with neighbors.

Configuration Steps

Running ISIS on Super VLAN

- Optional. Run this command to enable ISIS on a super VLAN if required.

Verification

- Run the **show isis neighbor** command to display the protocol status.
- Run the **show isis interface** command to view interface configuration.

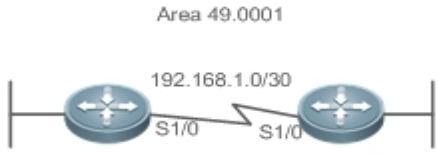
Related Commands

Running ISIS on Super VLAN

Command	**isis subvlan [all | *vid*]**
Parameter Description	**all:** Indicates that packets are allowed to be sent to all sub VLANs. *vid:* Specifies the sub VLAN ID. The value ranges from 1 to 4094.

Command Mode	Interface configuration mode
Usage Guide	In normal cases, a super VLAN contains multiple sub VLANs. Multicast packets of a super VLAN are also sent to its sub VLANs. In this case, when ISIS multicast packets are sent over a super VLAN containing multiple sub VLANs, the ISIS multicast packets are replicated multiple times, and the device processing capability is insufficient. As a result, a large number of packets are discarded, causing the neighbor down error. In most scenarios, the ISIS function does not need to be enabled on a super VLAN. Therefore, the ISIS function is disabled by default. However, in some scenarios, the ISIS function must be run on the super VLAN, but packets only need to be sent to one sub VLAN. In this case, run this command to specify a particular sub VLAN. You must be cautious in configuring packet transmission to all sub VLANs, as the large number of sub VLANs may cause a device processing bottleneck, which will lead to the neighbor down error.

Configuration Example

Scenario	
Configuration Steps	● Configure the ISIS basic functions on all devices. ● Specify a particular sub VLAN on all devices.
A	A# configure terminal A(config)# interface VLAN 300 A(config-if-VLAN 300)# isis subvlan 1024
B	B# configure terminal B(config)# interface VLAN 300 B(config-if-VLAN 300)# isis subvlan 1024
Verification	● Verify that an ISIS interface neighbor is established on Device A. ● Verify ISIS interface configuration on Device A.
A	A# show isis neighbor A# show isis interface

4.4.14 Configuring IS-IS Two-way Maintenance

Configuration Effect

- Enable IS-IS two-way maintenance.

Notes

- The IS-IS basic functions must be configured.
- The neighbor relationship is successfully established.

Configuration Steps

Configuring IS-IS Two-way Maintenance

- Configure the two-way maintenance function as required.
- Run the **two-way-maintain** command in IS-IS routing process configuration mode on the required devices unless otherwise specified.

Verification

- Run the **show isis neighbor** command to check the neighbor update time.
- Run the **show isis protocol** command to check whether the two-way maintenance function is enabled.

Related Commands

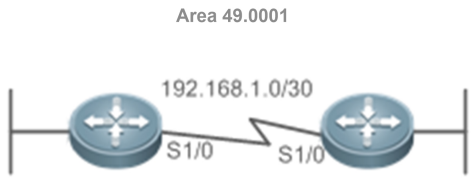
Configuring IS-IS Two-way Maintenance

Command	**two-way-maintain**
Parameter Description	N/A
Command Mode	IS-IS routing process configuration mode
Usage Guide	In a large-scale network, a large number of packets are sent and received, which occupies lots of CPU and memory resources, causing some IS-IS packets to be delayed or discarded. If the time required for processing HELLO packets exceeds the neighbor relationship maintenance time, the corresponding neighbor relationship times out and is removed. When the two-way maintenance function is enabled, if a large number of packets exist on the network, the LSP packets, CSNP packets, and PSNP packets from a neighbor in addition to HELLO packets can also be used to maintain the two-way relationship with the neighbor, preventing neighbor failure caused by delay or discard of HELLO packets.

Configuration Example

The following example is implemented based on IS-IS basic functions. For details about the IS-IS basic functions, see preceding description

Configuring IS-IS Two-way Maintenance

Scenario	
Configuration Steps	- Configure the ISIS basic functions on all devices. - The neighbor relationship is successfully established.

Verification	Verify that an ISIS interface neighbor is established on device A. Check the status of an ISIS instance on device A.

Common Errors

4.4.15 Configuring Other IS-IS Parameters

Configuration Effect

- **maximum-paths:** Configures the maximum number of IS-IS equal-cost paths to be installed to a routing table.
- **lsp-length receive:** Configures the maximum length allowed for received LSPs.
- **lsp-length originate:** Configures the maximum length allowed for sent LSPs.
- **passive-interface:** Prevents passive interfaces from receiving and sending IS-IS packets. That is, IS-IS neighbor relationships will not be established on passive interfaces. The IP addresses of passive interfaces are flooded through other interfaces.
- **isis metric**: Stores the metric, which is used in SPF calculation, in the IP reachability information TLV. The greater the metric, the greater the routing consumption of the interface and the longer the path obtained by SPF calculation.
- **isis priority:** In a broadcast network, IS-IS needs to elect a DIS among all devices. The DIS will generate a pseudonode and related LSPs. The device with the highest priority is elected as the DIS. You can configure different priorities for different Levels.
- **default-information originate:** Generates a Level-2 default route, which will be advertised through LSPs.
- **spf-interval:** Configures the exponential backoff algorithm of SPF.
- **summary-address** and **summary-prefix:** Creates a summary route to represent a group of routes in a routing table. A summary route can include multiple routes of the specified Level. The interface metric of the summary route follows the smallest interface metric among all routes.
- **log-adjacency-changes:** Enables neighbor relationship event output to log IS-IS neighbor relationship changes.
- **redistribute:** Redistributes other routes to IS-IS; redistributes Level-1 routes to Level-2; redistributes Level-2 routes to Level-1.

Configuration Steps

Configuring the Maximum Number of Equal-Cost Paths

- Perform this configuration based on requirements.
- Run the **maximum-paths** command in IS-IS routing process configuration mode or IS-IS address-family ipv6 configuration mode on the desired device, unless otherwise specified.

Configuring the Maximum Length Allowed for Received LSPs

- Perform this configuration based on requirements.
- Run the **lsp-length receive** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring the Maximum Length Allowed for Sent LSPs

- Perform this configuration based on requirements.

- Run the **lsp-length originate** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring a Passive Interface

- Perform this configuration based on requirements.
- Run the **passive-interface** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring the IS-IS Interface Metric

- Perform this configuration based on requirements.
- Run the **isis metric** command in interface configuration mode on the desired device, unless otherwise specified.

Configuring the Priority of the DIS

- Perform this configuration based on requirements.
- Run the **isis priority** command in interface configuration mode on the desired device, unless otherwise specified.

Configuring the SPF Calculation Cycle

- Perform this configuration based on requirements.
- Run the **spf-interval** command in interface configuration mode on the desired device, unless otherwise specified.

Generating a Default Route

- Perform this configuration based on requirements.
- Run the **default-information originate** command in IS-IS routing process configuration mode or IS-IS address-family ipv6 configuration mode on the desired device, unless otherwise specified.

Configure a Summary Route

- Perform this configuration based on requirements.
- Run the **summary-address** and **summary-prefix** commands in IS-IS routing process configuration mode or IS-IS address-family ipv6 configuration mode on the desired device, unless otherwise specified.

Enabling Neighbor Relationship Event Output

- Perform this configuration based on requirements.
- Run the **log-adjacency-changes** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Configuring Route Redistribution

- Perform this configuration based on requirements.
- Run the **redistribute** command in IS-IS routing process configuration mode on the desired device, unless otherwise specified.

Verification

- **maximum-paths:** Check whether the maximum number of equal-cost paths displayed by routing entries is the same as the configuration.
- **lsp-length receive:** Capture packets to check the length of LSPs.
- **lsp-length originate:** Capture packets to check the length of LSPs.
- **passive-interface:** Capture packets to check whether the interface receives and sends IS-IS packets.
- **isis metric:** Check the database details of IS-IS.
- **isis priority:** Check whether the device with the changed priority setting is elected as the DIS.
- **default-information originate:** Check whether a default route is generated.
- **spf-interval:** Check whether the SPF calculation cycle works.
- **summary-address and summary-prefix:** Capture packets to check whether the summary route instead of detailed routes is advertised through LSPs.
- **log-adjacency-changes:** Change the neighbor state and verify that the change is recorded when debugging is disabled.
- **redistribute:** Check IS-IS routing entries.

Related Commands

Configuring the Maximum Number of Equal-Cost Paths

Command	**maximum-paths** *maximum*
Parameter Description	*maximum*: Indicates the maximum number of IS-IS equal-cost routes to be installed to a routing table. The value range is 1 to device capacity.
Command Mode	IS-IS routing process configuration mode and IS-IS address-family ipv6 configuration mode
Usage Guide	This command is used by IS-IS to control the number of IS-IS equal-cost paths to be installed to a routing table. The routing table also has a command used to control the number of equal-cost paths. The number of effective equal-cost paths is determined by either of the two command values, whichever is smaller.

Configuring the Maximum Length Allowed for Received LSPs

Command	**lsp-length receive** *size*
Parameter Description	*size*: Indicates the maximum length allowed for received LSPs. According to RFC, the value range is 1,492 to 16,000, in the unit of bytes.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Use this command to control the maximum length allowed for LSPs received by the local device. Intermediate nodes with sufficient memory are required to receive LSPs whose maximum length is equal to the interface MTU in order to avoid a route convergence failure. From this perspective, the command is meaningless. The maximum length allowed for received LSPs cannot be smaller than that allowed for sent LSPs; otherwise, the former will be automatically adjusted to be equal to the latter.

Configuring the Maximum Length Allowed for Sent LSPs

Command	**lsp-length originate** *size* [**level-1** | **level-2**]

Parameter Description	*size*: Indicates the maximum length allowed for sent LSPs. The value range is 512 to 16,000, in the unit of bytes. **level-1**: Applies the setting only to Level-1 LSPs. **level-2**: Applies the setting only to Level-2 LSPs.
Command Mode	IS-IS routing process configuration mode
Usage Guide	In principle, the maximum length of LSPs and SNPs cannot be greater than the interface MTU; otherwise, the packets will be discarded when being sent.

Configuring a Passive Interface

Command	**passive-interface** [**default**] { *interface-type interface-number* }
Parameter Description	**default**: Configures all IS-IS interfaces that are not enabled as passive interfaces. *interface-type*: Indicates the interface type. *interface-number*: Indicates the interface number.
Command Mode	IS-IS routing process configuration mode
Usage Guide	This command prevents the specified interface from receiving and sending IS-IS packets, but the IP address of the interface will be flooded by other interfaces. If the **default** option is selected and there are more than 255 IS-IS interfaces not enabled, only the first 255 interfaces will be configured as passive interfaces. The remaining interfaces are non-passive interfaces.

Configuring the IS-IS Interface Metric

Command	**isis metric** *metric* [**level-1| level-2**]
Parameter Description	*metric*: Indicates the metric value. The value range is 1 to 63. The default value is 10. **level-1**: Applies the setting to Level-1 circuits. **level-2**: Applies the setting to Level-2 circuits.
Command Mode	Interface configuration mode
Usage Guide	The metric, which is used in SPF calculation, is stored in the IP reachability information TLV. The greater the metric, the greater the routing consumption of the interface and the longer the path obtained by SPF calculation. The metric belongs to the narrow type and is valid only when **metric-style** is set to **Narrow**.

Configuring the Wide Metric of an Interface

Command	**isis wide-metric** *metric* [**level-1 | level-2**]
Parameter Description	*metric*: Indicates the metric value. The value range is 1 to 16,777,214. The default value is 10. **level-1**: Applies the setting to Level-1 circuits. **level-2**: Applies the setting to Level-2 circuits.
Command Mode	Interface configuration mode
Usage Guide	The metric, which is used in SPF calculation, is stored in the IP reachability information TLV. The greater the metric, the greater the routing consumption of the interface and the longer the path obtained by SPF calculation. The metric is valid only when **metric-style** is set to **Wide**.

Configuring the Priority of the DIS

Command	**isis priority** *value* [**level-1 | level-2**]
Parameter Description	*value*: Indicates the priority. The value range is 0 to 127. The default value is 64. **level-1**: Applies the setting to Level-1 circuits. **level-2**: Applies the setting to Level-2 circuits.
Command Mode	Interface configuration mode
Usage Guide	Use this command to change the priority carried in Hello packets in a LAN. The device with a lower priority is less likely to be elected as the DIS. The command is invalid on a P2P network interface. The **no isis priority** command, with or without parameters, restores the priority to its default value. To change the configured priority, run the **isis priority** command with the priority specified to overwrite the existing configuration, or you can first restore the priority to its default value and then configure a new priority.

Generating a Default Route

Command	**default-information originate** [**route-map** *map-name*]
Parameter Description	**route-map** *map-name*: Associates with a route map.
Command Mode	IS-IS routing process configuration mode and IS-IS address-family ipv6 configuration mode
Usage Guide	Because Level-2 domains do not generate any default route, use this command to allow a default route to enter a Level-2 domain.

Configuring SPF Calculation Cycle

Command	**spf-interval** [**level-1 | level-2**] *maximum-interval* [*initial-interval hold-interval*]
Parameter Description	**level-1**: Applies the configuration only to Level-1. **level-2**: Applies the configuration only to Level-2. *maximum-interval*: Indicates the maximum interval for performing two consecutive SPF calculations. The value range is **1** to **120** (in seconds). The default value is **10**. *initial-interval*: Indicates the waiting time for performing the SPF calculation for the first time. The value range is **0** to **60000** (in milliseconds). The default value is **50**. *hold-interval*: Indicates the minimum interval for performing the SPF calculation for the second time. The value range is **10** to **60000** (in milliseconds). The default value is **200**.
Command Mode	IS-IS routing process configuration mode
Usage Guide	Increasing the maximum interval for performing SPF calculations can avoid frequent SPF calculations and waste of CPU resources. However, a larger minimum interval also leads to slower responses to route changes.
	The waiting time for performing the SPF calculation for the first time is the initial interval. If the network becomes unstable, the SPF calculation interval is less than the maximum interval, and the interval for performing the SPF calculation for the second time becomes the hold interval. A corresponding penalty is added to this interval: The next interval for the SPF calculation doubles the previous interval for the same SPF calculation, until the SPF calculation

	interval reaches the maximum interval. Subsequent SPF calculations are performed at the maximum interval. When the network becomes stable, the interval for performing the SPF calculation becomes greater than the maximum interval, and the waiting time for performing the SPF calculation is restored to the initial interval. Link changes have high requirements for convergence. The initial interval can be set to a small value. The preceding parameters can also be adjusted to larger values to reduce CPU consumption. The value of **initial-interval** cannot be greater than that of **maximum-interval**. Otherwise, the value of **initial-interval** will be used as the value of **maximum-interval**. The value of **hold-interval** cannot be greater than that of **maximum-interval**. Otherwise, the value of **hold-interval** will be used as the value of **maximum-interval**. The value of **initial-interval** cannot be greater than that of **hold-interval**. Otherwise, the value of **initial-interval** will be used as the value of **hold-interval**.

Configuring an IPv4 Summary Route

Command	**summary-address** *ip-address net-mask* [**level-1 | level-2 | level-1-2**] [metric *number*]
Parameter Description	*ip-address*: Indicates the IP address of the summary route. *net-mask*: Indicates the subnet mask of the summary route. **level-1**: Applies the setting only to Level-1. **level-2**: Applies the setting only to Level-2. By default, the setting takes effect for Level-2. **level-1-2**: Applies the setting to Level-1 and Level-2. *number*: Indicates the metric of the summary route.
Command Mode	IS-IS routing process configuration mode
Usage Guide	If the configured summary route contains routing information about a reachable address or network segment, the summary route, instead of detailed routes, is advertised externally.

Configuring an IPv6 Summary Route

Command	**summary-prefix** *ipv6-prefix/prefix-length* [**level-1 | level-2 | level-1-2**]
Parameter Description	*ipv6-prefix/prefix-length*: Indicates the network address of the summary route and its IPv6 prefix length. The address format is X:X:X:X::X/<0-128>. **level-1**: Applies the setting only to Level-1. **level-2**: Applies the setting only to Level-2. By default, the setting takes effect for Level-2. **level-1-2**: Applies the setting to Level-1 and Level-2.
Command Mode	IS-IS address-family ipv6 configuration mode
Usage Guide	If the configured summary route contains routing information about a reachable address or network segment, the summary route, instead of detailed routes, is advertised externally.

Enabling Neighbor Relationship Event Output

Command	**log-adjacency-changes**
Parameter Description	N/A
Command	IS-IS routing process configuration mode

Mode	
Usage Guide	You can also use the **debug** command to record IS neighbor state changes, but the command consumes many system resources.

Redistributing Other Routes to IS-IS

Command	**redistribute** { **bgp** | **ospf** *process-id* [**match** { **internal** [**external** [**1 | 2**]] [**nssa-external** [**1 | 2**]] | **external** [**1 | 2**] [**internal**] [**nssa-external** [**1 | 2**]] **| nssa-external** [**1 | 2**] [**internal**] [**external** [**1 | 2**]] }] | **rip** | **connected** | **static** } [**metric** *metric-value*] [**metric-type** *type-value*] [**route-map** *map-tag*] [**level-1 | level-1-2 | level-2**]
Parameter Description	*process-id*: Indicates the OSPF process ID. The range is 1 to 65,535. **match** { **internal** | **external** [**1 | 2**] **| nssa-external** [**1 | 2**] }: When OSPF routes are redistributed, the routes are filtered by subtype. If the **match** option is not selected, routes of all OSPF types will be received. If **match external** is not followed by the number 1 or 2, OSPF routes specified by **external 1** and **external 2** will be redistributed. If **match nssa-external** is not followed by the number 1 or 2, OSPF routes specified by **nssa-external 1** and **nssa-external 2** will be redistributed. **metric** *metric-value*: Indicates the metric of redistributed routes. The value range is 0 to 4,261,412,864. The metric of external routes is used when the **metric** option is not specified. **metric-type** { **internal | external** }: Indicates the metric type of redistributed routes. **internal**: Indicates that the metric belongs to the internal type. **external**: Indicates that the metric belongs to the external type. If **metric-type** is not specified, the metric belongs to the internal type. **route-map** *map-tag*: Indicates the route map used for external route redistribution. It is used to filter redistributed routes or configure the attributes of redistributed routes. The value of *map-tag* cannot exceed 32 characters. By default, **route-map** is not configured. **level-1 | level-1-2 | level-2**: Indicates the Level of redistributed routes received by IS-IS. If no Level is specified, routes are redistributed to Level-2. **level-1**: Redistributes routes to Leve-1. **level-1-2**: Redistributes routes to Level-1 and Level-2. **level-2**: Redistributes routes to Leve-2.
Command Mode	IS-IS routing process configuration mode and IS-IS address-family ipv6 configuration mode
Usage Guide	The **no redistritbue** { **bgp** | **ospf** *processs-id* | **rip** | **connected** | **static** } command is used to cancel the redistribution of routes mapped to the specified protocol. If **no redistribute** is followed by other parameters, the command will restore the default parameter settings, rather than cancel route redistribution. For example, **no redistribute bgp** cancels BGP route redistribution, whereas **no redistribute bgp route-map aa** cancels the route map named **aa** used for BGP route redistribution. When external routes are redistributed in IPv4 mode, the routing information is stored in LSPs' IP External Reachability Information TLV. When external routes are redistributed in IPv6 mode, the routing information is stored in LSPs' IPv6 Reachable TLV. In the old versions of some vendors, if **metric-type** is set to **external**, the metric of redistributed routes is added by 64 during route calculation and used to determine routing. This practice does not comply with the related protocol. In the actual application, external routes may be preferred over internal routes. If this happens during interworking with old versions of some vendors, you can modify the related setting (such as **metric** or **metric-type**) of each device to ensure that internal routes are preferred over external routes.

Redistributing the Level-1 Reachable Routing Information of the Specified IS-IS Instance to Level-2 of the Current Instance

Command	**redistribute isis** [*tag*] **level-1 into level-2** [**route-map** *route-map-name* | **distribute-list** *access-list-name*]
Parameter Description	*tag*: Indicates the name of the IS-IS instance whose routing information will be redistributed. **route-map** *route-map-name*: Indicates the route map used for rout redistribution. It is used to filter redistributed routes or configure the attributes of redistributed routes. The value of *route-map-name* cannot exceed 32 characters. By default, **route-map** is not configured. **distribute-list** *access-list-name*: Filters redistributed routes by using **distribute-list**. *access-list-name* indicates the associated prefix list, which can be a standard prefix list, an extended prefix list, or a name prefix list. It is in the format of {*<1-99>* | *<100-199>* | *<1300-1999>* | *<2000-2699>* | *acl-name* }. When the IS-IS address-family ipv6 configuration mode is applied, only the name prefix list can be used, in the format of *acl-name*.
Command Mode	IS-IS routing process configuration mode and IS-IS address-family ipv6 configuration mode
Usage Guide	You can use the **route-map** or **distribute-list** parameter to filter the specified instance's Level-1 routes to be redistributed. Only the routes that meet specific criteria can be redistributed to Level-2 of the current instance. The **route-map** and **distribute-list** parameters cannot be used at the same time. The **no redistritbue isis** [*tag*] **level-2 into level-1** command is used to cancel the redistribution of the specified instance's routes. If **no redistribute** is followed by other parameters, the command will restore the default parameter settings, rather than cancel route redistribution. For example, **no redistritbue isis** *tag1* **level-1 into level-2** cancels the redistribution of the routes of the IS-IS instance name **tag1**. **no redistritbue isis** *tag1* **level-1 into level-2 route-map** *aa* cancels the use of the route map named **aa** to filter redistributed routes.

↘ Redistributing the Level-2 Reachable Routing Information of the Specified IS-IS Instance to Level-1 of the Current Instance

Command	**redistribute isis** [*tag*] **level-2 into level-1** [**route-map** *route-map-name* | **distribute-list** *access-list-name* |**prefix** *ip-address net-mask*]
Parameter Description	*tag*: Indicates the name of the IS-IS instance whose routing information will be redistributed. **route-map** *route-map-name*: Indicates the route map used for route redistribution. It is used to filter redistributed routes or configure the attributes of redistributed routes. The value of *route-map-name* cannot exceed 32 characters. By default, **route-map** is not configured. **Distribute-list** *access-list-name*: Filters redistributed routes by using **distribute-list**. *access-list-name* indicates the associated prefix list, which can be a standard prefix list, an extended prefix list, or a name prefix list. It is in the format of {*<1-99>* | *<100-199>* | *<1300-1999>* | *<2000-2699>* | *acl-name* }. When the IS-IS address-family ipv6 configuration mode is applied, only the name prefix list can be used, in the format of *acl-name*. **prefix** *ip-address net-mask*: Determines the routes to be redistributed by address and prefix length.
Command Mode	IS-IS routing process configuration mode and IS-IS address-family ipv6 configuration mode
Usage Guide	You can use the **route-map**, **distribute-list**, or **prefix** parameter to filter the specified instance's Level-2 routes to be redistributed. Only the routes that meet specific criteria can be redistributed to Level-1 of the current instance. The **no redistritbue isis** [**tag**] **level-2 into level-1** command is used to cancel the redistribution of the specified instance's routes. If **no redistribute** is followed by other parameters, the command will restore the default parameter settings, rather than cancel route redistribution. For example: **no redistritbue isis** *tag1* **level-2 into level-1** cancels the redistribution of the routes of the IS-IS instance name **tag1**. **no**

	redistritbue isis *tag1* **level-2 into level-1 route-map** *aa* cancels the use of the route map named **aa** to filter redistributed routes.

Configuration Example

Configuring the Maximum Number of Equal-Cost Paths

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the maximum number of equal-cost paths.
	A(config)# router isis A(config-router)# maximum-paths 5
Verification	● Check whether the maximum number of equal-cost paths displayed by routing entries is the same as the configuration.
	A# show ip route isis

Configuring the Maximum Length Allowed for Received LSPs

Configuration	● Configure IS-IS neighbors. (Omitted) ● Configure the maximum length allowed for received LSPs.
	A(config)# router isis A(config-router)# lsp-length receive 512
Verification	Capture packets to check the length of received LSPs.

Configuring the Maximum Length Allowed for Sent LSPs

Configurations	● Configure IS-IS neighbors. (Omitted) ● Configure the maximum length allowed for sent LSPs.
	A# configure terminal A(config)# router isis 1 A(config-router)# lsp-length originate 512 level-2
Verification	Capture packets to check the length of sent LSPs.

Configuring a Passive Interface

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure a passive interface.
	A# configure terminal A(config)# router isis 1 A(config-router)# passive-interface GigabitEthernet 0/0
Verification	Capture packets to check whether the interface receives and sends IS-IS packets.

Configuring the Metric of an IS-IS Interface

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure metric of the IS-IS interface.
	A(config)# interface GigabitEthernet 0/1 A(config-if)#isis metric 1
Verification	Check the database details of IS-IS.
	A# show isis database detail

Configuring the Priority of the DIS

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configure the priority of the DIS.
	A(config)# interface GigabitEthernet 0/1 A(config-if)# isis priority 127 level-1
Verification	Check whether the device with the changed priority setting is elected as the DIS.
	A# show isis database detail

Generating a Default Route

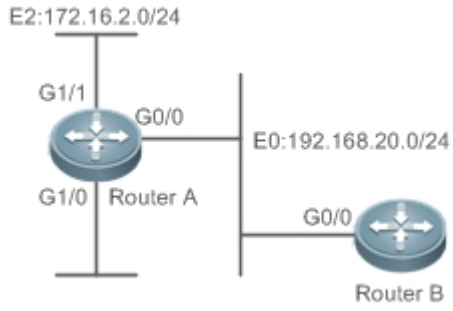
Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Generate a default route.
	A(config)# router isis A(config-router)# default-information originate

Verification	Capture packets to check whether the sent LSP contains a default route.

Configuring SPF Calculation Cycle

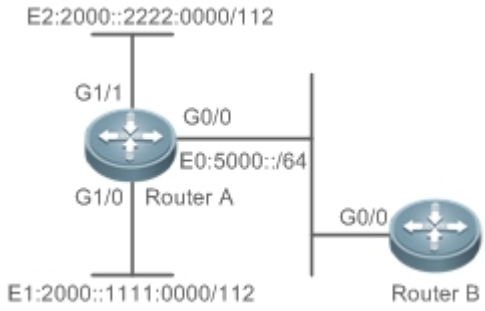
Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Configuring SPF calculation cycle.
	A(config)# router isis A(config-router)# spf-interval 5 100 200
Verification	Check whether the SPF calculation cycle works.

Configuring an IS-IS Summary Route

Configuration Requirements	Router A and Router B are connected through the Ethernet and run IS-IS. Configure Router A to advertise only the 172.16.0.0/22 route instead of the 172.16.1.0/24 and 172.16.2.0/24 routes.
Figure 4- 20 IS-IS Route Summary Topology	
Configuration Steps	● Configure IS-IS. ● Configure Ethernet interfaces. ● Configure the password for IS-IS authentication.
A	Configure IS-IS.
	A(config)# router isis A(config-router)# net 49.0001.0000.0000.0001.00 A(config-router)# summary-address 172.16.0.0/16 level-1-2
	Configure Ethernet interfaces. A(config)# interface GigabitEthernet 0/0 A(config-if)# ip address 192.168.20.1 255.255.255.0 A(config-if)# ip router isis A(config)# interface GigabitEthernet 1/0

	A(config-if)# ip address 172.16.1.1 255.255.255.0 A(config-if)# ip router isis A(config)# interface GigabitEthernet 1/1 A(config-if)# ip address 172.16.2.1 255.255.255.0 A(config-if)# ip router isis
B	Configure IS-IS.
	B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00
	Configure an Ethernet interface.
	B(config)# interface GigabitEthernet 0/0 B(config-if)# ip address 192.168.20.2 255.255.255.0 B(config-if)# ip router isis
Verification	Run the **show ip route** command on Router B to check whether only one summary route exists.
B	B(config)# show ip route i L1 172.16.0.0/16 [115/20] via 192.168.20.1, FastEthernet0/0

Configuring an IS-ISv6 Summary Route

	Router A and Router B are connected through the Ethernet and run IS-ISv6. Configure Router A to advertise only the 2000::/96 route instead of the 2000::1111:0/112 and 2000::2222::0/112 routes.
Figure 4-21 IS-ISv6 Route Summary Topology	
Configuration Steps	● Configure IS-IS. ● Configure Ethernet interfaces. ● Configure the password for IS-IS authentication.
A	Configure IS-IS.
	A(config)# ipv6 unicast-routing

	A(config)# router isis A(config-router)# net 49.0001.0000.0000.0001.00 A(config-router)# address-family ipv6 unicast A (config-router-af)# summary-prefix 2000::/96 level-1-2 A (config-router-af)# exit-address-family
	Configure Ethernet interfaces. A(config)# interface GigabitEthernet 0/0 A(config-if)# ipv6 address 5000::1/64 A(config-if)# ipv6 router isis A(config)# interface GigabitEthernet 1/0 A(config-if)# ipv6 address 2000::1111:0001/112 A(config-if)# ipv6 router isis A(config)# interface GigabitEthernet 1/1 A(config-if)# ipv6 address 2000::2222:0001/112 A(config-if)# ipv6 router isis
B	Configure IS-IS.
	B(config)# ipv6 unicast-routing B(config)# router isis B(config-router)# net 49.0001.0000.0000.0002.00
	Configure an Ethernet interface.
	B(config)# interface GigabitEthernet 0/0 B(config-if)# ipv6 address 5000::2/64 B(config-if)# ipv6 router isis
Verification	Run the **show ipv6 route** command on Router B to check whether only one summary route exists.
B	B(config)# show ipv6 route I1 2000::/96 [115/20] via FE80::C800:1BFF:FEF8:1C, FastEthernet1/0

Enabling Neighbor Relationship Event Output

Configuration Steps	● Configure IS-IS neighbors. (Omitted) ● Enable neighbor relationship event output.
	A(config-router)# log-adjacency-changes

Verification	Change the neighbor state and verify that the change is recorded when debugging is disabled.

Configuring Route Redistribution

Configuration Steps	• Configure IS-IS neighbors. (Omitted) • Configure OSPF routes. (Omitted) • Configure route redistribution
	A(config)# router isis A(config-router)# redistribute ospf 1 metric 10 level-1
Verification	Check whether routing entries with redistributed routes exist.
	A# show ip route isis

4.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears all IS-IS neighbor relationship tables.	**clear clns neighbors**
Clears all IS-IS data structures.	**clear isis** *
Clears all IS-IS counters.	**clear isis** [*tag*] **counter**

Displaying

Description	Command
Displays all IS neighbors and inter-device neighbor relationships.	**show clns** [*tag*] **is-neighbors** [*interface-type interface-number*] [**detail**]
Displays all IS neighbors and provides device information and information about the neighbor relationship with ESs.	**show clns** [*tag*] **neighbors** [*interface-type interface-number*] [**detail**]
Displays all IS-IS counters.	**show isis** [*tag*] **counter**
Displays the LSDB information.	**show isis** [*tag*] **database** [*FLAGS*] [*LEVEL*] [*LSPID*]
Displays the state information related to IS-IS GR.	**show isis** [*tag*] **graceful-restart**
Displays the relationship between the device name and system ID.	**show isis** [*tag*] **hostname**

Displays the details of an IS-IS interface.	**show isis** [*tag*] **interface** [*interface-type interface-number*] [*counter*]
Displays the mesh group configuration of all interfaces.	**show isis** [*tag*] **mesh-groups**
Displays IS-IS neighbor information.	**show isis** [*tag*] **neighbors** [**detail**]
Displays the neighbor information of virtual systems in IS-IS.	**show isis** [*tag*] **virtual-neighbors**
Displays IS-IS information.	**show isis** [*tag*] **protocol**
Displays the topology of IS-IS device connection.	**show isis** [*tag*] **topology** [**l1** | **l2** | **level-1** | **level-2**]
Displays information of an IS-IS IPv6 unicast topology.	**show isis** [*tag*] **ipv6 topology** [**l1** | **l2** | **level-1** | **level-2**]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Enables IS-IS debugging.	**debug isis** { **all** | **auth** | **events** | **gr** | **ifsm** | **lsp** | **mtr** | **nfsm** | **nsm** | **pdu** | **spf** | **warn** }

5 Configuring BGP

5.1 Overview

The Border Gateway Protocol (BGP) is an Exterior Gateway Protocol (EGP) used for communication between routers in different autonomous systems (ASs). BGP is used to exchange network accessibility information between different ASs and eliminate routing loops by using its own mechanism.

BGP uses TCP as the transmission protocol. The reliable transmission mechanism of TCP is used to ensure the transmission reliability of BGP.

Routers running BGP are called BGP speakers. BGP speakers between which a BGP session is established are called BPG peers.

Two modes can be used to establish peers between BGP speakers: Internal BGP (IBGP) and External BGP (EBGP).

- IBGP refers to a BGP connection established within an AS and completes transition of routing information within the AS.
- EBGP refers to a BGP connection established between different ASs and completes exchange of routing information between different ASs.

Rules for BGP to select an optimum route:

5. Invalid routing table entries are not involved in optimum route selection.

Invalid entries include entries of inaccessible next hops and flapping entries.

6. Otherwise, select a route with a large value of **LOCAL_PREF**.
7. Otherwise, select a route generated by a BGP speaker.

Routes generated by a BGP speaker include routes generated by the **network**, **redistribute** and **aggregate** commands.

8. Otherwise, select a route with the shortest AS length.

9. Otherwise, select a route with a smaller value of **ORIGIN**.

10. Otherwise, select a route with the smallest value of **MED**.

11. Otherwise, EBGP routes have higher priorities than IBGP routes and routes in the AS alliance, and the IBGP routes have the same priorities as the routes in the AS alliance.

12. Otherwise, select a route with the smallest IGP metric value to the next hop.

13. Otherwise, select an EBGP route that is received first.

14. Otherwise, select a route advertised by a BGP speaker with a smaller router ID.

15. Otherwise, select a route with a large cluster length.

16. Otherwise, select a route with a large neighbor address.

The preceding shows the route selection process under the default configurations. By using CLI commands, you can change the route selection process. For example, you can run the **bgp bestpath as-path ignore** command to make step 4 of the route selection process lose effect or run the **bgp bestpath compare-routerid** command to make step 9 lose effect.

Protocols and Standards

- RFC4271: A Border Gateway Protocol 4 (BGP-4)
- RFC4273: Definitions of Managed Objects for BGP-4
- RFC4360: Proposed Standard: BGP Extended Communities Attribute
- RFC4364: Proposed Standard: BGP/MPLS IP Virtual Private Networks (VPNs)
- RFC4486: Proposed Standard: Subcodes for BGP Cease Notification Message
- RFC4724: Proposed Standard: Graceful Restart Mechanism for BGP
- RFC4760: Draft Standard: Multiprotocol Extensions for BGP-4
- RFC5492: Draft Standard: Capabilities Advertisement with BGP-4
- RFC7313： Enhanced Route Refresh Capability for BGP-4
- RFC7432: Proposed Standard: BGP MPLS-based Ethernet VPN

5.2 Applications

Application	Description
Inter-AS Route Advertisement	Implement inter-AS route advertisement by using BGP.
Intra-AS Route Reflection	Set up a route reflection topology within an AS to reduce BGP connections.

5.2.1 Inter-AS Route Advertisement

Scenario

BGP implements route advertisement and maintenance across different ASs.

As shown in Figure 5- 1, BGP transfers the route of AS 65536 to AS 65538 through AS 65537.

Figure 5- 1

Remarks	R1 is a device at the network edge of AS 65536. R2 and R3 are devices at the network edge of AS 65537. R4 is a device at the network edge of AS 65538.

Deployment

- Establish the EBGP neighborship between R1 and R2 to implement inter-AS route advertisement.
- Establish the IBGP neighborship between R2 and R3 to implement intra-AS route advertisement.
- The Internet runs OSPF to ensure network accessibility between R2 and R3.
- Establish the EBGP neighborship between R3 and R4 to implement inter-AS route advertisement.

5.2.2 Intra-AS Route Reflection

Scenario

According to the BGP route advertisement principles, routes learned by an IBGP neighbor will not be advertised to the next IBGP neighbor by default. Therefore within an AS, a device running BGP must implement full-mesh. When there are many BGP devices within the AS, implementing full-mesh may cause large difficulties for network deployment. In this case, route reflection can be used to solve this problem.

As shown in Figure 5- 2, route reflection is deployed to implement BGP full-mesh among R1 to R4 and RR.

Figure 5- 2

Remarks	RR is a route reflector. R1 to R4 are route reflection clients.

Deployment

- Establish IBGP neighborships between R1 to R4 and RR respectively.
- Configure R1 to R4 as the route reflection clients of RR.

5.3 Features

Basic Concept

BGP Speaker and AS Number

A router enabled with BGP is called a BGP speaker.

After a router is enabled with BGP, a local AS number must be specified for the router. An AS number is a globally unique number allocated by IANA, ranging from 1 to 4294967295.

BGP Neighbor and Peer

Before a route is advertised between BGP speakers, a neighborship must be established in advance. You need to manually configure BGP neighbors on both BGP speakers. That is, configure the peer as a neighbor on the two BGP speakers respectively. Therefore, BGP neighbors are also called BGP peers.

Neighbor Type and Route Type

BGP neighborships are classified into the following types:

- IBGP neighborship: The neighborship between BGP speakers within an AS is called IBGP neighborship. Routes learned from IBGP neighbors are called IBGP routes.
- EBGP neighborship: The neighborship between BGP speakers in different ASs is called EBGP neighborship. Routes learned from EBGP neighbors are called EBGP routes.

BGP route attribute

When a BGP speaker advertises routes to its neighbors, the BGP speaker also advertises the attributes carried by the routes. Common BGP attributes are as follows:

- ORIGIN: Specifies the origin of a BGP route and can be set to **IGP**, **EGP**, or **INCOMPLETE**.
- AS-PATH: Lists the ASs passed by a route in a reverse order. The last AS is placed at the beginning of the list.
- NEXT-HOP: Specifies the IP address of the next hop to be reached by a BGP route.
- MULTI-EXIT-DISC: Distinguishes multiple output/input interfaces for reaching the same neighbor AS. A smaller value means a higher priority.
- LOCAL-PREF: Distinguishes the priorities of IBGP routes in an AS. A larger value means a higher priority.

Overview

Feature	Description
Creating a BGP Neighbor	Create a BGP neighbor.
Configuring a BGP Route Reflector	Set up a BGP route reflection topology to simplify network deployment for BGP neighbor full-mesh.
Configuring a BGP Alliance	Configure a BGP alliance to simplify network deployment for BGP neighbor full-mesh.

Feature	Description
Re-distributing Local AS Network Information to BGP	Re-distribute routing information to BGP and advertise local routes through BGP.
Controlling Route Exchange Between BGP Peers	Configure the route exchange policy for a BGP peer and control routes to be received by and to be advertised to this peer.
Obtaining Accessible Networks of Other ASs from BGP	Re-distribute routing information in BGP into a core routing table or IGP.
Configuring Synchronization Between BGP and IGP	Configure BGP to check whether BGP routes are synchronized with IGP routes.
Configuring BGP Soft Reset	After a routing policy changes, use soft reset to apply a new policy.
Configuring the Route Attributes of BGP	Configure the route selection algorithms and routing policy control of BGP.
Configuring BGP Route Aggregation	Reduce routes by means of route aggregation.
Configuring BGP Route Dampening	Reduce the impacts of route flapping on a network topology.
Configuring the Management Distance of BGP	Change the priorities of BGP routes.
Configuring Multi-path Load Balancing of BGP	Configure multi-path load balancing for BGP to enhance the network reliability and increase the network bandwidth.
Configuring BGP FRR	Configure fast rerouting for BGP to enhance the network reliability.
Configuring BGP Timers	Modify the internal timer time of BGP.
Configuring BGP Route Update Mechanisms	Disable/Enable regular scanning for BGP routes and configure the route scanning interval.
Configuring the Next-Hop Triggering Update Function of BGP	Configure the next hop triggering update function of BGP.
Configuring BGP LOCAL AS	Configure the LOCAL AS for a BGP neighbor.
Configuring BGP Capacity Protection	Avoid non-predictable running status caused by consumption of device capacity.
Configuring BGP GR	Configure the BGP GR function to enhance the network reliability.
Configuring 4-Byte AS Numbers of BGP	Configure the display mode of a 4-byte AS number.
Configuring a Regular Expression	Use a regular expression to filter routing information.
Configuring BGP Session Retention	Configure BGP to ensure that after an address family with incorrect routing attributes is detected for a neighbor, other address family routes advertised by the neighbor will not be affected.

Feature	Description
Configuring BGP Delayed Advertisement upon System Restart	Configure BGP to delay route advertisement to a neighbor within a period after the system is restarted.
Configuring BGP Tracking	Configure BGP tracking function.
Configuring Outbound Loop Detection for a BGP Neighbor	Enable outbound loop detection for a BGP neighbor.
Configuring Enhanced VPN Route Import	Configure the enhanced VPN route import function.
Configuring Enhanced Route-Refresh	Indicate the BGP route update group, which is used to improve the handling performance for route advertisement to neighbors.
EVPN Route Attribute	Indicate EVPN route attribute.
Other Related Configurations	Configure extended BGP functions.

5.3.1 Creating a BGP Neighbor

A BGP neighbor is manually configured by a user. Two connection modes are supported: IBGP and EBGP. You can identify the connection mode between BGP speakers based on the AS where the BGP peer resides and the AS where the BGP speaker resides.

Generally, BGP speakers between which an EBGP connection is established are directly connected whereas BGP speakers between which an IBGP connection is established can be at any location within an AS.

Working Principle

A BGP speaker can initiate a TCP connection request to a BGP peer specified by a user. After the TCP connection is successfully created, the peers will exchange BGP packets to negotiate about connection parameters. The BGP neighborship is successfully established after the negotiation succeeds.

Creating a TCP Connection

A BGP speaker initiates a TCP connection request to a neighbor. The destination IP address is the peer IP address specified by the user and the port number is fixed to 179.

The BGP speaker also listens on the port number 179 of the local TCP connection to receive connection requests from its peer.

Negotiating about Protocol Parameters

After the TCP connection is successfully created, the BGP speakers exchange OPEN packets to negotiate about BGP connection parameters. The parameters for negotiation include:

- Version: Indicates the BGP version number. At present, only version 4 is supported.
- Neighbor AS number: Determines whether the AS number of the neighbor is consistent with the local AS number. If not, the connection request will be denied.
- Hold Time: Negotiates about the timeout duration for the BGP connection. The default value is 180 seconds.
- Neighbor capability: Negotiates about various extended capabilities supported by the neighbor, including the address family, dynamic route update, and GR functions.

Maintaining Neighborship

The Keepalive message is periodically sent between BGP speakers. If a new Keepalive packet is not received from the BGP neighbor after the Hold Time expires, the BGP speaker considers that the neighbor is not accessible, disconnects the TCP connection from the neighbor, and attempts to reconnect to it. The interval for a BGP speaker to send the Keepalive message is one third of the Hold Time determined through negotiation and is 60 seconds by default.

Related Configuration

Creating a BGP Neighbor

By default, a BGP speaker does not specify any neighbor. You can manually configure a BGP neighbor.

You can run the **neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number* command to manually create a BGP neighbor and specify the AS number of the neighbor.

Setting the Neighbor TTL

By default, The TTL field in a TCP packet sent by an IBGP neighbor is set to the maximum value (255). It is set to 1 by an EBGP neighbor.

You can run the **neighbor** { *peer-address* | *peer-group-name* } **ebgp-multihop** [*ttl*] command to set the TTL field of a TCP packet sent by a BGP neighbor.

A larger value of TTL means a longer distance between BGP neighbors. When TTL is 1, the BGP neighbor devices must be directly connected.

Setting the Source Address of TCP

By default, BGP automatically selects the source IP address of a TCP connection based on the IP address of the neighbor. Generally, the IP address of a local packet output interface is used.

You can run the **neighbor** { *peer-address* | *peer-group-name* } **update-source** {*interface-type interface-number* | *address* } command to adjust the source IP address of the neighbor's TCP connection.

Setting MD5 Encryption

By default, a BGP connection is not encrypted through MD5.

You can run the **neighbor** { *peer-address* | *peer-group-name* } **password** [**0** | **7**] *string* command to set encryption for a BGP neighbor's TCP connection.

Activating the Address Family Capability of a Neighbor

By default, a neighbor created in the BGP configuration mode activates only the IPv4 Unicast address family capability.

You can run the **address-family** command to enter a corresponding address family mode, and then run the **neighbor** { *peer-address* | *peer-group-name* } **activate** command to activate the address family capability for the BGP neighbor.

5.3.2 Configuring a BGP Route Reflector

According to the principle of BGP route advertisement, full mesh must be established for all BGP speakers within an AS (neighborships need to be established between each two BGP speakers). Too many BGP speakers within an AS will increase the resource overhead of the BGP speakers, increase the network administrator's workload and complexity of configuration and decrease the network expansion capability.

Using a route reflector is a method for reducing IBGP peer connections within an AS.

The methods for reducing the IBGP peer connections within an AS include using a route reflector and using an AS alliance.

Working Principle

Configure a BGP speaker as a route reflector which classifies IBGP peers in an AS into two types: clients and non-clients.

The rules for implementing a route reflector within an AS are as follows:

- Configure a route reflector and specify clients for the route reflector. The route reflector and its clients form a cluster. The route reflector will connect to its clients.
- The clients of a route reflector in a cluster cannot connect to other BGP speakers out of the cluster.
- Within an AS, full mesh is established among IBGP peers of non-clients. The IBGP peers of non-clients include the following situations: Multiple route reflectors in a cluster; a route reflector in a cluster and BGP speakers (generally not supporting the route reflector function) not involved in the route reflector function out of the cluster; a route reflector in a cluster and route reflectors in other clusters.

The rules for processing a route received by a route reflector are as follows:

- A route update message received by an EBGP speaker will be sent to all clients and non-clients.
- A route update message received by a client will be sent to other clients and all non-clients.
- A route update message received by an IBGP speaker will be sent to all the other clients.

Generally, only one route reflector is configured in a cluster. In this case, the Router ID of the route reflector can be used to identify this cluster. To increase the redundancy, you can set multiple route reflectors in a cluster. In this case, you must configure the cluster ID so that a route reflector can identify the route update messages from other route reflectors in the cluster.

If multiple route reflectors are configured for a cluster, you must configure a cluster ID for the cluster.

Generally, it is unnecessary to create connections between the clients of a route reflector in a cluster because the route reflector will reflect the routes between the clients. However, if full mesh has been established among all clients, you can cancel the client route reflection function of the route reflector.

Related Configuration

Configuring a BGP Route Reflector and Reflected Clients

By default, BGP is not configured with route reflection.

You can run the **neighbor** *peer-address* **route-reflector-client** command to configure a device as a route reflector and its neighbor devices as reflected clients.

Configuring BGP Client-Client Reflection

By default, BGP client-client route reflection is enabled, which means that routes received from a reflected client can be advertised to other clients.

You can run the **bgp client-to-client reflection** command to enable or disable (using the **no** form of this command) client-client reflection.

Configuring a BGP Reflection Cluster ID

By default, a BGP reflection cluster ID is the Router-ID of BGP. If multiple reflection clusters are deployed within an AS, different reflection cluster IDs must be configured for these reflection clusters.

You can run the **bgp cluster-id** *cluster-id* command to manually configure the cluster ID of a route reflector.

5.3.3 Configuring a BGP Alliance

An alliance is another method for reducing the IBGP peer connections within an AS.

Working Principle

Divide an AS into multiple sub ASs and configure a unified alliance ID (namely, the alliance AS NUMBER) for these sub ASs to form an alliance. Outside the alliance, the entire alliance is still considered as an AS and only the AS number of the alliance is visible. Inside the alliance, full mesh of IBGP peers can be established for BGP speakers within a sub AS, and EBGP connections can be established for BGP speakers in different sub ASs. Though EBGP connections are established between BGP speakers within a sub AS, when information is exchanged, NEXT_HOP, MED, LOCAL_PREF and other path attributes keep unchanged.

Related Configuration

Configuring a BGP Alliance ID

By default, no alliance ID is configured for a BGP speaker.

You can run the **bgp confederation identifier** *as-number* command to configure a BGP alliance ID. After the configuration is successful, the local AS (specified by the **router bgp** *as-number* command) of BGP becomes the private AS inside the alliance and is invisible to other ASs.

Configuring a BGP Alliance Neighbor

By default, no alliance neighbor is configured for BGP.

You can run the **bgp confederation peers** *as-number* [… *as-number*] command to configure a BGP alliance neighbor. After the configuration succeeds, the AS specified by this command and the local AS belong to the same alliance.

5.3.4 Re-distributing Local AS Network Information to BGP

BGP cannot automatically discover or learn accessible networks. The accessible network information of a local AS must be re-distributed to BGP. Then, BGP can advertise the information to neighbors.

Working Principle

Two methods can be used to re-distribute local AS network information to BGP:

- Manual static configuration: re-distribute the accessible network information within a specified range to BGP.
- Configuring route re-distribution: re-distribute accessible IGP network information to BGP.

In addition, you can also re-distribute local AS network information to BGP routes by configuring route aggregation.

Related Configuration

Configuring a BGP Network

By default, no network is configured for BGP.

You can run the **network** *network-number* [**mask** *mask*] [**route-map** *map-tag*] [**backdoor**] command to configure a BGP network to re-distribute specified accessible network information to BGP. The prerequisite for successfully re-distributing routing information to BGP is that a route is available in the core routing table and this route can be an IGP, directly-connected or static route.

Configuring BGP Route Re-distribution

By default, BGP is not configured with route re-distribution.

You can run the **redistribute** *protocol-type* command to re-distribute the routing information of other protocols to BGP, including OSPF, RIP, ISIS, static and directly-connected routes.

Importing Routes with Multiple Paths or Next Hops to BGP

By default, routes imported to BGP have only one next hop.

Run the **bgp sourced-paths *protocol-type* all** command to import routes with multiple next hops of other protocols to BGP.

5.3.5 Controlling Route Exchange Between BGP Peers

BGP provides powerful route management functions. You can actively control the route exchange between BGP peers.

Working Principle

Configure the route exchange policy for a BGP peer and control routes to be received by and to be advertised to this peer.

Related Configuration

Configuring the Default Route to Be Advertised to a Peer

By default, BGP does not advertise the default route.

You can run the **neighbor** { *address* | *peer-group-name* } **default-originate** [**route-map** *map-tag*] command to advertise the default route to a peer (or a peer group).

Configuring Next-Hop-Self for a Peer

By default, BGP does not change the next hop of a route when it advertises the route to an IBGP neighbor and sets the next hop to the local BGP speaker when it advertises the route to an EBGP neighbor.

You can run the **neighbor** { *address* | *peer-group-name* } **next-hop-self** command to configure the next hop of a route to the local BGP speaker when distributing the route to a specified BGP peer (group).

Configuring Remove-Private-AS for a Peer

By default, BGP does not delete the private AS in the AS-PATH attribute when it advertises routing information to a peer.

You can run the **neighbor** { *address* | *peer-group-name* } **remove-private-as** command to require that the private AS number recorded in the AS path attribute should be deleted when routing information is distributed to an EBGP peer (group). This command does not apply to an IBGP neighbor.

Configuring Send-Community for a Peer

By default, BGP does not send the community attribute when it advertises routing information to a peer.

You can run the **neighbor** { *address* | *peer-group-name* } **send-community** command to specify that the community attribute can be sent to a specified BGP peer (group).

Configuring Maximum-Prefix for a Peer

By default, BGP does not restrict the records of routing information that can be received by a peer.

You can run the **neighbor** { *address* | *peer-group-name* } **maximum-prefix** *maximum* [**warning-only**] command to specify the records of routing information received from a specified peer (group).

Configuring Route Filtering for a BGP Neighbor

By default, a BGP neighbor is not enabled with any filtering policy and receives all legal routing information advertised by a neighbor.

BGP supports multiple methods of configuring the route filtering policies for a neighbor, including:

- **neighbor** { peer-address | peer-group-name } **distribute-list** { access-list-number | access-list-name } { **in** | **out** }

Use an ACL to filter routes in the input and output directions of the neighbor.

- **neighbor** { *peer-address* | *peer-group-name* } **filter-list** *access-list-number* { **in** | **out** }

Use an AS-PATH list to filter routes in the input and output directions of the neighbor.

- **neighbor** { *peer-address* | *peer-group-name* } **prefix-list** *prefix-list-name* { **in** | **out** }

Use a prefix-list to filter routes in the input and output directions of the neighbor.

- **neighbor** { *peer-address* | *peer-group-name* } **route-map** *map-tag* { **in** | **out** }

Use a route map to filter routes in the input and output directions of the neighbor.

- **neighbor** { *address* | *peer-group-name* } **unsuppress-map** *map-tag*

Allow for advertising certain routing information previously suppressed by the **aggregate-address** command when distributing routing information to a specified peer.

5.3.6 Obtaining Accessible Networks of Other ASs from BGP

Send routing information of other ASs exchanged by BGP to the routing table of a device so that the device can forward packets to other ASs.

Send routing information of other ASs exchanged by BGP to the routing table of a device so that the device can forward packets to other ASs.

Working Principle

BGP Sends Routing Information to a Core Routing Table

BGP controls routing information sent to the core routing table by using **table-map**. **table-map** can modify the attributes of routing information sent to the core routing table. If the route is matched, BGP modifies the attribute of the routing information and sends the route. If the route is not matched or route matching is denied, BGP does not modify the attribute of the routing information but sends the route.

Changes of **table-map** are not reflected in the core routing table immediately, but reflected a moment later. To update the application of **table-map** immediately, you can run the **clear ip bgp** [vrf *vrf-name*] **table-map** command to update the routing information in the core routing table immediately. This command does not clear the existing routes in the core routing table, but directly applies **table-map** to send the updated routing information, thereby not causing forwarding flapping.

Re-distributing BGP Routes to IGP

Re-distribute BGP routes on a BGP speaker to IGP to ensure that routers within an AS can obtain routes to other ASs.

Related Configuration

Configuring table-map

By default, BGP is not configured with a table-map and allows for sending all routes without modifying the attributes of the routes.

You can run the **table-map** *route-map-name* command to set a table-map and control the routing information to be sent to the core routing table. *route-map-name* specifies a route-map to be associated.

Run the **table-map** command in the BGP configuration mode or in the IPv4 address family mode.

The Match and Set statements supported in the table-map are as follows:

Match statements: as-path, community, ip address, ip next-hop, metric, origin and route-type

Set statements: metric, tag and next-hop

You can run the **no table-map** command to delete the table-map configurations.

Configuring BGP Route Re-distribution by IGP

By default, IGP does not re-distribute BGP routes.

You can run the **redistribute bgp** [**route-map** *map-tag*] [**metric** *metric-value*] command to re-distribute BGP routes to IGP (RIP\OSPF\ISIS).

The **bgp redistribute-internal** command controls only whether to re-distribute routes learned from IBGP to IGP. By default, routes learned from IBGP can be re-distributed to IGP.

You can run the **bgp redistribute-internal** command in the BGP configuration mode, IPv4/IPv6 address family mode or the IPv4 VRF address family mode.

You can run the **no bgp redistribute-internal** command to delete the configuration.

5.3.7 Configuring Synchronization Between BGP and IGP

Generally, BGP speakers working as mutual IBGP neighbors are not directly connected. IGP devices between the BGP speakers may fail to learn routing information same as that learned by the BGP speakers. When a BGP speaker at the border of an AS forwards packets received from other domains to the next-hop IBGP neighbor, the packets pass an IGP device in the middle. In this case, the packets may be lost due to no routing information on the IGP device.

Working Principle

To keep synchronization between BGP and IGP, you must ensure that all routers within an AS can learn routing information to be sent to another AS before the routing information is advertised to this AS.

Synchronization between BGP and IGP is not required only in the following cases:

- Routing information passing through an AS is not available. For example, the AS is an end AS.
- All routers within an AS run BPG. Full mesh is established among all BGP speakers (neighborship is established between each two BGP speakers).

Related Configuration

Configuring BGP Route Synchronization

By default, synchronization between BGP and IBGP routes is disabled.

You can run the **synchronization** command to enable synchronization between BGP and IGP.

Note

You can run the **no synchronization** command to disable synchronization between BGP and IGP.

5.3.8 Configuring BGP Soft Reset

If routing policies (including **neighbor distribute-list, neighbor route-map, neighbor prefix-list** and **neighbor filter-list**) change, an effective method must be provided to implement new routing policies. A traditional method is to terminate a BGP connection and then create a new BGP connection. By configuring BGP Soft Reset, you can execute a new routing policy without terminating a BGP session connection.

Working Principle

Routing policies that affect inbound routing information are called inbound routing policies (such as In-route-map and In-dist-list) and routing policies that affect outbound routing information are called outbound routing policies (such as Out-route-map and Out-dist-list).

When outbound routing policies change, BGP soft reset will re-advertise all routing information of a BGP speaker to its neighbors.

If inbound routing policies change, the operation is more complex than that when outbound routing policies change. This is because outbound routing policies are executed in the routing table of the local BGP speaker whereas inbound routing policies are executed for routing information received from the BGP peer. To reduce cost, the local BGP speaker does not store the original routing information received from the BGP peer.

If inbound routing policies change and a neighbor device supports route update, you can configure soft reset to send a route update request to the neighbor device. After receiving the request, the neighbor device re-advertises all routing information. You can also perform configuration to ensure that each BGP peer stores original routing information on the local BGP speaker and provides original routing information basis for modifying inbound routing policies subsequently.

The "route update capability" allows for modifying and executing routing policies without storing original routing information. This product supports the route update capability. You can run the **show ip bgp neighbors** command to check whether a BGP peer supports route update. If yes, you do not need to run the **neighbor soft-reconfiguration inbound** command when inbound routing policies change.

Related Configuration

Configuring BGP Soft Reset

Run the **clear ip bgp** { * | *peer-address* | **peer-group** *peer-group-name* | **external** } **soft out** command to soft reset a BGP connection. You can activate execution of a routing policy without restarting the BGP session.

Saving Original Routing Information of Neighbors

By default, BGP does not save original routing information of neighbors.

Run the **neighbor** { *address* | *peer-group-name* } **soft-reconfiguration inbound** command to save unmodified routing information sent by a BGP peer (group).

5.3.9 Configuring the Route Attributes of BGP

BGP provides various control policies for route attributes. You can apply the policies based on actual conditions.

Working Principle

AS_PATH Attribute

BGP can control distribution of routing information in three modes:

- IP address. You can run the **neighbor distribute-list** and **neighbor prefix-list** commands for implementation.
- AS_PATH attribute. See the description in this section.
- COMMUNITY attribute. See the related configuration of the COMMUNITY attribute.

You can use an AS path-based access control list (ACL) to control the distribution of routing information. Where, the AS path-based ACL uses a regular expression to parse the AS path.

Based on the standard (RFC1771), BGP does not consider the AS path length when selecting the optimum path. Generally, a shorter AS path length means a higher path priority; therefore, FS considers the AS path length when selecting the optimum path. You can determine whether to consider the AS path length when selecting the optimum path based on the actual conditions.

Within an AS, whether to consider the AS path should be consistent for all BGP speakers when the optimum path is selected; otherwise, the optimum paths selected by the BGP speakers may be different.

MULTI_EXIT_DISC Attribute

BGP uses the MED value as the basis for comparing priorities of paths learned from EBGP peers. A smaller MED value means a higher path priority.

- By default, the MED value is compared only for paths of peers from the same AS when the optimum path is selected.
- By default, the MED value is not compared for paths of peers from other sub ASs within an AS alliance.
- By default, if a path not configured with the MED attribute is received, it is considered that the MED value of this path is 0. Since a smaller MED value means a higher path priority, this path has the highest priority.
- By default, the MED value is not compared with paths from different ASs; instead, the sequence of receiving the paths is compared.

LOCAL_PREF Attribute

When sending routes received from EBGP peers to IBGP peers, a BGP speaker adds the LOCAL_PREF attribute. BGP uses the LOCAL_PREF attribute as the basis for comparing priorities of paths learned from IBGP peers. A larger value of LOCAL_PREF means a higher path priority.

You can also run the **set local-preference** command of a route map to modify the LOCAL_PREF attribute of the specified path.

COMMUNITY Attribute

The COMMUNITY attribute is another mode for controlling distribution of routing information.

A community is a set of destination addresses. The COMMUNITY attribute is intended to facilitate execution of a routing policy based on a community and thereby simplify the configuration of routing information distribution control on BGP speakers. Each destination address may belong to multiple communities. An AS administrator can define the communities, to which a destination address belongs.

By default, all destination addresses belong to the Internet community and are carried in the community attribute of the path.

At present, four common community attribute values are pre-defined:

- Internet: Indicates the Internet community. All paths belong to this community.
- no-export: Indicates that the path is not advertised to EBGP peers.

- no-advertise: Indicates that the path is not advertised to any BGP peer.
- local-as: Indicates that a path is not advertised to other ASs. When an AS alliance is configured, the path is not advertised to other ASs or sub ASs.

By using the community attribute, you can control the receiving, prioritization and distribution of routing information. BGP speakers can set, add or modify the community attribute when learning, advertising or re-distributing routes. An aggregation path will contain the community attribute values of all aggregated paths.

BGP supports up to 32 COMMUNITY attributes for each route and allows for up to 32 COMMUNITY attributes when match and set COMMUNITY of a route map are configured.

Others

During selection of the optimum path, if two paths with the same path attributes are received from different EBGP peers, the optimum path is selected based on the receiving sequence by default. You can disable comparison of the receiving sequence but use the path with a smaller router ID as the optimum path.

Related Configuration

Configuring AS_PATH Attribute

- **ip as-path access-list** *path-list-name* { **permit** | **deny** } *as-regular-expression*

Defines an AS path list.

- **neighbor** { *address* | *peer-group-name* } filter-list *path-list-name* { **in** | **out** }

By default, no filtering policy is configured for BGP peers.

The configuration is the same as that for routing information receiving and sending for a specified BGP peer (group). Routing policies are executed based on the AS path list to advertise or receive only routes that match the policies.

- **neighbor** { *address* | *peer-group-name* } **route-map** *map-tag* { **in** | **out** }

By default, no filtering policy is configured for BGP peers.

The configuration is the same as when receiving and sending routing information for a specified BGP peer (group). Routing policies are executed based on a route map or the set rules in the route map are used to modify routing attributes.

In the route-map configuration mode, you can run the **match as-path** command to modify AS path attributes by using an AS path list or directly run the **set as-path** command to modify AS attribute values.

- **bgp bestpath as-path ignore**

Allows BGP not to consider the AS path length when selecting the optimum path. The AS path length is compared by default.

By default, a smaller AS path length means a higher path priority.

Configuring MULTI_EXIT_DISC Attribute

- **bgp always-compare-med**

Allows for comparing the MED values of paths from different ASs, which is disabled by default.

- **bgp bestpath med confed**

Allows for comparing the MED values of paths of peers from other sub ASs in the same AS alliance, which is disabled by default.

- **bgp bestpath med missing-as-worst**

Sets a path not configured with the MED attribute to the lowest priority, which is disabled by default.

- **bgp deterministic-med**

Allows for comparing the paths of peers within the same AS, which is disabled by default.

Configuring LOCAL_PREF Attribute

- **bgp default local-preference** *value*

Changes the default local preference value, ranging from 0 to 4,294,967,295. A larger value means a higher priority. The default value is 100.

Configuring COMMUNITY Attribute

- **ip community-list standard** *community-list-name* { **permit** | **deny** } *community-number*

Creates a community list. **community-list-name** indicates the name of the community list.

community-number: Indicates a value (0 to 4,294,967,295) specified by a user or a known community attribute (internet, local-AS, no-advertise or no-export).

- **neighbor** { *address* | *peer-group-name* } **send-community**

Allows for sending the community attribute to a specified BGP peer (group), which is not configured by default.

- **neighbor** { *address* | *peer-group-name* } **route-map** *map-tag* { **in** | **out** }

The configuration is the same as that for routing information receiving and sending for a specified BGP peer (group). Routing policies are executed based on a route map. No filtering policy is configured for peers by default.

In the route-map configuration mode, you can run the **match community-list [exact]** and **set community-list delete** commands to modify the community attribute by using a community list or directly run the **set community** command to modify the community value.

Others

- **bgp bestpath compare-routerid**

Allows BGP to compare the router ID when selecting the optimum path, which is disabled by default.

5.3.10 Configuring BGP Route Aggregation

BGP-4 supports CIDR and therefore allows for creating aggregation entries to reduce the size of a BGP routing table. BGP aggregation entries can be added to a BGP routing table only when valid paths are available within the aggregation range.

Working Principle

Aggregate one or more detailed BGP routes into a BGP route with a shorter network mask.

By default, BGP advertises all path information before and after aggregation. If you hope that only aggregated path information is advertised, you can run the **aggregate-address summary-only** command.

When the **aggregate-address** command is used to configure an aggregated route, the aggregated route takes effect immediately as long as there are routes in the configured address range.

Related Configuration

Configuring BGP Route Aggregation

- **aggregate-address** *address mask*

Configures BGP route aggregation. By default, BGP does not create any aggregated routing entry.

- **aggregate-address** *address mask* **as-set**

Configures an aggregation address and stores the AS path information within the aggregation address range. By default, BGP does not store AS path information.

- **aggregate-address** *address mask* **summary-only**

Configures an aggregation address and advertises only an aggregated path. By default, BGP advertises all path information within the aggregation range.

- **aggregate-address** *address mask* **as-set summary-only**

Configures an aggregation address, stores the AS path information within the aggregation address range and advertises only aggregated paths.

5.3.11 Configuring BGP Route Dampening

If a route changes between being valid and invalid, route flapping occurs.

Route flapping often causes transmission of unstable routes in a network, and thereby causes network instability. BGP route dampening is a method for reducing route flapping. It reduces possible route flapping by monitoring routing information from EBGP peers.

Working Principle

Terms used in BGP route dampening are as follows:

- Route Flap: A route changes between being valid and invalid.
- Penalty: Once route flapping occurs, a BGP speaker enabled with route dampening adds a value to the penalty for this route. The penalty is accumulated until the Suppress Limit is reached.
- Suppress Limit: When the penalty of a route is greater than this value, the route will be suppressed.
- Half-life-time: The time used for the penalty to be halved.
- Reuse Limit: When the penalty value of a route is smaller than this value, route suppression will be canceled.
- Max-suppress-time: The longest time that a route can be suppressed.

A brief description of route dampening processing: BGP speaker punishes a route once (adds to the penalty) route flapping occurs. When the penalty reaches the Suppress Limit, the route will be suppressed. When the Half-life-time reaches, the penalty is halved. When the penalty is reduced to the Reuse Limit, the route is activated again. The Max-suppress-time indicates the longest time that the route can be suppressed.

Related Configuration

Configuring BGP Route Dampening

- **bgp dampening**

Enables BGP dampening, which is disabled by default.

- **bgp dampening** *half-life-time reuse suppress max-suppress-time*

Configures the parameters of route dampening.

half-life-time (1~45minutes): The default value is 15 minutes. A larger value means a longer flapping suppression and dampening period.

reuse (1~10000): The default value is 750. A smaller value means longer time for continuous stabilization before a flapping route is enabled again.

suppress (1~20000): The default value is 2,000. A smaller value means more flapping times allowed before suppression.

max-supress-time (1~255minutes): The default value is 4*half-life-time. A larger value means longer maximum suppression time.

Displaying BGP Route Dampening

- **show ip bgp dampening flap-statistics**

Displays the flapping statistics about all routes.

- **show ip bgp dampening dampened-paths**

Displays the statistics about suppressed routes.

Resetting BGP Route Dampening

- **clear ip bgp flap-statistics**

Clears the flapping statistics about all routes that are not suppressed.

- **clear ip bgp flap-statistics** *address mask*

Clears the flapping statistics about specified routes (excluding suppressed routes).

- **clear ip bgp dampening** [*address* [*mask*]]

Clears the flapping statistics about all routes, including routes whose suppression is cancelled.

5.3.12 Configuring the Management Distance of BGP

The management distance is used to evaluate the reliability of various route sources. A smaller management distance means a better route.

Working Principle

Management Distance of BGP

The management distance indicates the reliability of a route source, ranging from 1 to 255. A larger value means lower reliability. BGP sets different management distances for routing information learned from different sources, including External-distance, Internal-distance and Local-distance.

- External-distance: Indicates the management distance of routes learned from EBGP peers.
- Internal-distance: Indicates the management distance of routes learned from IBGP peers.
- Local-distance: Indicates the management distance for routes learned from peers but it is considered that better routes can be learned from IGP. Generally, these routes can be indicated by the **Network Backdoor** command.

You are not advised to change the management distance of BGP. If you really need to change the management distance of BGP, please remember:

The external-distance should be shorter than the management distances of other IGP routing protocols (OSPF and RIP).
The internal-distance and local-distance should be longer than the management distances of other IGP routing protocols.

Backdoor Route

If you prefer an IGP route but do not use an EBGP route, you can set the EBGP route as the backdoor route. By default, the management distance for routes learned from a BGP speaker for which an EBGP connection is established is 20. You can run the **network backdoor** command to set the management distance of the network information to 200 so that the same network information learned from IGP has the highest priority. The networks learned from IGP are considered backdoor networks and are not advertised.

Related Configuration

Configuring the Management Distance of BGP

You can run the **distance bgp** *external-distance internal-distance local-distance* command to configure the management distance of BGP. The value ranges from 1 to 255.

The default value of *external-distance* is 20; the default value of *internal-distance* is 200; the default value of *local-distance* is 200.
A longer management distance means a lower route priority.

Configuring a Backdoor Route

Run the **network** *network-number* **mask** *network-mask* **backdoor** command to configure a backdoor route. By default, no backdoor route is configured.

5.3.13 Configuring Multi-path Load Balancing of BGP

Multi-path load balancing means that there are multiple paths to the same network and data packets are evenly forwarded by these paths. In a routing table, one route has multiple next hops.

According to the types of equivalent routes, multi-path load balancing of BGP is classified into the following types:

- EBGP load balancing: implement load balancing for routes learned from EBGP neighbors.
- IBGP load balancing: implement load balancing for routes learned from IBGP neighbors.

Both the IPv4 and IPv6 protocol stacks support multi-path load balancing.

Load balancing cannot be implemented between IBGP and EBGP routes (including EBGP routes in an alliance).

Working Principle

If a BGP routing table has multiple paths to the same network, BGP calculates the route with the highest priority by default. If there are optimum multiple routes with the same priorities, BGP still selects a unique route by using comparison rules, notifies the route to the forwarding plane and controls the forwarding of data streams. After multi-path load balancing is enabled, BGP calculates a unique optimum route and also lists paths with the same priorities as equivalent routes. Then, BGP notifies the optimum route and the equivalent routes to the forwarding plane to implement load balancing.

Equivalent routes have the same basic attributes and priorities. That is, according to the optimum path selection rules of BGP, the paths have the same priorities before router-IDs are compared.

AS_PATH Loose Comparison

By default, equivalent routes must have the same AS-PATH attributes. Under such strict conditions, load balancing cannot be implemented in certain environments. In this case, you are advertised to enable the AS-PATH loose comparison mode. In the AS-PATH

loose comparison mode, when other conditions for equivalent routes are met, as long as the AS-PATH lengths of routes and the AS-PATH lengths of routes from an alliance are the same respectively, it is considered that the conditions for equivalent routes are met.

Router ID Multi-path Comparison

By default, equivalent routes do need to come from the same device (Router ID of the source route do not need to be the same). Enable this function so that only the routes from the same Router ID can be equivalent.

When the next hops of multiple BGP equivalent paths recur to the same IGP output interface, load balancing cannot be implemented.

Related Configuration

Configuring Multi-path Load Balancing of BGP

- **maximum-paths ebgp** *number*

Enables the multi-path load balancing function of EBGP.

number indicates the number of equivalent next hops, ranging from 1 to device capacity. The default value is 1. A larger value means more equivalent next hops allowed.

- **maximum-paths ibgp** *number*

Enables the multi-path load balancing function of IBGP.

number indicates the number of equivalent next hops, ranging from 1 to device capacity. The default value is 1. A larger value means more equivalent next hops allowed.

Configuring AS_PATH Loose Comparison

- **bgp bestpath as-path multipath-relax**

Enables the BGP AS-PATH loose comparison mode.

Configuring Router ID Multi-path Comparison

- **bgp bestpath multipath-compare-routerid**

Enables the router ID multi-path comparison mode.

5.3.14 Configuring BGP FRR

With high-speed development of IP technologies and application of various complex services, the requirements for network security and stability become increasingly higher. Especially, certain real-time services (audios and videos) are sensitive to network running status and may be largely affected by unstable networks. Therefore, more and more focus and importance are attached to network reliability. With these requirements, the IP FRR function comes into being. It is intended to use a backup link to maintain data forwarding during route platform convergence after a faulty link is detected, in order to achieve the ideal targets of "zero delay" and "zero loss" in packet forwarding.

BGP FRR is shorted for Fast Reroute.

Working Principle

If a BGP routing table has multiple paths to the same network, BGP calculates the route with the highest priority by default. After the BGP FRR function is used, BGP selects a backup route for each optimum route. After BFD FRR detects that the master link is faulty, it switches the data to the originally calculated backup link for forwarding. After route convergence is completed, data is switched to the optimum route re-calculated for forwarding. In this way, BGP FRR can avoid route disconnection due to a link fault before BGP route convergence is completed.

- BGP FRR is supported only in the IPv4 Unicast and IPv4 VRF address families of BGP.
- Only one backup route can be generated and the next hop of the backup route cannot be the same as that of the preferred route.
- A backup next hop cannot be generated for an Equal-Cost Multi-Path Routing (ECMP) route.
- In the BGP IPv4 VRF configuration mode, BGP FRR has a lower priority than VPN FRR. That is, if VPN FRR is enabled in the VRF mode, BGP FRR takes effect only when VPN FRR fails to calculate a backup route.

Related Configuration

Configuring BGP FRR

Run the **bgp fast-reroute** command to enable the BGP FRR function, which is disabled by default.

Configuring a BFD Session to a BGP Neighbor

Run the **neighbor** *peer-address* **fall-over bfd** command to configure a BFD session to a BGP neighbor, which is not configured by default.

5.3.15 Configuring BGP Timers

You can manually configure various timers within BGP to meet the neighbor keepalive and route management requirements in different network environments.

Working Principle

BGP Neighbor Keepalive Timer

BGP uses the Keepalive timer to maintain a valid connection with a peer and uses the Holdtime timer to identify whether a peer is valid. By default, the value of the Keepalive timer is 60 seconds and the value of the Holdtime timer is 180 seconds. When a BGP connection is established between two BGP speakers, the two BGP speakers negotiate about the Holdtime timer value and select a smaller value. 1/3 of the negotiated Holdtime timer value and the configured Keepalive timer value are compared and the smaller value is used as the Keepalive timer value.

Neighbor Reconnection Timer

To reduce the impacts of frequent BGP reconnection to a neighbor on the network bandwidth, after a BGP speaker detects failure of a neighbor connection, the BGP speaker attempts to reconnect the neighbor after the connect-retry timer expires. By default, the value of the connect-retry timer is 15s.

Route Advertisement Timer

To reduce the impacts of route update packets on the network bandwidth, after a BGP speaker detects a network topology change, the BGP speaker does not advertise the route update to its neighbors immediately. Instead, the BGP speaker uses a regular update mechanism to advertise all changed routing information to its neighbors.

Related Configuration

Configuring the BGP Neighbor Keepalive Timer

- **timers bgp** *keepalive holdtime*

Adjusts the BGP *keepalive* and *holdtime* values for all peers.

The *keepalive* value ranges from 0 to 65,535. The default value is 60 seconds.

The *holdtime* value ranges from 0 to 65,535. The default value is 180 seconds.

- **neighbor** { *address* | *peer-group-name* } **timers** *keepalive holdtime*

Configures the *keepalive* and *holdtime* values used for connecting to a specified BGP peer (group).

The *keepalive* value ranges from 0 to 65,535. The default value is 60 seconds.

The *holdtime* value ranges from 0 to 65,535. The default value is 180 seconds.

Configuring the Neighbor Re-connection Timer

- **neighbor** { *address* | *peer-group-name* } **timers connect** *connect-retry*

Configures the *connect-retry* value used for reconnecting to a specified BGP peer (group).

The value of *connect-retry* ranges from 1 to 65,535. The default value is 15 seconds.

Configuring the Route Advertisement Timer

- **neighbor** { *address* | *peer-group-name* } **advertisemet-interval** *seconds*

Configures the minimum interval for sending route updates to a specified BGP peer (group). The value of advertisemet-interval ranges from 0 to 600 seconds. The default value for IBGP peers is 0 seconds and the default value for EBGP peers is 30 seconds.

- **neighbor** { *address* | *peer-group-name* } **as-origination-interval** *seconds*

Configures the minimum interval for sending local initial route updates to a specified BGP peer (group). The value of As-origination-interval ranges from 1 to 65,535. The default value is 1 second.

5.3.16 Configuring BGP Route Update Mechanisms

Working Principle

BGP provides two route update mechanisms: regular-scanning update and event-triggering update. Regular-scanning update indicates that BGP uses an internal timer to start scanning regularly and update the routing table. Event-triggering update indicates that BGP starts scanning and updates the routing table when the BGP configuration commands are changed due to user configuration or the next hop of a BGP route changes.

This function is configured based on address families and can be configured in the IPv4, IPv6, IPv4 vrf and IPv6 VRF address family modes.

If you set the BGP route update mechanism to event-triggering update (by running the **bgp scan-rib disable** command), you must disable synchronization (by running the **no synchronization** command) and enable the BGP next-hop triggering update function (by running the **bgp nexthop trigger enable** command). On the other hand, if you enable synchronization or disable the BGP next-hop triggering update function, the BGP routing table must be updated in the regular scanning mode.

Related Configuration

Configuring Route Update Mechanisms

- **bgp scan-rib disable**

Sets the BGP route update mechanism to event-triggering update. Regular-scanning update is used by default.

- **bgp scan-time** *scan-time*

Configures the regular update interval of BGP. The value of *scan-time* ranges from 5 to 60 seconds. The default value is 60 seconds.

5.3.17 Configuring the Next-Hop Triggering Update Function of BGP

The next-hop triggering update function of BGP is a method for reducing the BGP convergence time. This function is used to optimize the method for monitoring the next hop of a route to ensure that BGP can increase the BGP route convergence speed when the network topology is stable.

Working Principle

When BGP connects to a neighbor, BGP automatically monitors the next hop of the BGP route learned from the neighbor. When the next hop changes in the core routing table, BGP receives an advertisement about the next hop change and updates the BGP routing table. This optimization measure improves the BGP route convergence performance by reducing the time for detecting next-hop changes.

If this function is disabled, BGP next hop update will be discovered through regular scanning specified by scan-timer.

This function is configured based on address families and can be configured in the IPv4, IPv6, and IPv4 vrf address family modes.

bgp nexthop trigger delay and **bgp scan-time** control the same timer. When bgp scan is enabled (it is enabled by default and can be disabled by the **bgp scan-rib disable** command), if the value of **bgp nexthop trigger delay** is larger than 60s, bgp scan does not take effect because the scan timer is always triggered before the delay.

If the network environment is unstable (with frequent next-hop changes), especially with many routes, this function performs unnecessary route calculations, which consumes more CPU resources. Therefore, you are advised to disable this function in this environment.

Related Configuration

Configuring the Next-Hop Triggering Update Function of BGP

- **bgp nexthop trigger enable**

Enables the BGP next-hop triggering function, which is enabled by default.

- **bgp nexthop trigger** delay *delay-time*

Configures the delay of BGP next-hop triggering update. The value of *delay-time* ranges from 0 to 100 seconds. The default value is 5 seconds.

5.3.18 Configuring BGP LOCAL AS

The Local AS function of BGP is used to configure a local AS different from a router BGP AS for a specific peer. This is similar to deploying a new virtual AS between the peer devices. When the local router BGP AS changes, you can establish a BGP connection without changing the BGP configurations on the peer device. This function is mainly used for AS migration and merging of large networks and ensures that the device configurations in other interconnected ASs are not affected.

Working Principle

In BGP, when a local device connects to a peer, the local device advertises the local AS number to the peer by using an Open message. The peer checks whether the BGP AS number advertised is the same as the local AS number. If the AS numbers are different, the peer will deny the BGP connection. By default, the local AS in the BGP connection is a route BGP AS. However, if a local AS is configured for the peer, the configured local AS will replace the route BGP AS when a BGP connection is established between the local device and the peer.

The **neighbor** *peer-address* **local-as** *as-num* command for configuring the BGP Local AS function can be followed by more options. For details, see the Command Reference.

The BGP Local AS function is applied only to EBGP peers, but is not applied to IBGP peers and alliance EBGP peers. In addition, the BGP Local AS function has the following restrictions:

1) The configured local AS cannot be the same as the remote AS of a peer.

2) The local AS cannot be configured independently for a member of a peer group.

3) The configured local AS cannot be the same as the route BGP AS.

4) If a device is a member of an AS alliance, the local AS cannot be the same as the AS alliance number.

Related Configuration

Configuring BGP LOCAL AS

- **neighbor** { *address* | *peer-group-name* } **local-as** *as-number*

Configures a local AS for a peer. By default, no local AS is configured for any peer. The local AS of a peer is the route BGP AS.

5.3.19 Configuring BGP Capacity Protection

There are often a large number of BGP routes, which may cause overload of a device, especially for a device with small memory. Protecting BGP capacity helps avoid non-predictable running status caused by consumption of device capacity.

Working Principle

Restricting the Number of BGP Routes

Restrict the number of BGP routes by setting the maximum number of routes in a BGP address family and the maximum number of routes that can be learned by a BGP neighbor.

Entering the OVERFLOW State in case of Insufficient Memory

If the memory is insufficient, BGP can enter the OVERFLOW state. In the OVERFLOW state, BGP generates a default route pointing to a NULL interface. If a newly learned route is not a refined route other than the default route in the current routing table, the route is discarded. In other words, general newly learned routes are discarded to ensure that the system memory is stable. The purpose of not discarding all routes is to avoid route loops in the entire network. Therefore, it is safe for BGP to enter the OVERFLOW state. BGP is allowed to enter the OVERFLOW state by default.

By default, BGP enters the OVERFLOW state in case of insufficient memory. If you do not want to BGP to enter the OVERFLOW state, you can run the **no overflow memory-lack** command to disable this function.

In the OVERFLOW state, BGP supports only the **clear bgp** { *addressfamily* | **all** } * command at present. You can also exit from the OVERFLOW state by disabling and enabling BGP again. When the memory becomes sufficient again, BGP can also automatically exit from the OVERFLOW state.

Related Configuration

Restricting the Number of BGP Routes

- **neighbor** { *address* | *peer-group-name* } **maximum-prefix** *maximum* [**threshold**] [**warning-only**]

Restricts the maximum number of routes that can be learned from a BGP neighbor, which is not restricted by default.

- **maximum-prefix** *maximum*

Restricts the maximum number of routes in a BGP address family. The default maximum number of routes for the BGP IPv4 VRF, IPv6 VRF and IPv4 MDT address families are 10,000 and is not configured for other address families.

- Run the **bgp maximum-prefix** *maximum* [**vrf** *vrf-name*] command to restrict the maximum number of routes in the BGP global or specified VRF. When a route advertisement in an address family causes the current number of BGP routes to exceed the maximum number, a prompt indicating route overflow in the global or specified VRF is displayed, and the BGP global or specified VRF is set to the overflow state. This function is disabled by default.

Configuring BGP OVERFLOW

- **overflow memory-lack**

Enable BGP to enter the OVERFLOW state in case of insufficient memory, which is enabled by default.

5.3.20 Configuring BGP GR

Graceful Restart (GR) is intended to implement uninterrupted data forwarding during restart of BGP. During active/standby switching of the management boards, the GR function keeps the network topology stable, maintains the forwarding table and ensures that key services are not interrupted.

Working Principle

Comply with RFC4724: Graceful Restart Mechanism for BGP. [BGP GR] is used in the following description to indicate the RFC.

BGP GR is not an independent process, but is jointly completed by the Restarter and Helper.

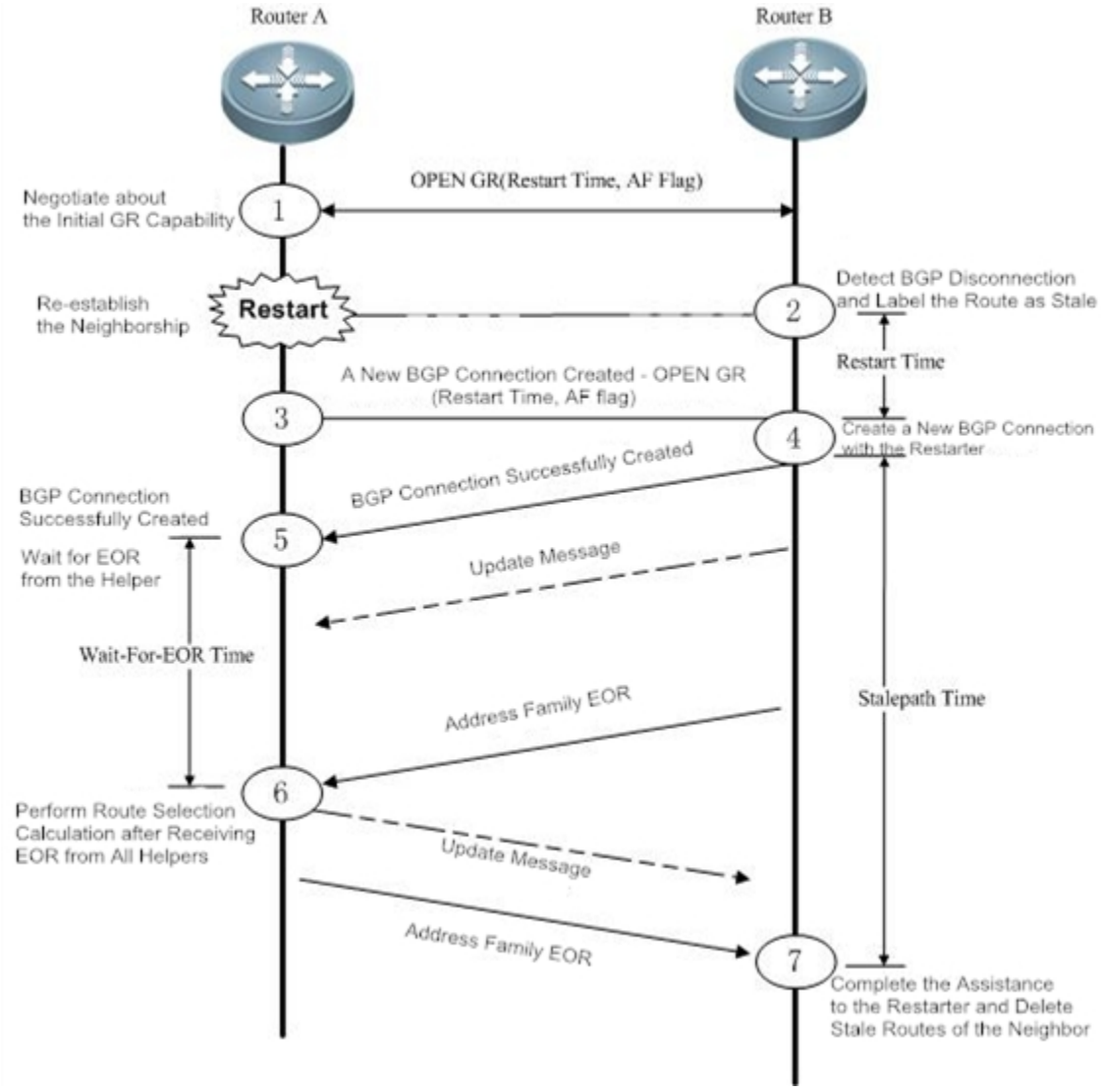
- The Restarter performs restart and maintains the working capability of the route forwarding plane when the route control plane is faulty.

- The Helper is the BGP neighbor of the Restarter and helps the Restarter to complete GR.

A capability indicating GR is added to the OPEN message of BGP, which is called "Graceful Restart Capability". This capability is used by BGP to tell its neighbor it supports the graceful restart capability. During initialization of a BGP connection, two neighbors negotiate about the GR capability.

The route update end flag (End-of-RIB, shorted as EOR) is added to the Update packet of BGP, which indicates that the routing information update to the neighbor is completed.

Figure 5- 3 BGP GR Interaction Process



5. ① When BGP establishes neighborship at the beginning, BGP uses the GR capability field in the OPEN message to negotiate about the GR capabilities of the two neighbors.
6. ②At a moment, the Restarter starts restart, and the BGP session is disconnected. The Helper detects the disconnection, keeps the route of the Restarter valid but adds the "Stale (aged but not updated)" flag to the route.
7. ③ and ④ The Restarter completes restart and connects to the Helper again.
8. ⑤ The Restarter waits for the route update message and EOR flag from the Helper.
9. ⑥ After receiving the EOR flag from all neighbors, the Restarter performs route calculation, update routing entries and then sends updated routes to the Helper.
10. ⑦ After receiving the updated routes, the Helper cancels the "Stale" flag of the routes. After receiving the EOR flag from the Restarter, the Helper deletes routes with the "Stale" flag (these routes are not updated), performs route calculation, and updates the routing entries. The entire GR process is completed.

BGP GR defines several extended and important timers:

- **Restart-Timer**: The GR Restarter advertises the time value to the GR Helper, which indicates the maximum waiting time that the GR Restarter hopes the Helper to wait before a new connection is established between them. You can run the **bgp graceful-restart restart-time** command to modify the time value.

- **Wait-For-EOR Timer**: Indicates the maximum time that the GR Restarter waits for the EOR flag from all GR Helpers. After receiving the EOR flag from all GR Helpers or after the Wait-For-EOR timer expires, the GR Restarter calculates the preferred route and updates the routing entries. You can run the **bgp update-delay** command to modify the time value.

- **StalePath Timer**: Indicates the maximum time that the GR-Helper waits for the EOR flag from the GR Restarter after a new connection is established between them. Within this period, the Helper keeps the original route of the Restarter valid. After receiving the EOR flag or after the StalePath timer expires, the Helper clears the routing entries still with the "Stale" tag. You can run the **bgp graceful-restart stalepath-time** command to modify the time value.

Related Configuration

Configuring BGP GR

- **bgp graceful-restart**

Enables the Restarter capability, which is enabled by default.

- **bgp graceful-restart restart-time** *time*

Sets the Restart Timer. The default value is 120 seconds.

- **bgp update-delay** *delay*

Sets the Wait-For-EOR Timer. The default value is 120 seconds.

- **bgp graceful-restart stalepath-time** *time*

Sets the StalePath Timer. The default value is 360 seconds.

- **bgp graceful-restart disable**

Disables the address family GR capability. The address family GR capability is enabled by default. After the global BGP GR is enabled, the GR capability is automatically enabled for all address families.

When BGP GR is implemented, all BGP peers must enable the BGP GR capability. If certain peers do not support or enable GR, BGP GR may fail to be implemented. GR failure may cause a short route black-hole or route loop, which may affect the network. Therefore, you are advised to verify that all neighbors are enabled with the BGP GR capability. You can run the **show ip bgp neighbors** command to display the capabilities successfully negotiated between BGP peers and verify that the GR capability negotiation is successful. In the BGP route configuration mode, run the **bgp graceful-restart** command to enable the BGP GR capability.

The **bgp graceful-restart** command will not be applied to a successfully established BGP connection immediately. That is, when the BGP connection is in the Established state, the BGP peers will not re-negotiate about the GR capability immediately. To enable the BGP peers of the BGP connection to negotiate about the GR capability immediately, you need to forcibly restart the BGP peers to re-negotiate about the GR capability by running the **clear ip bgp** 192.168.195.64 command (for example).To make GR enabling or disabling take effect immediately, you must restart the neighborship for capability negotiation, which may cause network flapping and affect normal use of users. Therefore, you can explicitly control whether to restart the neighborship.

Supporting BGP GR does not mean that a device can be used as the Restarter to implement BGP GR. Whether to implement BGP GR also depends on the hardware capabilities of the device. FS devices must support the dual-engine hot backup when being used as the GR Restarter.

The restart period configured by the **bgp graceful-restar restart-time** command should not be longer than the Hold Time of the BGP peers; otherwise, the Hold Time will be used as the restart time to be advertised to the BGP peers during GR capability negotiation.

The **bgp graceful-restart disable** command is used to disable the GR capability in an address family in the address family configuration mode, which is not configured by default.

5.3.21 Configuring 4-Byte AS Numbers of BGP

A traditional AS number consists of 2 bytes, ranging from 1 to 65,535. A newly defined AS number consists of 4 bytes, ranging from 1 to 4,294,967,295. Newly defined AS numbers are used to cope with exhaustion of AS number resources.

Working Principle

4-byte AS numbers support two expression modes: the decimal mode and dot mode. The decimal mode is the same as the original expression mode, that is, expressing the 4 bytes of an AS number as decimal digits. The dot mode is expressed as ([higher 2 bytes.]lower 2 bytes). If the higher 2 bytes are 0, they will not be displayed.

For example, an AS number is 65534 in the decimal mode and is 65,534 in the dot mode (the 0 at the beginning is not displayed).

For example, an AS number is 65,536 in the decimal mode, and is 1.0 in the dot mode.

For example, an AS number is 65,538 in the decimal mode, and is 1.2 in the dot mode.

Related protocols are as follows: RFC 4893 and RFC 5396.

Configuring the Display Mode of a 4-Byte AS Number

A 4-byte AS number is displayed in the decimal mode by default. You can manually set the display mode to the dot mode. After the setting, a regular expression will use the dot mode for matching 4-byte AS numbers.

Compatibility with Devices Supporting Only 2-Byte AS Numbers

With introduction of 4-byte AS numbers, BGP connections may be established between old BPG speakers supporting only 2-byte AS numbers and new BGP speakers supporting 4-byte AS numbers. If the AS where a new BGP speaker resides has a 4-byte AS number, when an old BGP speaker creates neighborship with the new BGP speaker, the old BGP speaker uses the reserved AS number 23,456 to replace the 4-byte AS number of the new BGP speaker. In the OPEN packets sent by the new BGP speaker to the old BGP speaker, the 4-byte AS number in the **My Autonomous System** field will be replaced by 23,456. In addition, in UPDATE packets sent to the old BGP speaker, the 4-byte AS number in the AS-PATH and AGGREGATOR attributes will also be replaced by 23,456. In addition, new optional transfer attributes AS4-PATH and AS4- AGGREGATOR will be used to record the real 4-byte AS number so that the real AS-PATH and AGGREGATOR attributes can be restored when the route reaches a next new BGP speaker.

In other cases, the real AS number of the remote end is used to create neighborship.

Related Configuration

Configuring the Display Mode of a 4-Byte AS Number

- **bgp asnotation dot**

Displays a 4-byte AS number in the dot mode. The decimal mode is used by default.

5.3.22 Configuring a Regular Expression

A regular expression is a formula that matches strings based on a template.

The formula is used to assess text data and return True or False to indicate whether the expression can correctly describe the data.

Working Principle

Regular expressions are used in BGP path attributes. The following table describes the usages of special characters in a regular expression.

Character	Symbol	Special Meaning
Period	.	Matches any single character.
Asterisk	*	Matches zero or any sequence in a string.
Plus sign	+	Matches one or any sequence in a string.
Question mark	?	Matches zero or one symbol in a string.
Caret	^	Matches the start of a string.
Dollar sign	$	Matches the end of a string.
Underline	_	Matches the start, end and space of commas, brackets and strings.
Square brackets	[ ]	Matches a single character within a range.

Related Configuration

Using a Regular Expression in a show Command

- **show ip bgp regexp** *regexp*

Displays the BGP routing information in a specified regular expression matched by the AS-PATH attribute.

- **show ip bgp quote-regexp** *regexp*

Displays the BGP routing information in a regular expression within the specified double quotation marks matched by the AS-PATH attribute.

5.3.23 Configuring BGP Session Retention

By default, when an UPDATE packet is received from a neighbor, a BGP session will be disconnected if an error is detected on the multi-protocol routing attribute. This will cause flapping of the routes in all address families of this neighbor. That is, the routing error in an address family will affect the route stability in other address families.

Working Principle

After the BGP session retention function is enabled, if an error occurs in the routing attribute of an address family, only the routing information in this address family related to the neighbor is deleted. In addition, the BGP session and other address families are not affected, which enhances the stability of BGP.

recovery-time is used to configure the time for waiting for automatic route recovery, which requires that a neighbor should support the route-refresh capability. After the recovery-time, BGP sends the route-refresh message of the address family to the neighbor and re-advertises all routing information in the address family to this neighbor.

In the session retention state, you can manually reset the neighbor to exit from the session retention state.

Related Configuration

Configuring BGP Session Retention

- **bgp mp-error-handle session-retain** [**recovery-time** *time*]

Enables the BGP session retention function, which is disabled by default.

recovery-time *time* configures the time for waiting for automatic route recovery, ranging from 10 to 4,294,967,296 seconds. The default value is 120.

5.3.24 Configuring BGP Delayed Advertisement upon System Restart

By default, after the neighborship is established after system restart, a BGP peer can advertise route information to its neighbors. This is normal in most cases. However, in certain cases, for example, there are many neighbors or routes during startup but writing entries into the hardware is slow. In this case, the neighbors have learned the routes and started forwarding traffic, but the hardware has not completed writing of entries at the local end, which causes failure of traffic forwarding.

Working Principle

The BGP delayed advertisement upon system restart ensures that routes are not advertised to neighbors immediately after the neighborship is established upon system restart and that the routes are advertised after a period. This function has no effect on other behaviors such as route receiving performed by the neighbors. If part of the routes is not affected by the delay, configure prefix-list policy to match this part of routes so that route advertisement can be more flexible.

delay-time is used to configure the waiting time before routes are advertised to the neighbors. **startup-time** is used to configure the startup time. Within the startup-time, BGP sends routing information to the neighbors at the interval specified by **delay-time**.

After the startup-time ends, the default route advertisement behavior recovers.

Related Configuration

Configuring BGP Delayed Advertisement upon System Restart

- **bgp initial-advertise-delay** *delay-time* [*startup-time*] [**wait-for-controller**]

Enables BGP delayed advertisement upon system restart, which is disabled by default.

delay-time configures the delay time for advertising routes after the BGP neighborship is established upon system restart, ranging from 1 to 600 seconds. The default value is 1s.

startup-time configures the time range for system restart, ranging from 5 to 58,400 seconds. The delayed route advertisement mechanism is used within this range. The default value is 600s.

- **bgp initial-advertise-delay prefix-list** *prefix-list-name*

By default, the BGP delayed advertisement upon system restart is disabled. If enable it, the route will be immediately sent after the prefix-list policy is matched.

prefix-list-name: indicates the name of prefix-list policy.

5.3.25 Configuring BGP Delayed Advertisement for First Routes

By default, after the neighborship is established, a BGP peer can advertise route information to its neighbors. However, neighbors with the neighborship newly established will send out the route information after a delayed period of time.

Working Principle

After BGP starts, BGP peers negotiate to establish the neighborship before sending route information (update packets). In addition, after **update-delay** is configured on the local end, the local end will sends out the route information after the delay time.

If BGP delayed advertisement upon system restart and BGP delayed advertisement for first routes are enabled at the same time, BGP delayed advertisement upon system restart takes precedence over BGP delayed advertisement for first routes.

BGP GR is not affected by either BGP delayed advertisement upon system restart or BGP delayed advertisement for first routes, that is, the BGP GR route advertisement is not affected by the delay time.

Related Configuration

Creating a BGP Neighbor

By default, no neighbor is specified on a BGP speaker. You need to manually configure a BGP neighbor.

Run the **neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number* command to manually create a BGP neighbor and specify the AS number of the neighbor.

Configuring BGP Delayed Advertisement for First Routes

By default, BGP delayed advertisement for first routes is disabled for neighbors.

Run the **neighbor** { *peer-address* | *peer-group-name* } **update-delay** *delay-time* command to enable BGP delayed advertisement for first routes.

5.3.26 Configuring BGP Tracking

The BGP tracking function provides fast link fault detection for BGP speakers, accelerating route convergence.

Working Principle

When the BGP tracking function is enabled for a BGP speaker, the BGP speaker is associated with the corresponding track session of the track module to monitor status change. In normal cases, the BGP speaker associates with the track module to perceive link changes. When a link is faulty, the track module notifies the BGP speaker rapidly, implementing fast route convergence. BGP tracking configuration is simpler than BFD configuration because only local configuration is required.

For details about the track session configuration and related commands, see *TRACK-RNS-SCG.doc*.

Related Configuration

Configuring Association Between the BGP Neighbor and Track Instance

Run the **neighbor** *neighbor-address* **track** *track-obj-number* command to configure association between a BGP neighbor and a track instance, which is not configured by default.

5.3.27 Configuring Outbound Loop Detection for a BGP Neighbor

By default, BGP conducts loop detection on BGP routes when receiving the BGP routes from a neighbor. When the **AS Path** attribute carried in a BGP route contains the local AS number, BGP filters out the BGP route. The outbound loop detection function of a neighbor is to conduct loop detection on routes in advance when the routes are transmitted to a neighbor, so as to filter out loop routes.

Working Principle

When sending a route to an EBGP neighbor, the device judges whether the **AS Path** attribute carried in the BGP route contains the AS number of the neighbor. If yes, the route is looped and the device does not send the route to the EBGP neighbor.

Related Configuration

Creating a BGP Neighbor

By default, no neighbor is specified for a BGP speaker. You need to manually configure a BGP neighbor.

Run the **neighbor** { *neighbor-address* | *peer-group-name* } **remote-as** *as-number* command to manually create a BGP neighbor and specify an AS number for the neighbor.

Enabling Outbound Loop Detection for a Neighbor

The outbound loop detection is disabled for a neighbor by default.

Run the **neighbor** { *neighbor-address* | *peer-group-name* } **as-loop-check out** command to enable the outbound loop detection for the BGP neighbor.

5.3.28 Configuring Enhanced VPN Route Import

Working Principle

During inter-VRF route import, import of L3VPN remote routes to VRF, or import of EVPN routes to the IP route table, only routes with preferred next hops are imported by default.

The enhanced VPN route import function is an extension of the inter-import of the preceding routes. It enables all routes with next hops or equivalent next hops to be imported.

Related Configuration

Configuring the Route Import Policy

By default, only routes with preferred next hops are imported.

Run the **import path selection { all | bestpath | multipath }** command to import all routes with next hops, routes with preferred next hops, or routes with equivalent next hops.

5.3.29 BGP Route Update Group

The BGP route update group function is used to enhance the performance for advertising routes to neighbors.

Working Principle

The BGP route update group function automatically classifies neighbors with the same outbound policy to the same update group. When routes are sent to neighbors, the update packet is encapsulated based on the update group and sent to all neighbors in the update group. In this case, the update packet is encapsulated for once and sent multiple times, improving the performance of route advertisement to neighbors.

5.3.30 Other Related Configurations

For configuration and application of BGP MCE, see section "VRF Configuration Guide".

For configuration and application of BGP L2VPN, see section "L2VPN Configuration Guide".

For configuration and application of BGP/MPLS VPN, see section "BGP/MPLS VPN Configuration Guide".

For configuration and application of the BGP MDT address family, see section "Multicast VPN (MD Configuration Guide)".

5.4 Configuration

Configuration	Description and Command	
Configuring a BGP Peer (Group)	(Mandatory) It is used to create a BGP neighbor.	
	router bgp	Enables BGP.
	neighbor { *peer-address* | *peer-group-name* } **remote-as** *as-number*	Creates a BGP neighbor.
Configuring MD5 Authentication	(Optional) It is used to perform encrypted authentication for the BGP neighbor.	
	neighbor { *peer-address* | *peer-group-name* } **password** [0 | 7] *string*	Configures the password for encryption.
Configuring a Route Reflector	(Optional) It is used to reduce the number of BGP neighbor connections.	
	neighbor { *peer-address* | *peer-group-name* } **route-reflector-client**	Specifies a peer (group) as a reflector client.
Configuring an AS Alliance	(Optional) It is used to reduce the number of BGP neighbor connections.	
	bgp confederation identifier *as-number*	Configures the BGP alliance ID.
	bgp confederation peers *as-number* [… *as-number*]	Configures a BGP alliance neighbor.
Configuring Multi-path Load Balancing of BGP	(Optional) It is used to implement multi-path load balancing.	
	maximum-paths ibgp *number*	Configures IBGP load balancing.
	maximum-paths ebgp *number*	Configures EBGP load balancing.
	bgp bestpath as-path multipath-relax	Enables the BGP AS-PATH loose comparison mode.
Configuring EBGP FRR	(Optional) It is used to increase the convergence speed when a network fault occurs.	
	bgp fast-reroute	Configures BGP FRR.
	neighbor { *peer-address* | *peer-group-name* } **fall-over bfd**	Configures a BFD session to a BGP neighbor.
Configuring FRR in an IBGP Route Reflection Environment	(Optional) It is used to increase the convergence speed when a network fault occurs.	
	bgp fast-reroute	Configures BGP FRR.
Configuring Local ASs	(Optional) It is used for transitional deployment during network migration.	
	neighbor { *peer-address* | *peer-group-name* } **local-as** *as-number* [**no-prepend** [**replace-as** [**dual-as**]]]	Configures the local AS for a BGP neighbor.
Configuring BGP GR	(Recommended) It is used to improve the network reliability.	
	bgp graceful-restart	Enables the BGP GR capability.
	bgp graceful-restart restart-time *restart-time*	Configures the maximum time for BGP GR.
	bgp graceful-restart stalepath-time *time*	Configures the maximum retention time for BGP stable route.

Configuration	Description and Command	
Configuring a BGP IPv6 Address Family	(Optional) It is used to deploy an IPv6 network by using BGP.	
	address-family ipv6 unicast	Enters the BGP IPv6 unicast configuration mode.
	neighbor { *peer-address* | *peer-group-name* } **activate**	Activates the address family capability of a BGP neighbor in the current configuration mode.
Configuring a BGP MDT Address Family	(Optional) It is used to deploy a multicast VPN network by using BGP.	
	address-family ipv4 mdt	Enters the BGP IPv4 multicast VPN configuration mode.
	neighbor { *neighbor-address* | *peer-group-name* } **activate { ipv4 | ipv6 }**	Activates the address family capability of a BGP neighbor in the current configuration mode.
Configuring BGP EVPN	Optional. It is used to configure the EVPN VXLAN network.	
	address-family l2vpn evpn	**address-family l2vpn evpn**
	vni *vni-id*	Create EVI instance
	rd { **auto** | *rd_value* }	Configure RD
	route-target { **import** | **export** | **both** } { **auto** | *rt_value* }	
	export map *routemap-name*	Configure the extended group attribute policy of the local end to EVPN route
	import map *routemap-name*	Configure the policy of the EVPN route to the local VNI instance
Configuring Interconnection with Devices Supporting Only 2-Byte AS Numbers	Optional. It is used for interconnecting with an old device that supports only 2-byte AS numbers.	
	neighbor { *neighbor-address* | *peer-group-name* } **remote-as** *as-number*	Creates a BGP neighbor.

5.4.1 Configuring a BGP Peer (Group)

Configuration Effect

- Configure BGP and create IBGP and EBGP neighbors.

Notes

- If an IBGP neighbor is not directly connected, you need to configure IGP or a static routing protocol to implement interconnection.
- If an EBGP neighbor is not directly connected, you need to configure the **ebgp-multihop** parameter for the neighbor.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring a Source Interface for a BGP Neighbor

- (Optional) Perform this configuration in the BGP configuration mode. By default, BGP automatically selects a local interface that reaches the destination IP address of a peer as the source interface.

For an IBGP neighbor, you are advised to use a Loopback interface as the source interface.

Verification

- Run the **show** command to display the neighbor status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

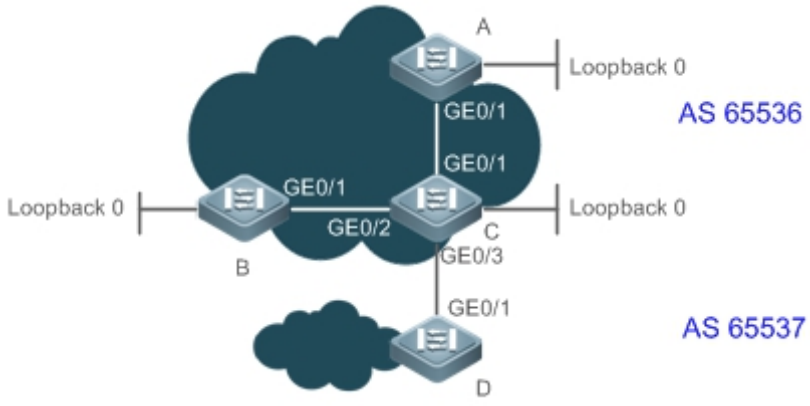
Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Creating a Source Interface for a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **update-source** { *interface-type interface-number* | *address* }
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *interface-type interface-number*: Indicates an interface name. *address*: Directly specifies the network interface address used for creating a BGP connection.
Command Mode	BGP configuration mode
Usage Guide	The source interface of a neighbor must be a local valid interface or address.

Configuration Example

Configuring a BGP Peer (Group)

Scenario **Figure 5- 4**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5- 4. ● Configure a loopback interface on A, B, and C and create an IBGP neighbor based on the loopback interface. ● Create an EBGP neighborship by using the directly connected interfaces on C and D. ● Create an IBGP peer group on C.
A	A# configure terminal A(config)# interface loopback 0 A(config-if-Loopback 0)# ip address 10.1.1.1 255.255.255.255 A(config-if-Loopback 0)# exit A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0 A(config-if-GigabitEthernet 0/1)# exit A(config)# router bgp 65536 A(config-router)# neighbor 10.1.1.3 remote-as 65536 A(config-router)# neighbor 10.1.1.3 update-source loopback 0
B	B# configure terminal B(config)# interface loopback 0 B(config-if-Loopback 0)# ip address 10.1.1.2 255.255.255.255 B(config-if-Loopback 0)# exit B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip address 192.168.2.2 255.255.255.0 B(config-if-GigabitEthernet 0/1)# exit B(config)# router bgp 65536 B(config-router)# neighbor 10.1.1.3 remote-as 65536 B(config-router)# neighbor 10.1.1.3 update-source loopback 0

C	C# configure terminal C(config)# interface loopback 0 C(config-if-Loopback 0)# ip address 10.1.1.3 255.255.255.255 C(config-if-Loopback 0)# exit C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 192.168.1.3 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# interface GigabitEthernet 0/2 C(config-if-GigabitEthernet 0/2)# ip address 192.168.2.3 255.255.255.0 C(config-if-GigabitEthernet 0/2)# exit C(config)# interface GigabitEthernet 0/3 C(config-if-GigabitEthernet 0/3)# ip address 192.168.3.3 255.255.255.0 C(config-if-GigabitEthernet 0/3)# exit C(config)# router bgp 65536 C(config-router)# neighbor ibgp-group peer-group C(config-router)# neighbor ibgp-group remote-as 65536 C(config-router)# neighbor ibgp-group update-source loopback 0 C(config-router)# neighbor 10.1.1.1 peer-group ibgp-group C(config-router)# neighbor 10.1.1.2 peer-group ibgp-group C(config-router)# neighbor 192.168.3.4 remote-as 65537
D	D# configure terminal D(config)# interface GigabitEthernet 0/1 D(config-if-GigabitEthernet 0/1)# ip address 192.168.3.4 255.255.255.0 D(config-if-GigabitEthernet 0/1)# exit D(config)# router bgp 65537 D(config-router)# neighbor 192.168.3.3 remote-as 65536
Verification	Run the **show** command to display the BGP neighbor status.
A	A# show ip bgp neighbor BGP neighbor is 10.1.1.3, remote AS 65536, local AS 65536, internal link BGP version 4, remote router ID 10.1.1.3 BGP state = Established, up for 00:00:05 Last read , hold time is 180, keepalive interval is 60 seconds

	Neighbor capabilities: Route refresh: advertised and received (old and new) Four-octets ASN Capability: advertised and received Address family IPv4 Unicast: advertised and received Received 2 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:1 refresh message:0 dynamic cap:0 notifications:0 Sent 2 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:1 refresh message:0 dynamic cap:0 notifications:0 Route refresh request: received 0, sent 0 Minimum time between advertisement runs is 0 seconds Update source is Loopback 0 For address family: IPv4 Unicast BGP table version 1, neighbor version 1 Index 0, Offset 0, Mask 0x1 0 accepted prefixes 0 announced prefixes Connections established 1; dropped 0 Local host: 10.1.1.1, Local port: 1039 Foreign host: 10.1.1.3, Foreign port: 179 Nexthop: 10.1.1.1 Nexthop global: :: Nexthop local: :: BGP connection: non shared network Last Reset: , due to BGP Notification received Notification Error Message: (Cease/Other Configuration Change.)
B	B# show ip bgp neighbor BGP neighbor is 10.1.1.3, remote AS 65536, local AS 65536, internal link BGP version 4, remote router ID 10.1.1.3 BGP state = Established, up for 00:00:07

	Last read , hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Four-octets ASN Capability: advertised and received Address family IPv4 Unicast: advertised and received Received 2 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:1 refresh message:0 dynamic cap:0 notifications:0 Sent 2 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:1 refresh message:0 dynamic cap:0 notifications:0 Route refresh request: received 0, sent 0 Minimum time between advertisement runs is 0 seconds Update source is Loopback 0 For address family: IPv4 Unicast BGP table version 1, neighbor version 1 Index 0, Offset 0, Mask 0x1 0 accepted prefixes 0 announced prefixes Connections established 1; dropped 0 Local host: 10.1.1.2, Local port: 1041 Foreign host: 10.1.1.3, Foreign port: 179 Nexthop: 10.1.1.2 Nexthop global: :: Nexthop local: :: BGP connection: non shared network Last Reset: , due to BGP Notification received Notification Error Message: (Cease/Other Configuration Change.)
C	C# show ip bgp neighbor BGP neighbor is 10.1.1.1, remote AS 65536, local AS 65536, internal link Member of peer-group ibgp-group for session parameters

```
  BGP version 4, remote router ID 10.1.1.1
  BGP state = Established, up for 00:01:13
  Last read                , hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received (old and new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
  Received 3 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:2
    refresh message:0 dynamic cap:0 notifications:0
  Sent 3 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:2
    refresh message:0 dynamic cap:0 notifications:0
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 0 seconds
  Update source is Loopback 0

 For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Index 1, Offset 0, Mask 0x2
  ibgp-group peer-group member
  0 accepted prefixes
  0 announced prefixes

 Connections established 1; dropped 0
Local host: 10.1.1.3, Local port: 179
Foreign host: 10.1.1.1, Foreign port: 1039
Nexthop: 10.1.1.3
Nexthop global: ::
Nexthop local: ::
BGP connection: non shared network

BGP neighbor is 10.1.1.2, remote AS 65536, local AS 65536, internal link
 Member of peer-group ibgp-group for session parameters
```

```
  BGP version 4, remote router ID 10.1.1.2
  BGP state = Established, up for 00:01:17
  Last read                , hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received (old and new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
  Received 3 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:2
    refresh message:0 dynamic cap:0 notifications:0
  Sent 3 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:2
    refresh message:0 dynamic cap:0 notifications:0
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 0 seconds
  Update source is Loopback 0

 For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Index 1, Offset 0, Mask 0x2
  ibgp-group peer-group member
  0 accepted prefixes
  0 announced prefixes

 Connections established 1; dropped 0
Local host: 10.1.1.3, Local port: 179
Foreign host: 10.1.1.2, Foreign port: 1041
Nexthop: 10.1.1.3
Nexthop global: ::
Nexthop local: ::
BGP connection: non shared network

BGP neighbor is 192.168.3.4, remote AS 65536, local AS 65536, internal link
 Member of peer-group ibgp-group for session parameters
```

	BGP version 4, remote router ID 192.168.3.4 BGP state = Established, up for 00:01:01 Last read , hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Four-octets ASN Capability: advertised and received Address family IPv4 Unicast: advertised and received Received 3 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:2 refresh message:0 dynamic cap:0 notifications:0 Sent 3 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:2 refresh message:0 dynamic cap:0 notifications:0 Route refresh request: received 0, sent 0 Minimum time between advertisement runs is 0 seconds Update source is Loopback 0 For address family: IPv4 Unicast BGP table version 1, neighbor version 1 Index 1, Offset 0, Mask 0x2 ibgp-group peer-group member 0 accepted prefixes 0 announced prefixes Connections established 1; dropped 0 Local host: 192.168.3.3, Local port: 179 Foreign host: 192.168.3.4, Foreign port: 1018 Nexthop: 192.168.3.3 Nexthop global: :: Nexthop local: :: BGP connection: non shared network
D	D# show ip bgp neighbor BGP neighbor is 192.168.3.3, remote AS 65536, local AS 65536, internal link

```
 Member of peer-group ibgp-group for session parameters
   BGP version 4, remote router ID 10.1.1.3
   BGP state = Established, up for 00:01:01
   Last read                , hold time is 180, keepalive interval is 60 seconds
   Neighbor capabilities:
     Route refresh: advertised and received (old and new)
     Four-octets ASN Capability: advertised and received
     Address family IPv4 Unicast: advertised and received
   Received 3 messages, 0 notifications, 0 in queue
     open message:1 update message:0 keepalive message:2
     refresh message:0 dynamic cap:0 notifications:0
   Sent 3 messages, 0 notifications, 0 in queue
     open message:1 update message:0 keepalive message:2
     refresh message:0 dynamic cap:0 notifications:0
   Route refresh request: received 0, sent 0
   Minimum time between advertisement runs is 0 seconds
   Update source is Loopback 0

 For address family: IPv4 Unicast
   BGP table version 1, neighbor version 1
   Index 1, Offset 0, Mask 0x2
   ibgp-group peer-group member
   0 accepted prefixes
   0 announced prefixes

 Connections established 1; dropped 0
Local host: 192.168.3.4, Local port: 1018
Foreign host: 192.168.3.3, Foreign port: 179
Nexthop: 192.168.3.4
Nexthop global: ::
Nexthop local: ::
BGP connection: non shared network
```

Common Errors

- IGP is not enabled and the interconnection between the local loopback address and the loopback address on the IBGP neighbor fails, which causes that the neighbor fails to be created.
- ebgp-multihop is not configured when an EBGP is not directly connected, which causes that a TCP connection fails to be created.

5.4.2 Configuring MD5 Authentication

Configuration Effect

- Configure MD5 for encrypted authentication between EBGP and IBGP neighbors.

Notes

- If an IBGP neighbor is not directly connected, you need to configure IGP or a static routing protocol to implement interconnection.
- If an EBGP neighbor is not directly connected, you need to configure the **ebgp-multihop** parameter for the neighbor.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Verification

- Run the **show** command to display the neighbor status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

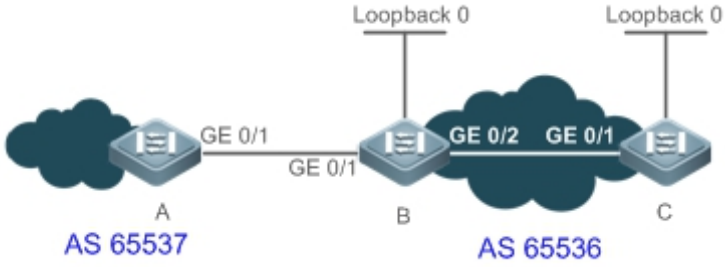
Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring an MD5 Password for a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **password** [**0** | **7**] *string*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. **0**: Displays a password not encrypted. **7**: Displays a password encrypted. *string*: Indicates a password for TCP MD5 authentication, consisting of a maximum of 80 characters.
Command Mode	BGP configuration mode
Usage Guide	The same passwords must be configured on the two ends of a BGP neighborship.

Configuration Example

Configuring BGP MD5 Authentication

Scenario **Figure 5- 5**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5- 5. ● Configure a loopback interface on B and C and create an IBGP neighbor based on the loopback interface. ● Create an EBGP neighborship by using the directly connected interfaces on A and B. ● Configure the passwords on A, B and C for their neighbors.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0 A(config-if-GigabitEthernet 0/1)# exit A(config)# router bgp 65537 A(config-router)# neighbor 192.168.1.2 remote-as 65536 A(config-router)# neighbor 192.168.1.2 password 7 ebgpneighbor
B	B# configure terminal B(config)# interface loopback 0 B(config-if-Loopback 0)# ip address 10.1.1.1 255.255.255.255 B(config-if-Loopback 0)# exit

	B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip address 192.168.1.2 255.255.255.0 B(config-if-GigabitEthernet 0/1)# exit B(config)# interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)# ip address 192.168.2.2 255.255.255.0 B(config-if-GigabitEthernet 0/2)# exit B(config)# router bgp 65536 B(config-router)# neighbor 10.1.1.2 remote-as 65536 B(config-router)# neighbor 10.1.1.2 update-source loopback 0 B(config-router)# neighbor 10.1.1.2 password ibgpneighbor B(config-router)# neighbor 192.168.1.1 remote-as 65537 B(config-router)# neighbor 192.168.1.1 password 7 ebgpneighbor
C	C# configure terminal C(config)# interface loopback 0 C(config-if-Loopback 0)# ip address 10.1.1.2 255.255.255.255 C(config-if-Loopback 0)# exit C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 192.168.2.3 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# router bgp 65536 C(config-router)# neighbor 10.1.1.1 remote-as 65536 C(config-router)# neighbor 10.1.1.1 update-source loopback 0 C(config-router)# neighbor 10.1.1.1 password ibgpneighbor
Verification	Run the **show** command to display the BGP neighbor status.

A	A#show ip bgp neighbors
	BGP neighbor is 192.168.1.2, remote AS 65536, local AS 65537, external link
	  BGP version 4, remote router ID 10.1.1.1
	  BGP state = Established, up for 00:04:54
	  Last read , hold time is 180, keepalive interval is 60 seconds
	  Neighbor capabilities:
	  Route refresh: advertised and received (old and new)
	  Four-octets ASN Capability: advertised and received

	Address family IPv4 Unicast: advertised and received Received 7 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:6 refresh message:0 dynamic cap:0 notifications:0 Sent 7 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:6 refresh message:0 dynamic cap:0 notifications:0 Route refresh request: received 0, sent 0 Minimum time between advertisement runs is 30 seconds For address family: IPv4 Unicast BGP table version 1, neighbor version 0 Index 1, Offset 0, Mask 0x2 0 accepted prefixes 0 announced prefixes Connections established 2; dropped 1 Local host: 192.168.1.1, Local port: 1026 Foreign host: 192.168.1.2, Foreign port: 179 Nexthop: 192.168.1.1 Nexthop global: :: Nexthop local: :: BGP connection: non shared network Last Reset: 00:04:54, due to BGP Notification sent Notification Error Message: (Cease/Administratively Reset.)
B	B# show ip bgp neighbors BGP neighbor is 10.1.1.2, remote AS 65536, local AS 65536, internal link BGP version 4, remote router ID 10.1.1.2 BGP state = Established, up for 00:04:01 Last read , hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Four-octets ASN Capability: advertised and received

```
    Address family IPv4 Unicast: advertised and received
  Received 8 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:7
    refresh message:0 dynamic cap:0 notifications:0
  Sent 8 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:7
    refresh message:0 dynamic cap:0 notifications:0
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 30 seconds

 For address family: IPv4 Unicast
  BGP table version 1, neighbor version 0
  Index 1, Offset 0, Mask 0x2
  0 accepted prefixes
  0 announced prefixes

 Connections established 2; dropped 1
Local host: 10.1.1.1, Local port: 179
Foreign host: 10.1.1.2, Foreign port: 1038
Nexthop: 10.1.1.1
Nexthop global: ::
Nexthop local: ::
BGP connection: non shared network
Last Reset: 00:05:27, due to BGP Notification received
Notification Error Message: (Cease/Administratively Reset.)

BGP neighbor is 192.168.1.1, remote AS 65537, local AS 65536, external link
  BGP version 4, remote router ID 192.168.1.1
  BGP state = Established, up for 00:05:27
  Last read                , hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received (old and new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
```

	Received 8 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:7 refresh message:0 dynamic cap:0 notifications:0 Sent 8 messages, 0 notifications, 0 in queue open message:1 update message:0 keepalive message:7 refresh message:0 dynamic cap:0 notifications:0 Route refresh request: received 0, sent 0 Minimum time between advertisement runs is 30 seconds For address family: IPv4 Unicast BGP table version 1, neighbor version 0 Index 1, Offset 0, Mask 0x2 0 accepted prefixes 0 announced prefixes Connections established 2; dropped 1 Local host: 192.168.1.2, Local port: 179 Foreign host: 192.168.1.1, Foreign port: 1026 Nexthop: 192.168.1.2 Nexthop global: :: Nexthop local: :: BGP connection: non shared network Last Reset: 00:05:27, due to BGP Notification received Notification Error Message: (Cease/Administratively Reset.)
C	C# show ip bgp neighbors BGP neighbor is 10.1.1.1, remote AS 65536, local AS 65536, internal link BGP version 4, remote router ID 10.1.1.1 BGP state = Established, up for 00:04:01 Last read , hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Four-octets ASN Capability: advertised and received Address family IPv4 Unicast: advertised and received

```
  Received 8 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:7
    refresh message:0 dynamic cap:0 notifications:0
  Sent 8 messages, 0 notifications, 0 in queue
    open message:1 update message:0 keepalive message:7
    refresh message:0 dynamic cap:0 notifications:0
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 30 seconds

 For address family: IPv4 Unicast
  BGP table version 1, neighbor version 0
  Index 1, Offset 0, Mask 0x2
  0 accepted prefixes
  0 announced prefixes

 Connections established 2; dropped 1
Local host: 10.1.1.2, Local port: 1038
Foreign host: 10.1.1.1, Foreign port: 179
Nexthop: 10.1.1.2
Nexthop global: ::
Nexthop local: ::
BGP connection: non shared network
Last Reset: 00:05:27, due to BGP Notification received
Notification Error Message: (Cease/Administratively Reset.)
```

Common Errors

- The passwords for MD5 encrypted authentication at the two ends of a BGP neighborship are different.

5.4.3 Configuring a Route Reflector

Configuration Effect

- Configure a route reflector in the IBGP environment to reduce the number of BGP neighbor connections.

Notes

- If an IBGP neighbor is not directly connected, you need to configure IGP or a static routing protocol to implement interconnection.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Creating a BGP Reflector

- (Mandatory) Perform this configuration in the BGP configuration mode.

Verification

- Run the **show** command to display the neighbor status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

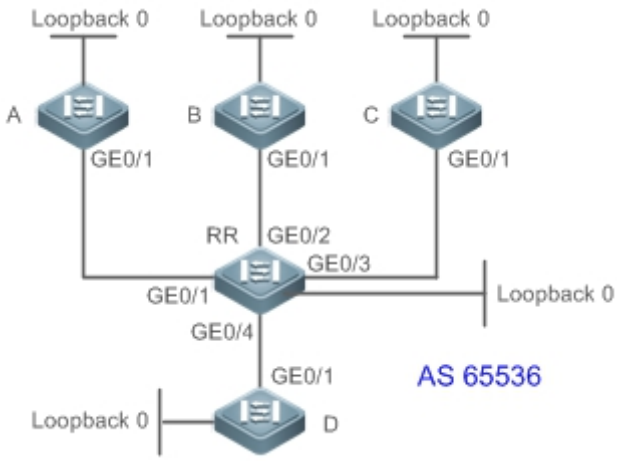
Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Creating a BGP Reflector

Command	**neighbor** { *peer-address* | *peer-group-name* } **route-reflector-client**
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters.
Command Mode	BGP configuration mode
Usage Guide	-

Configuration Example

Configuring a BGP Route Reflector

Scenario **Figure 5-6**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5-6. ● Configure a loopback interface on all devices and create an IBGP neighborship by using the loopback interface according to the connection lines as shown in Figure 5-6. ● Configure route reflection on the device RR and specify A, B, C and D as reflector clients.
A	A# configure terminal A(config)# interface loopback 0 A(config-if-Loopback 0)# ip address 10.1.1.1 255.255.255.255 A(config-if-Loopback 0)# exit A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0 A(config-if-GigabitEthernet 0/1)# exit A(config)# router bgp 65536 A(config-router)# neighbor 10.1.1.5 remote-as 65536 A(config-router)# neighbor 10.1.1.5 update-source loopback 0 A(config-router)# network 192.168.1.0 mask 255.255.255.0
B	B# configure terminal B(config)# interface loopback 0 B(config-if-Loopback 0)# ip address 10.1.1.2 255.255.255.255 B(config-if-Loopback 0)# exit B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip address 192.168.2.2 255.255.255.0 B(config-if-GigabitEthernet 0/1)# exit B(config)# router bgp 65536 B(config-router)# neighbor 10.1.1.5 remote-as 65536

	B(config-router)# neighbor 10.1.1.5 update-source loopback 0
C	C# configure terminal C(config)# interface loopback 0 C(config-if-Loopback 0)# ip address 10.1.1.3 255.255.255.255 C(config-if-Loopback 0)# exit C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 192.168.3.3 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# router bgp 65536 C(config-router)# neighbor 10.1.1.5 remote-as 65536 C(config-router)# neighbor 10.1.1.5 update-source loopback 0
D	C# configure terminal C(config)# interface loopback 0 C(config-if-Loopback 0)# ip address 10.1.1.4 255.255.255.255 C(config-if-Loopback 0)# exit C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 192.168.4.4 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# router bgp 65536 C(config-router)# neighbor 10.1.1.5 remote-as 65536 C(config-router)# neighbor 10.1.1.5 update-source loopback 0
RR	RR# configure terminal RR(config)# interface loopback 0 RR(config-if-Loopback 0)# ip address 10.1.1.5 255.255.255.255 RR(config-if-Loopback 0)# exit RR(config)# interface GigabitEthernet 0/1 RR(config-if-GigabitEthernet 0/1)# ip address 192.168.1.5 255.255.255.0 RR(config-if-GigabitEthernet 0/1)# exit RR(config)# interface GigabitEthernet 0/2 RR(config-if-GigabitEthernet 0/2)# ip address 192.168.2.5 255.255.255.0 RR(config-if-GigabitEthernet 0/2)# exit RR(config)# interface GigabitEthernet 0/3 RR(config-if-GigabitEthernet 0/3)# ip address 192.168.3.5 255.255.255.0

```
RR(config-if-GigabitEthernet 0/3)# exit
RR(config)# interface GigabitEthernet 0/4
RR(config-if-GigabitEthernet 0/4)# ip address 192.168.4.5 255.255.255.0
RR(config-if-GigabitEthernet 0/4)# exit
RR(config)# router bgp 65536
RR(config-router)# neighbor 10.1.1.1 remote-as 65536
RR(config-router)# neighbor 10.1.1.1 update-source loopback 0
RR(config-router)# neighbor 10.1.1.1 route-reflector-client
RR(config-router)# neighbor 10.1.1.2 remote-as 65536
RR(config-router)# neighbor 10.1.1.2 update-source loopback 0
RR(config-router)# neighbor 10.1.1.2 route-reflector-client
RR(config-router)# neighbor 10.1.1.3 remote-as 65536
RR(config-router)# neighbor 10.1.1.3 update-source loopback 0
RR(config-router)# neighbor 10.1.1.3 route-reflector-client
RR(config-router)# neighbor 10.1.1.4 remote-as 65536
RR(config-router)# neighbor 10.1.1.4 update-source loopback 0
RR(config-router)# neighbor 10.1.1.4 route-reflector-client
```

Verification

Run the **show** command to display the BGP neighbor status.

RR

```
RR# show ip bgp summary
BGP router identifier 10.1.1.5, local AS number 65536
BGP table version is 1
0 BGP AS-PATH entries
0 BGP Community entries
1 BGP Prefix entries (Maximum-prefix:4294967295)

Neighbor        V         AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.1.1.1        4      65536       8       9        1    0    0 00:05:11        1
10.1.1.2        4      65536       9       9        1    0    0 00:05:24        0
10.1.1.3        4      65536       8       7        1    0    0 00:05:10        0
10.1.1.4        4      65536       9       8        1    0    0 00:05:14        0

RR# show ip bgp
BGP table version is 1, local router ID is 10.1.1.5
```

	Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale, b - backup entry Origin codes: i - IGP, e - EGP, ? - incomplete Network Next Hop Metric LocPrf Weight Path *>i192.168.1.0 10.1.1.1 0 100 0 i Total number of prefixes 1
D	D# show ip bgp summary BGP router identifier 10.1.1.4, local AS number 65536 BGP table version is 1 0 BGP AS-PATH entries 0 BGP Community entries 1 BGP Prefix entries (Maximum-prefix:4294967295) Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd 10.1.1.5 4 65536 8 9 1 0 0 00:05:20 1 D# show ip bgp BGP table version is 1, local router ID is 10.1.1.4 Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale, b - backup entry Origin codes: i - IGP, e - EGP, ? - incomplete Network Next Hop Metric LocPrf Weight Path * i192.168.1.0 10.1.1.1 0 100 0 i Total number of prefixes 1

5.4.4 Configuring an AS Alliance

Configuration Effect

- Configure a BGP alliance to reduce the number of BGP neighbor connections.

Notes

- It is advised to use private AS numbers for sub ASs (also called member ASs) within an alliance. Private AS numbers range from 64,512 to 65,535.
- Within a sub AS of an alliance, full mesh must be established for all BGP speakers (route reflectors can be further configured within the sub AS).
- An EBGP neighborship must be established between sub ASs of an alliance.
- All BGP speakers within an alliance must belong to a sub AS within the alliance.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Configuring a BGP Alliance ID

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring a BGP Alliance Member

- (Mandatory) Perform this configuration in the BGP configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring Multiple Hops for an EBGP Neighbor

- Perform this configuration in the BGP configuration mode. It is mandatory when an EBGP neighbor is not directly connected.

Configuring BGP Route Re-distribution to a Network

- (Optional) Perform this configuration in the BGP configuration mode. Perform this configuration when a local route needs to be advertised. You can also configure an alternative network by means of re-distribution.

Verification

- Run the **show** command to display the BGP neighbor status.
- Run the **show** command to display the BGP routing table information.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Enabling a BGP Alliance ID

Command	**bgp confederation identifier** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	BGP configuration mode
Usage Guide	-

Configuring a BGP Alliance Member

Command	**bgp confederation peers** *as-number* [… *as-number*]
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	BGP configuration mode
Usage Guide	All member ASs of a local EBGP alliance must be identified.

Creating a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring Multiple Hops for an EBGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **ebgp-multihop** [*ttl*]
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *ttl*: Indicates the maximum number of hops that are allowed, ranging from 1 to 255.
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

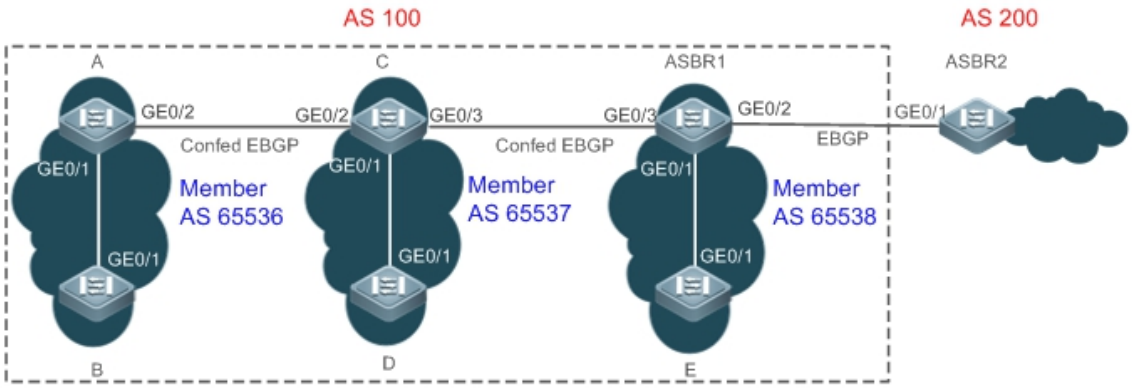
Configuring BGP Route Re-distribution to a Network

Command	**network** *network-number* [**mask** *mask*] [**route-map** *map-tag*] [**backdoor**]
Parameter Description	*network-number*: Indicates the network address. *mask*: Indicates the subnet mask. *map-tag*: Indicates the name of a route map, consisting of no more than 32 characters. **backdoor**: Indicates that the route is a backdoor route.
Command Mode	BGP configuration mode

Usage Guide	The core routing table must contain same IGP (or static and directly connected) routes.

Configuration Example

Configuring a BGP Alliance

Scenario **Figure 5-7**	
Configuration Steps	● Configure BGP on A and B, set the AS number to 65,536 and configure an IBGP neighborship. ● Configure BGP on C and D, set the AS number to 65,537 and configure an IBGP neighborship. ● Configure BGP on ASBR1 and E, set the AS number to 65,538 and configure an IBGP neighborship. ● Configure an alliance ID 100 on A, B, C, D, E and ASBR1. ● Configure the alliance member 65,537 on A, configure C as an EBGP neighbor, and set the peer AS number to 65,537. ● Configure the alliance members 65,536 and 65,538 on C, configure A as an EBGP neighbor and set the peer AS number to 65,536, configure ASBR1 as an EBGP neighbor and set the peer AS number to 65,538. ● Configure the alliance members 65,537 on ASBR1, configure C as an EBGP neighbor and set the peer AS number to 65,537, configure ASBR2 as an EBGP neighbor and set the peer AS number to 200. ● Configure BGP on ASBR2 and set the AS number to 200; configure ASBR1 as an EBGP neighbor and set the peer AS number to 100.
A	A# configure terminal A(config)# interface loopback 0 A(config-if-Loopback 0)# ip address 10.1.1.1 255.255.255.255 A(config-if-Loopback 0)# exit A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0 A(config-if-GigabitEthernet 0/1)# exit A(config)# interface GigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)# ip address 192.168.2.1 255.255.255.0 A(config-if-GigabitEthernet 0/2)# exit A(config)# router bgp 65536 A(config-router)# bgp confederation identifier 100

	A(config-router)# bgp confederation peers 65537 A(config-router)# neighbor 10.1.1.2 remote-as 65536 A(config-router)# neighbor 10.1.1.2 update-source loopback 0 A(config-router)# neighbor 10.1.1.3 remote-as 65537 A(config-router)# neighbor 10.1.1.3 ebgp-multihop 2 A(config-router)# neighbor 10.1.1.3 update-source loopback 0 A(config-router)# network 192.168.1.0 mask 255.255.255.0
B	B# configure terminal B(config)# interface loopback 0 B(config-if-Loopback 0)# ip address 10.1.1.2 255.255.255.255 B(config-if-Loopback 0)# exit B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip address 192.168.1.2 255.255.255.0 B(config-if-GigabitEthernet 0/1)# exit B(config)# router bgp 65536 B(config-router)# neighbor 10.1.1.1 remote-as 65536 B(config-router)# neighbor 10.1.1.1 update-source loopback 0
C	C# configure terminal C(config)# interface loopback 0 C(config-if-Loopback 0)# ip address 10.1.1.3 255.255.255.255 C(config-if-Loopback 0)# exit C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 192.168.3.3 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# interface GigabitEthernet 0/2 C(config-if-GigabitEthernet 0/2)# ip address 192.168.2.3 255.255.255.0 C(config-if-GigabitEthernet 0/2)# exit C(config)# interface GigabitEthernet 0/3 C(config-if-GigabitEthernet 0/3)# ip address 192.168.4.3 255.255.255.0 C(config-if-GigabitEthernet 0/3)# exit C(config)# router bgp 65537 C(config-router)# bgp confederation identifier 100 C(config-router)# bgp confederation peers 65536 65538

	C(config-router)# neighbor 10.1.1.1 remote-as 65536 C(config-router)# neighbor 10.1.1.1 update-source loopback 0 C(config-router)# neighbor 10.1.1.1 ebgp-multihop 2 C(config-router)# neighbor 10.1.1.4 remote-as 65537 C(config-router)# neighbor 10.1.1.4 update-source loopback 0 C(config-router)# neighbor 10.1.1.5 remote-as 65538 C(config-router)# neighbor 10.1.1.5 update-source loopback 0 C(config-router)# neighbor 10.1.1.5 ebgp-multihop 2
D	D# configure terminal D(config)# interface loopback 0 D(config-if-Loopback 0)# ip address 10.1.1.4 255.255.255.255 D(config-if-Loopback 0)# exit D(config)# interface GigabitEthernet 0/1 D(config-if-GigabitEthernet 0/1)# ip address 192.168.3.4 255.255.255.0 D(config-if-GigabitEthernet 0/1)# exit D(config)# router bgp 65537 D(config-router)# neighbor 10.1.1.3 remote-as 65537 D(config-router)# neighbor 10.1.1.3 update-source loopback 0
E	E# configure terminal E(config)# interface loopback 0 E(config-if-Loopback 0)# ip address 10.1.1.6 255.255.255.255 E(config-if-Loopback 0)# exit E(config)# interface GigabitEthernet 0/1 E(config-if-GigabitEthernet 0/1)# ip address 192.168.5.6 255.255.255.0 E(config-if-GigabitEthernet 0/1)# exit E(config)# router bgp 65538 E(config-router)# neighbor 10.1.1.5 remote-as 65538 E(config-router)# neighbor 10.1.1.5 update-source loopback 0
ASBR1	ASBR1# configure terminal ASBR1(config)# interface loopback 0 ASBR1(config-if-Loopback 0)# ip address 10.1.1.5 255.255.255.255 ASBR1(config-if-Loopback 0)# exit ASBR1(config)# interface GigabitEthernet 0/1

	ASBR1(config-if-GigabitEthernet 0/1)# ip address 192.168.5.5 255.255.255.0 ASBR1(config-if-GigabitEthernet 0/1)# exit ASBR1(config)# interface GigabitEthernet 0/2 ASBR1(config-if-GigabitEthernet 0/2)# ip address 192.168.6.5 255.255.255.0 ASBR1(config-if-GigabitEthernet 0/2)# exit ASBR1(config)# interface GigabitEthernet 0/3 ASBR1(config-if-GigabitEthernet 0/3)# ip address 192.168.4.5 255.255.255.0 ASBR1(config-if-GigabitEthernet 0/3)# exit ASBR1(config)# router bgp 65538 ASBR1(config-router)# bgp confederation identifier 100 ASBR1(config-router)# bgp confederation peers 65537 ASBR1(config-router)# neighbor 10.1.1.3 remote-as 65537 ASBR1(config-router)# neighbor 10.1.1.3 update-source loopback 0 ASBR1(config-router)# neighbor 10.1.1.3 ebgp-multihop 2 ASBR1(config-router)# neighbor 10.1.1.6 remote-65538 ASBR1(config-router)# neighbor 10.1.1.6 update-source loopback 0 ASBR1(config-router)# neighbor 192.168.6.7 remote-as 200
ASBR2	ASBR2# configure terminal ASBR2(config)# interface GigabitEthernet 0/1 ASBR2(config-if-GigabitEthernet 0/1)# ip address 192.168.6.7 255.255.255.0 ASBR2(config-if-GigabitEthernet 0/1)# exit ASBR2(config)# router bgp 200 ASBR2(config-router)# neighbor 192.168.6.5 remote-as 100 ASBR2(config-router)# network 192.168.6.0 mask 255.255.255.0
Verification	Run the **show** command to display the information.
A	A# show ip bgp summary BGP router identifier 10.1.1.1, local AS number 65536 BGP table version is 1 1 BGP AS-PATH entries 0 BGP Community entries 1 BGP Prefix entries (Maximum-prefix:4294967295)

```
Neighbor        V         AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.1.1.2        4      65536       3       3        1    0    0 00:00:05        0
10.1.1.3        4      65537       3       3        1    0    0 00:00:06        1

Total number of neighbors 1

A# show ip bgp
BGP table version is 1, local router ID is 10.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale, b - backup entry
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric     LocPrf     Weight Path
*  192.168.6.0      192.168.6.7              0        100            0 (65537 65538) 200 i

Total number of prefixes 1
```

ASBR1

```
A# show ip bgp summary
BGP router identifier 10.1.1.5, local AS number 200
BGP table version is 2
2 BGP AS-PATH entries
0 BGP Community entries
2 BGP Prefix entries (Maximum-prefix:4294967295)

Neighbor        V         AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.1.1.3        4      65537       3       3        2    0    0 00:00:10        1
10.1.1.6        4      65538       3       3        2    0    0 00:00:08        0
192.168.6.7     4        200       3       3        2    0    0 00:00:05        1

Total number of neighbors 1

A# show ip bgp
BGP table version is 1, local router ID is 10.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
```

	S Stale, b - backup entry Origin codes: i - IGP, e - EGP, ? - incomplete Network Next Hop Metric LocPrf Weight Path * 192.168.1.0 10.1.1.1 0 100 0 (65537 65536) i *> 192.168.6.0 192.168.6.7 0 100 0 200 i Total number of prefixes 1
ASBR2	A# show ip bgp summary BGP router identifier 192.168.6.7, local AS number 200 BGP table version is 1 1 BGP AS-PATH entries 0 BGP Community entries 1 BGP Prefix entries (Maximum-prefix:4294967295) Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd 192.168.6.5 4 100 3 3 1 0 0 00:00:05 1 Total number of neighbors 1 A# show ip bgp BGP table version is 1, local router ID is 10.1.1.1 Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale, b - backup entry Origin codes: i - IGP, e - EGP, ? - incomplete Network Next Hop Metric LocPrf Weight Path *> 192.168.1.0 192.168.6.5 0 100 0 (65537 65538) 200 i Total number of prefixes 1

Common Errors

- No BGP alliance neighbor is configured.
- Full mesh is not established within sub ASs of an alliance.

5.4.5 Configuring Multi-path Load Balancing of BGP

Configuration Effect

- Implement multi-path load balancing for IBGP routes.
- Support AS-PATH loose comparison.

Notes

- Routes learned from an IBGP neighbor must have the same priority (the router-ID does not need to be compared).

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring BGP Load Balancing

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring AS-PATH Loose Comparison

- (Optional) Perform this configuration in the BGP configuration mode. Perform this configuration when load balancing needs to be implemented for routes learned from different ASs.

Verification

- Run the **show** command to display BGP routing information.
- Run the **show** command to display the core routing table information.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).

Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring BGP Load Balancing

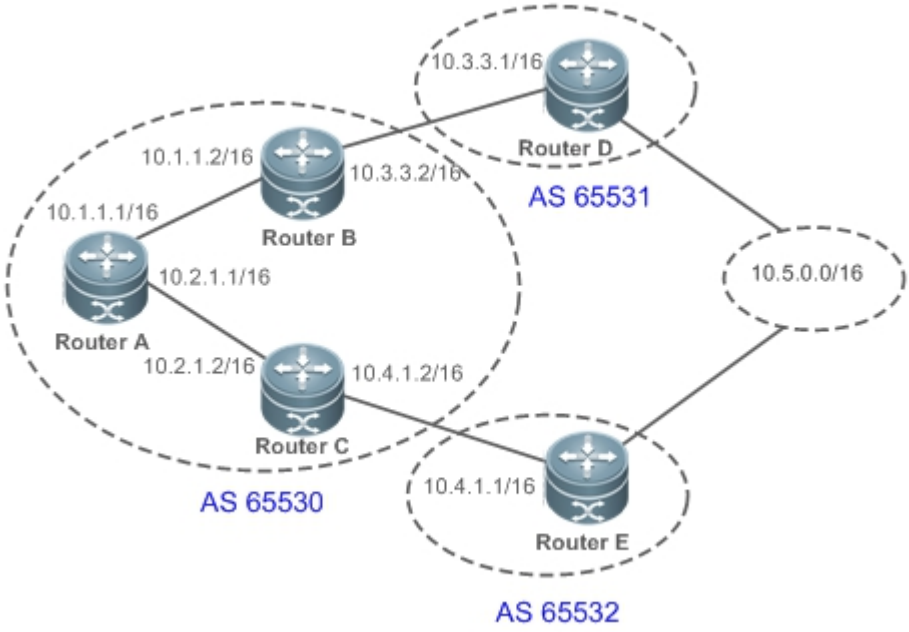
Command	**maximum-paths** { **ebgp** | **ibgp** } *number*
Parameter Description	*number*: Indicates the maximum number of equivalent paths, ranging from 1 to device capacity. If the value is **1**, multi-path load balancing of IBGP will be disabled.
Command Mode	BGP configuration mode
Usage Guide	-

Configuring AS-PATH Loose Comparison

Command	**bgp bestpath as-path multipath-relax**
Parameter Description	**-**
Command Mode	BGP configuration mode
Usage Guide	-

Configuration Example

Configuring Multi-path Load Balancing of IBGP

Scenario **Figure 5-8**	
Configuration	● Enable BGP on all devices and set the AS numbers as shown in Figure 5-8.

Steps	● Establish IBGP neighborships between A and B and between A and C by using directly connected interfaces. ● Establish EBGP neighborships between B and D and between C and E by using directly connected interfaces. ● Re-distribute the same routes to D and E. ● Configure IBGP load balancing on A and enable the AS-PATH loose comparison mode.
A	A# conf terminal A(config)# interface fastEthernet 0/0 A(config-if-FastEthernet 0/0)# ip address 10.1.1.1 255.255.0.0 A(config-if-FastEthernet 0/0)# exit A(config)# interface fastEthernet 0/1 A(config-if-FastEthernet 0/1)# ip address 10.2.1.1 255.255.0.0 A(config-if-FastEthernet 0/1)# exit A(config)# ip route 10.3.0.0 255.255.0.0 10.1.1.2 A(config)# ip route 10.4.0.0 255.255.0.0 10.2.1.2 A(config)# router bgp 65530 A(config-router)# neighbor 10.1.1.2 remote-as 65530 A(config-router)# neighbor 10.2.1.2 remote-as 65530 A(config-router)# bgp maximum-paths ibgp 2 A(config-router)# bgp bestpath as-path multipath-relax
B	B# conf terminal B(config)# interface fastEthernet 0/0 B(config-if-FastEthernet 0/0)# ip address 10.1.1.2 255.255.0.0 B(config-if-FastEthernet 0/0)# exit B(config)# interface fastEthernet 0/1 B(config-if-FastEthernet 0/1)# ip address 10.3.1.2 255.255.0.0 B(config-if-FastEthernet 0/1)# exit B(config)# router bgp 65530 B(config-router)# neighbor 10.1.1.1 remote-as 65530 B(config-router)# neighbor 10.3.1.1 remote-as 65531
C	C# conf terminal C(config)# interface fastEthernet 0/0 C(config-if-FastEthernet 0/0)# ip address 10.2.1.2 255.255.0.0 C(config-if-FastEthernet 0/0)# exit C(config)# interface fastEthernet 0/1 C(config-if-FastEthernet 0/1)# ip address 10.4.1.2 255.255.0.0

	C(config-if-FastEthernet 0/1)# exit C(config)# router bgp 65530 C(config-router)# neighbor 10.2.1.1 remote-as 65530 C(config-router)# neighbor 10.4.1.1 remote-as 65532
D	D# conf terminal D(config)# interface fastEthernet 0/0 D(config-if-FastEthernet 0/0)# ip address 10.3.1.1 255.255.0.0 D(config-if-FastEthernet 0/0)# exit D(config)# interface loopback 1 D(config-if)#ip address 10.5.1.1 255.255.0.0 D(config-if-FastEthernet 0/1)# exit D(config)# router bgp 65531 D(config-router)# neighbor 10.3.1.2 remote-as 65530 D(config-router)# redistribute connected
E	E# conf terminal E(config)# interface fastEthernet 0/0 E(config-if-FastEthernet 0/0)# ip address 10.4.1.1 255.255.0.0 E(config-if-FastEthernet 0/0)# exit E(config)# interface loopback 1 E(config-if)#ip address 10.5.1.2 255.255.0.0 E(config-if-FastEthernet 0/1)# exit E(config)# router bgp 65532 E(config-router)# neighbor 10.4.1.2 remote-as 65530 E(config-router)# redistribute connected
Verification	Run the **show** command to display the information.
A	A# show ip bgp summary BGP router identifier 10.2.1.1, local AS number 65530 BGP table version is 9 2 BGP AS-PATH entries 0 BGP Community entries 3 BGP Prefix entries (Maximum-prefix:4294967295)

```
Neighbor        V    AS      MsgRcvd   MsgSent   TblVer   InQ   OutQ   Up/Down     State/PfxRcd
172.16.23.140   4    65530   29        25        8        0     0      00:18:48    2
172.16.23.141   4    65530   24        21        8        0     0      00:17:58    2

A# show ip bgp
BGP table version is 9, local router ID is 10.2.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric     LocPrf     Weight Path
*>i10.3.0.0/16      10.3.1.1                 0        100          0 65531 ?
*>i10.4.0.0/16      10.4.1.1                 0        100          0 65532 ?
* i10.5.0.0/16      10.3.1.1                 0        100          0 65531 ?
*>i                 10.4.1.1                 0        100          0 65532 ?

Total number of prefixes 3
A# show ip bgp 10.5.0.0
BGP routing table entry for 10.5.0.0/16
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Not advertised to any peer
  65532
    10.4.1.1 from 10.2.1.2 (172.16.24.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, multipath, best
      Last update: Mon Mar 21 03:45:14 2011

  65531
    10.3.1.1 from 10.1.1.2 (172.16.25.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, multipath
      Last update: Mon Mar 21 03:45:14 2011

A# show ip route

Codes:  C - connected, S - static, R - RIP, B - BGP
```

```
            O - OSPF, IA - OSPF inter area
            N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
            E1 - OSPF external type 1, E2 - OSPF external type 2
            i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
            ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C       10.1.0.0/16 is directly connected, FastEthernet 0/0
C       10.1.1.1/32 is local host.
C       10.2.0.0/16 is directly connected, FastEthernet 0/1
C       10.2.1.1/32 is local host.
S       10.3.0.0/16 [1/0] via 10.1.1.2
S       10.4.0.0/16 [1/0] via 10.2.1.2
B       10.5.0.0/16 [200/0] via 10.3.1.1, 00:27:56
                    [200/0] via 10.4.1.1, 00:27:56
```

Common Errors

- The priorities of multi-hop BGP routes are different, which causes load balancing failure.

5.4.6 Configuring EBGP FRR

Configuration Effect

- Implement EBGP FRR.

Notes

- (Optional) Configure a neighbor BFD session to implement fast link fault detection.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring BGP FRR

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring a Neighbor BFD Session

- (Optional) Perform this configuration in the BGP configuration mode.

Verification

- Run the **show** command to display routing information.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Configuring BGP FRR

Command	**bgp fast-reroute**
Parameter Description	**-**
Command Mode	BGP configuration mode
Usage Guide	-

Creating a BGP Neighbor

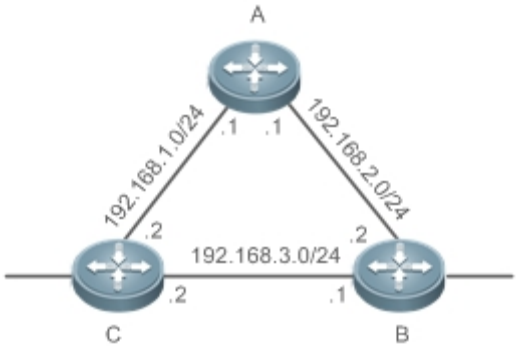
Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Creating a BFD Session to a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **fall-over bfd**
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters.
Command Mode	BGP configuration mode
Usage Guide	-

Configuration Example

Configuring EBGP FRR

Scenario **Figure 5- 9**	
Configuration Steps	● Enable BGP on all devices. ● Configure the addresses of the directly connected interfaces on A, B and C to establish EBGP neighborships. ● Configure a BFD session for the EBGP neighborship between B and C. ● Configure FRR on C.
A	A# conf terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0 A(config-if-GigabitEthernet 0/1)# exit A(config)# interface GigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)# ip address 192.168.2.1 255.255.255.0 A(config-if-GigabitEthernet 0/2)# exit A(config)# router bgp 100 A(config-router)# neighbor 192.168.1.2 remote-as 300 A(config-router)# neighbor 192.168.2.2 remote-as 200 A(config-router)# redistribute connect
B	B# configure terminal B(config)# interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)# ip address 192.168.3.1 255.255.255.0 B(config-if-GigabitEthernet 0/1)# bfd interval 200 min_rx 200 multiplier 5 B(config-if-GigabitEthernet 0/1)# exit B(config)# interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)# ip address 192.168.2.2 255.255.255.0 B(config-if-GigabitEthernet 0/2)# exit B(config)# router bgp 200 B(config-router)# neighbor 192.168.3.2 remote-as 300

	B(config-router)# neighbor 192.168.3.2 fall-over bfd B(config-router)# neighbor 192.168.2.1 remote-as 100 B(config-router)# redistribute connect
C	C# configure terminal C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ip address 192.168.1.2 255.255.255.0 C(config-if-GigabitEthernet 0/1)# exit C(config)# interface fastEthernet 0/2 C(config-if-GigabitEthernet 0/2)# ip address 192.168.3.2 255.255.0.0 C(config-if-GigabitEthernet 0/2)# bfd interval 200 min_rx 200 multiplier 5 C(config-if-GigabitEthernet 0/2)# exit C(config)# router bgp 300 C(config-router)# neighbor 192.168.1.1 remote-as 100 C(config-router)# neighbor 192.168.3.1 remote-as 200 C(config-router)# neighbor 192.168.3.1 fall-over bfd C(config-router)# address-family ipv4 unicast C(config-router-af)# bgp fast-reroute C(config-router-af)# redistribute connect
Verification	Run the **show** command to display the information.
C	C# show ip bgp summary BGP router identifier 10.10.10.10, local AS number 300 BGP table version is 12 4 BGP AS-PATH entries 0 BGP Community entries 3 BGP Prefix entries (Maximum-prefix:4294967295) Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd 192.168.1.1 4 100 76 77 12 12 0 00:59:27 3 192.168.3.1 4 200 30 30 12 12 0 00:19:03 3 Total number of neighbors 2

```
C# show ip bgp
BGP table version is 12, local router ID is 10.10.10.10
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
                    S Stale, b - backup entry
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric     LocPrf     Weight Path
*  192.168.1.0      192.168.3.1            0                        0 200 ?
*                   192.168.1.1            0                        0 100 ?
*>                  0.0.0.0                0                    32768     ?
*> 192.168.2.0      192.168.3.1            0                        0 200 ?
*b                  192.168.1.1            0                        0 100 ?
*  192.168.3.0      192.168.3.1            0                        0 200 ?
*                   192.168.1.1            0                        0 100 200 ?
*>                  0.0.0.0                0                    32768     ?

Total number of prefixes 3
C# show ip bgp 192.168.2.0
BGP routing table entry for 192.168.2.0/24
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  192.168.1.1
  200
    192.168.3.1 from 192.168.3.1 (3.3.3.3)
      Origin incomplete, metric 0, localpref 100, valid, external, best
      Last update: Tue Oct  5 00:26:52 1971

  100
    192.168.1.1 from 192.168.1.1 (44.44.44.44)
      Origin incomplete, metric 0, localpref 100, valid, external, backup
      Last update: Mon Oct  4 23:46:28 1971
C# show ip route

Codes:  C - connected, S - static, R - RIP, B - BGP
```

```
            O - OSPF, IA - OSPF inter area
            N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
            E1 - OSPF external type 1, E2 - OSPF external type 2
            i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
            ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C       192.168.1.0/24 is directly connected, GigabitEthernet 1/9
C       192.168.1.2/32 is local host.
B       192.168.2.0/24 [20/0] via 192.168.3.1, 00:21:39
C       192.168.3.0/24 is directly connected, GigabitEthernet 1/11
C       192.168.3.2/32 is local host.
```

Common Errors

- No BFD session is configured for BGP neighbors.

5.4.7 Configuring Local ASs

Configuration Effect

- Smoothly migrate the network configurations of router A from AS 23 to AS 3600.

Notes

N/A

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring the Local AS for a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Verification

- Run the **show** command to display the information.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

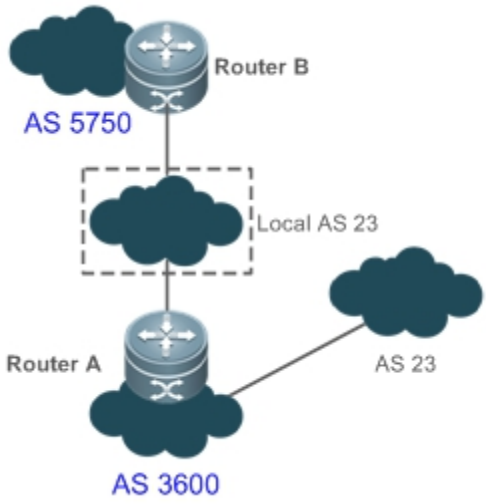
Configuring the Local AS for a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **local-as** *as-number* [**no-prepend** [**replace-as** [**dual-as**]]]
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates a local AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode. **no-prepend**: Does not add the local AS to the AS-PATH in the routing information received by a peer. This option is not available by default. **replace-as**: For the AS-PATH in the routing information sent by a peer, the local AS is used to replace the BGP AS. This option is not available by default. **dual-as**: Enables a peer to use the BGP AS or Local AS to establish a BGP connection with a device. This option is not available by default.
Command Mode	BGP configuration mode
Usage Guide	-

Configuration Example

Configuring BGP Local-AS

Scenario **Figure 5- 10**	

	
Configuration Steps	● Create an EBGP neighborship with B on A and specify the Local-AS for the EBGP neighborship. ● Create an EBGP neighborship for connecting to A on B.
A	A# configure terminal A(config)# router bgp 3600 A(config-router)# neighbor 57.50.1.1 remote-as 5750 A(config-router)# neighbor 57.50.1.1 update-source loopback 0 A(config-router)# neighbor 57.50.1.1 ebgp-multihop 255 A(config-router)# neighbor 57.50.1.1 local-as 23 no-prepend replace-as dual-as
B	B# configure terminal B(config)# router bgp 5750 B(config-router)# neighbor 36.0.1.1 remote-as 23 B(config-router)# neighbor 36.0.1.1 update-source loopback 0 B(config-router)# neighbor 36.0.1.1 ebgp-multihop 255
Verification	Run the **show** command to display the BGP neighbor status.

A	A# show ip bgp neighbors 57.50.1.1 BGP neighbor is 57.50.1.1, remote AS 5750, local AS 23(using Peer's Local AS, no-prepend, replace-as, dual-as), external link BGP version 4, remote router ID 0.0.0.0 BGP state = Idle Last read, hold time is 180, keepalive interval is 60 seconds Received 0 messages, 0 notifications, 0 in queue open message:0 update message:0 keepalive message:0 refresh message:0 dynamic cap:0 notifications:0 Sent 0 messages, 0 notifications, 0 in queue

5.4.8 Configuring BGP GR

Configuration Effect

- Configure BGP GR to implement network deployment with high reliability.

Notes

- To successfully deploy the BGP GR function, you need to use a neighbor device as the GR Helper.
- In an BGP environment, you also need to configure IGP GR.
- After BGP GR is enabled, you need to reset a BGP neighbor connection to make it take effect.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Configuring BGP GR

- Perform this configuration in the BGP configuration mode, which is configured by default.

Configuring a BGP GR Timer

- (Optional) Perform this configuration in the BGP configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Verification

- Run the **show** command to display the neighbor status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Configuring BGP GR

Command	**bgp graceful-restart**
Parameter Description	-
Command Mode	BGP configuration mode
Usage Guide	-

Configuring the BGP GR Restart Timer

Command	**bgp graceful-restart restart-time** *restart-time*
Parameter Description	*restart-time*: Indicates the maximum waiting time that the GR Restarter hopes the GR Helper to wait before a new connection is created, ranging from 1 to 3600 seconds.
Command Mode	BGP configuration mode
Usage Guide	-

Configuring the BGP GR Route Stale Timer

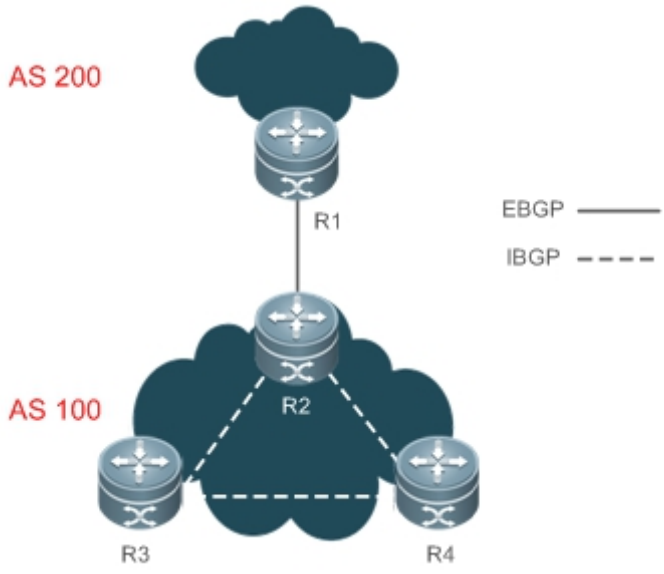
Command	**bgp graceful-restart stalepath-time** *time*
Parameter Description	*time*: Indicates the maximum time that a stale route keeps valid after the connection with a neighbor GR device is recovered, ranging from 1 to 3600 seconds.
Command Mode	BGP configuration mode
Usage Guide	-

Creating a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuration Example

Configuring BGP GR

Scenario **Figure 5- 11**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5- 11. ● Configure a loopback interface on R2, R3, and R4 and create an IBGP neighborship based on the loopback interface. ● Create an EBGP neighborship by using the directly connected interfaces on R1 and R2. ● Enable BGP GR on R1, R2, R3, and R4.
R1	R1# configure terminal R1(config-router)# exit R1(config)# router bgp 100 R1(config-router)# bgp graceful-restart
R2	R2# configure terminal R2(config)# router ospf 1 R2(config-router)# graceful-restart R2(config-router)# exit R2(config)# router bgp 100 R2(config-router)# bgp graceful-restart
R3	R3# configure terminal R3(config)# router ospf 1 R3(config-router)# graceful-restart R3(config-router)# exit R3(config)# router bgp 100 R3(config-router)# bgp graceful-restart

R4	R4# configure terminal R4(config)# router ospf 1 R4(config-router)# graceful-restart R4(config-router)# exit R4(config)# router bgp 100 R4(config-router)# bgp graceful-restart
Verification	Run the **show** command to display the BGP neighbor status.
R2	R2# show ip ospf Routing Process "ospf 1" with ID 10.0.0.2 Process uptime is 4 minutes Process bound to VRF default Conforms to RFC2328, and RFC1583Compatibility flag isenabled Supports only single TOS(TOS0) routes Supports opaque LSA This router is an ASBR (injecting external routing information) SPF schedule delay 5 secs, Hold time between two SPFs 10 secs LsaGroupPacing: 240 secs Number of incomming current DD exchange neighbors 0/5 Number of outgoing current DD exchange neighbors 0/5 Number of external LSA 4. Checksum 0x0278E0 Number of opaque AS LSA 0. Checksum 0x000000 Number of non-default external LSA 4 External LSA database is unlimited. Number of LSA originated 6 Number of LSA received 2 Log Neighbor Adjency Changes : Enabled Graceful-restart enabled Graceful-restart helper support enabled Number of areas attached to this router: 1 Area 0 (BACKBONE) .....

	R2# show ip bgp neighbors BGP neighbor is 192.168.195.183, remote AS 200, local AS 100, external link Using BFD to detect fast fallover - BFD session state up BGP version 4, remote router ID 10.0.0.1 BGP state = Established, up for 00:06:37 Last read 00:06:37, hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Address family IPv4 Unicast: advertised and received Graceful restart: advertised and received Remote Restart timer is 120 seconds Address families preserved by peer: None .....

Common Errors

- GR is not enabled for IGP.
- GR is not enabled for a BGP neighbor device.

5.4.9 Configuring a BGP IPv6 Address Family

Configuration Effect

- Configure BGP IPv6 routes to implement IPv6 network access in different ASs.

Notes

- Generally, BGP uses IPv6 addresses to create neighborships and implement exchange of IPv6 routes.
- In special scenarios (such as the 6PE function, see the MPLS-L3VPN-SCG.doc), BGP supports exchange of IPv6 routes on the neighbors with IPv4 addresses.
- Configurations related to BGP IPv6 services must be configured in the BGP IPv6 address family mode.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring the BGP IPv4 Address Family Mode

- (Optional) Perform this configuration in the BGP configuration mode.

Disabling the IPv4 Address Family Capability for a BGP Neighbor

- (Optional) Perform this configuration in the BGP IPv6 configuration mode.

Configuring the BGP IPv6 Address Family Mode

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring the IPv6 Address Family Capability for a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP IPv6 configuration mode.

Configuring IPv6 Route Advertisement in BGP

- (Optional) Perform this configuration in the BGP IPv6 configuration mode.

Verification

- Run the **show** command to display the neighbor status.
- Run the **show** command to display the routing status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Indicates the address of a peer, which is usually an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring the BGP IPv4 Address Family Mode

Command	**address-family ipv4 unicast**
Parameter Description	-
Command	BGP configuration mode

Mode	
Usage Guide	-

Disabling the IPv4 Address Family Capability for a BGP Neighbor

Command	**no neighbor** { *peer-address* | *peer-group-name* } **activate**
Parameter Description	*peer-address*: Indicates the address of a peer, which is usually an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters.
Command Mode	BGP IPv4 address family mode
Usage Guide	Neighbors with IPv6 addresses are used to exchange IPv6 routes. However, when a neighbor is configured in the BGP mode, BGP automatically activates the IPv4 unicast address family capability for the neighbor. Therefore, you are advised to manually disable the IPv4 unicast address family capability.

Configuring the BGP IPv6 Address Family Mode

Command	**address-family ipv6 unicast**
Parameter Description	-
Command Mode	BGP configuration mode
Usage Guide	-

Configuring the IPv6 Address Family Capability for a BGP Neighbor

Command	**neighbor** { *peer-address* | *peer-group-name* } **activate**
Parameter Description	*peer-address*: Indicates the address of a peer, which is usually an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters.
Command Mode	BGP IPv6 address family mode
Usage Guide	-

Configuring IPv6 Route Advertisement in BGP

Command	**network** *network-number* [**mask** *mask*] [**route-map** *map-tag*] [**backdoor**]
Parameter Description	*network-number*: Indicates the network number. *mask*: Indicates the subnet mask. *map-tag*: Indicates the name of a route map, consisting of no more than 32 characters. **backdoor**: Indicates that the route is a backdoor route.
Command Mode	BGP IPv6 address family mode
Usage Guide	-

Configuration Example

Configuring BGP to Implement IPv6 Route Exchange in Different ASs

Scenario **Figure 5- 12**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5- 12. ● Configure a BGP neighbor, disable the IPv4 address family capability for the neighbor and activate the IPv6 address family capability. ● Configure IPv6 route advertisement in BGP.
A	A# configure terminal A(config)# int loopback 0 A(config-if-Loopback)# ipv6 address 30::1/128 A(config-if-Loopback)# exit A(config)# router bgp 65530 A(config-router)# neighbor 100::1 remote-as 65531 A(config-router)# address-family ipv4 A(config-router-af)# no neighbor 100::1 activate A(config-router-af)# exit-address-family A(config-router)# address-family ipv6 A(config-router-af)# neighbor 100::1 activate A(config-router-af)# network 30::1/128
B	B# configure terminal B(config)# router bgp 65531 B(config-router)# neighbor 100::2 remote-as 65530 B(config-router)# address-family ipv4 B(config-router-af)# no neighbor 100::2 activate B(config-router-af)# exit-address-family B(config-router)# address-family ipv6 B(config-router-af)# neighbor 100::2 activate
Verification	Run the **show** command to display the BGP neighbor status.
A	A# show bgp ipv6 unicast summary BGP router identifier 1.1.1.1, local AS number 65530 BGP table version is 1

	1 BGP AS-PATH entries 0 BGP Community entries 1 BGP Prefix entries (Maximum-prefix:4294967295) Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd 100::1 4 65531 4 6 1 0 0 00:01:49 0 Total number of neighbors 1
B	Run the **show** command to display BGP routing information. B# show bgp ipv6 unicast BGP table version is 4, local router ID is 2.2.2.2 Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale, b - backup entry Origin codes: i - IGP, e - EGP, ? - incomplete Network Next Hop Metric LocPrf Weight Path *> 30::1/128 100::2 0 0 65530 i Total number of prefixes 1

Common Errors

- The IPv6 address family capability is not activated for BGP neighbors.
- In non-6PE scenarios, IPv4 addresses are used to establish IPv6 routes for exchange between neighbors.

5.4.10 Configuring a BGP MDT Address Family

Configuration Effect

- Configure BGP to implement multicast VPN deployment in different ASs.

Notes

- By default, the BGP routing mode is located in the IPv4 unicast address family and a BGP multicast VPN must be configured in the IPv4 MDT address family mode.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in BGP configuration mode.

Activating the MDT Capability for a BGP Neighbor

- (Mandatory) Perform this configuration in BGP IPv4 MDT mode.

Verification

- Run the **show** command to display the information.

Related Commands

Enabling BGP

Command	**router bgp** *as-number* [**instance** *instance-name*]
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode. **instance**: The instance should be specified for the non-default instance, and the instance name is required. *instance-name*: Instance name of 1 to 32 characters.
Command Mode	Global configuration mode
Usage Guide	-

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

Command	**neighbor** { *neighbor-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*neighbor-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Entering the IPv4 MDT Address Family

Command	**address-family ipv4 mdt**
Parameter Description	-
Command Mode	BGP configuration mode

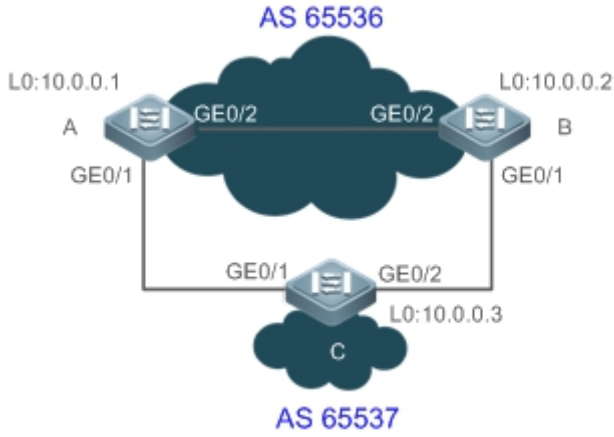
Usage Guide	-

Activating the IPv4 MDT Capability for a BGP Neighbor

Command	**neighbor** { *neighbor-address* | *peer-group-name* } **activate { ipv4 | ipv6 }**
Parameter Description	*neighbor-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group. ipv4: Activate only IPv4 neighbors in the peer group ipv6: Activate only IPv6 neighbors in the peer group
Command Mode	BGP IPv4 MDT address family mode
Usage Guide	-

Configuration Example

Configuring a BGP MDT Address Family

Scenario **Figure 5- 13**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5- 13. ● Configure BGP neighborships as shown in Figure 5- 13. ● Activate the IPv4 MDT address family capability for BGP neighbors.
A	A# config terminal A(config)# ip vrf VRF1 A(config-vrf)# rd 100:1 A(config-vrf)# route-target both 123:123 A(config-vrf)# mdt default 232.1.1.1 A(config-vrf)# exit A(config)# interface gigabitEthernet 0/4 A(config-GigabitEthernet 0/1)# ip vrf forwarding VRF1 A(config-GigabitEthernet 0/1)# ip address 10.1.1.1 255.255.255.0 A(config-GigabitEthernet 0/1)# exit

	A(config)# router bgp 65536 A(config-router)# neighbor 10.0.0.2 remote-as 65536 A(config-router)# neighbor 10.0.0.2 update-source loopback 0 A(config-router)# neighbor 10.0.0.3 remote-as 65537 A(config-router)# address-family ipv4 mdt A(config-router-af)# neighobr 10.0.0.2 activate A(config-router-af)# neighobr 10.0.0.3 activate A(config-router)# address-family vpnv4 A(config-router-af)# neighobr 10.0.0.2 activate A(config-router-af)# neighobr 10.0.0.3 activate A(config-router-af)# exit-address-family A(config-router)# address-family ipv4 vrf VRF1 A(config-router-af)# exit-address-family
B	The same as that for A.
C	The same as that for A.
Verification	Run the **show** command to display the information.
A	A# show ip vrf interfaces Interface IP-Address VRF Protocol GigabitEthernet 0/4 10.1.1.1 VRF1 up A# show bgp ipv4 mdt all BGP table version is 1, local router ID is 10.0.0.1 Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale Origin codes: i - IGP, e - EGP, ? - incomplete Network Next Hop Metric LocPrf Weight Path Route Distinguisher: 100:1 *> 10.0.0.1/32 0.0.0.0 0 32768 ? *>i10.0.0.2/32 10.0.0.2 0 100 ? *> 10.0.0.3/32 10.0.0.3 0 200 ? Total number of prefixes 3

Common Errors

- No VPNv4 address family neighbor is configured.
- No MPLS infrastructure network is deployed.
- The MDT address family capability is not activated for BGP neighbors.

5.4.11 Configuring Interconnection with Devices Supporting Only 2-Byte AS Numbers

Configuration Effect

- Successfully interconnect devices supporting 4-byte AS numbers with devices supporting only 2-byte AS numbers.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in the global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in the BGP configuration mode.

Configuring the Display Mode of a 4-Byte AS Number

- (Optional) Perform this configuration in the BGP configuration mode. By default, a 4-byte AS number is displayed as decimal digits.

Verification

- Run the **show** command to display the neighbor status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

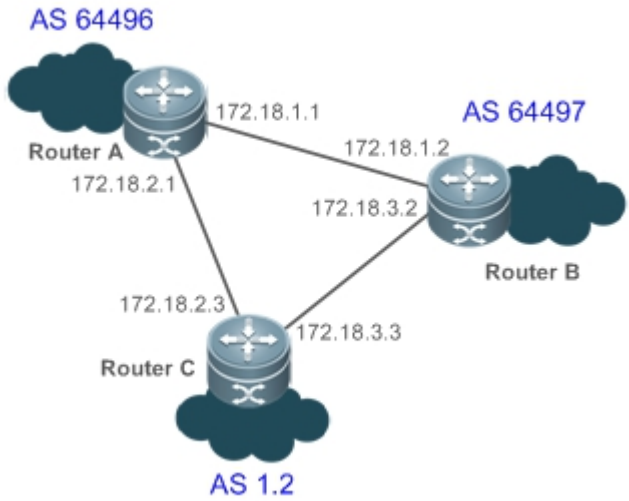
Command	**neighbor** { *peer-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*peer-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group, consisting of no more than 32 characters. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring the Display Mode of a BGP 4-Byte AS Number

Command	**bgp asnotation dot**
Parameter Description	-
Command Mode	BGP configuration mode
Usage Guide	-

Configuration Example

Configuring Compatibility Between BGP Devices Supporting 4-Byte AS Numbers and 2-Byte AS Numbers

Scenario **Figure 5- 14**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in Figure 5- 14. ● Configure BGP neighborships.
A	A# configure terminal A(config)# router bgp 64496 A(config-router)# neighbor 172.18.1.2 remote-as 64497 A(config-router)# neighbor 172.18.2.3 remote-as 23456
B	B# configure terminal B(config)# router bgp 64497 B(config-router)# neighbor 172.18.1.1 remote-as 64496 B(config-router)# neighbor 172.18.3.3 remote-as 1.2 B(config-router)# bgp asnotation dot B(config-router)# end
C	C# configure terminal C(config)# router bgp 1.2

	C(config-router)# neighbor 172.18.2.1 remote-as 64496 C(config-router)# neighbor 172.18.3.2 remote-as 64497
Verification	Run the **show** command to display the BGP neighbor status.
A	A# show ip bgp summary BGP router identifier 172.18.1.1, local AS number 64496 BGP table version is 1, main routing table version 1 Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down Statd 172.18.1.2 4 64497 7 7 1 0 0 00:03:04 0 172.18.2.3 4 23456 4 4 1 0 0 00:00:15 0
B	B# show ip bgp summary BGP router identifier 172.18.3.2, local AS number 64497 BGP table version is 1, main routing table version 1 Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down Statd 172.18.1.1 4 64496 7 7 1 0 0 00:00:04 0 172.18.3.2 4 1.2 4 4 1 0 0 00:00:16 0

Common Errors

N/A

5.4.12 Configuring BGP Tracking

Configuration Effect

- Configure the BGP tracking function to ensure fast route convergence.

Notes

N/A

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in BGP configuration mode.

Configuring the Tracking Function for BGP Neighbors

- (Mandatory) Perform this configuration in BGP configuration mode.

Verification

- Run the **show** command to display the information.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.
Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

Command	**neighbor** { *neighbor-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*neighbor-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring the Tracking Function for BGP Neighbors

Command	**neighbor** { *neighbor-address* | *peer-group-name* } **track** *track-obj-number*
Parameter Description	*neighbor-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group. *track-obj-number*: Specifies the number of the tracked object.
Command Mode	BGP configuration mode
Usage Guide	-

Configuration Example

Configuring BGP Tracking

Scenario	

Figure 5-15	
Configuration Steps	● Establish an EBGP neighborship between routers A and B. ● Configure the loopback interface (2.2.2.2/32) on A for tracking B.
A	A# configure terminal A(config)# ip rns 1 A(config-ip-rns)#icmp-echo 2.2.2.2 A(config-ip-rns-icmp-echo)#timeout 6000 A(config-ip-rns-icmp-echo)#frequency 10000 A(config-ip-rns-icmp-echo)#exit A(config)# ip rns schedule 1 start-time now life forever A(config)# track 3 rns 1 A(config)# router bgp 3600 A(config-router)# neighbor 192.168.182.34 remote-as 65531 A(config-router)# neighbor 192.168.182.34 track 3
Verification	Run the **show** command to query the BGP neighbor status.
A	A# BGP neighbor is 192.168.182.34, remote AS 65531, local AS 65530, external link Using TRACK to detect state BGP version 4, remote router ID 88.5.5.5 BGP state = Established, up for 00:00:16 Last read , hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Four-octets ASN Capability: advertised and received Address family IPv4 Unicast: advertised and received Address family L2VPN VPLS: advertised and received Address family L2VPN EVPN: advertised and received Graceful Restart Capability: advertised and received Remote Restart timer is 120 seconds Address families preserved by peer:

	IPv4 Unicast (was not preserved) IPv4 Labeled (was not preserved) VPNv4 Unicast (was not preserved) IPv6 Unicast (was not preserved) IPv6 Labeled (was not preserved) L2VPN VPLS (was not preserved) L2VPN EVPN (was not preserved) L2VPN VPWS (was not preserved) ……

Common Errors

N/A

5.4.13 Configuring Outbound Loop Detection for a BGP Neighbor

Configuration Effect

- Configure outbound loop detection for a BGP neighbor

Notes

- This feature is available only to EBGP neighbors.

Configuration Steps

Enabling BGP

- (Mandatory) Perform this configuration in global configuration mode.

Creating a BGP Neighbor

- (Mandatory) Perform this configuration in BGP configuration mode.

Configuring Outbound Loop Detection for a BGP Neighbor

- (Mandatory) Perform this configuration in BGP configuration mode.

Verification

- Run the **show** command to display the neighbor status.

Related Commands

Enabling BGP

Command	**router bgp** *as-number*
Parameter Description	*as-number*: Indicates an AS number, ranging from 1 to 4,294,967,295, which is 1 to 65535.65535 in the dot mode.

Command Mode	Global configuration mode
Usage Guide	-

Creating a BGP Neighbor

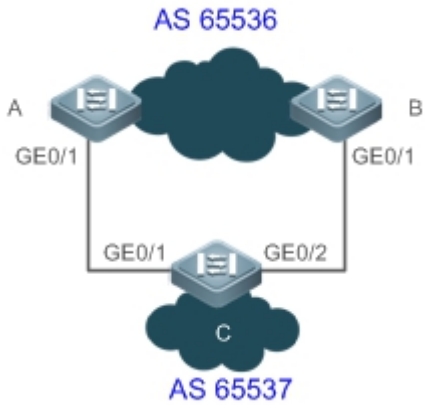
Command	**neighbor** { *neighbor-address* | *peer-group-name* } **remote-as** *as-number*
Parameter Description	*neighbor-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group. *as-number*: Indicates the AS number of a BGP peer (group).
Command Mode	BGP configuration mode
Usage Guide	The AS number specified for a peer (group) must be the same as the BGP AS number of a BGP speaker at the peer end.

Configuring Outbound Loop Detection for a BGP Neighbor

Command	**neighbor** { *neighbor-address* | *peer-group-name* } **as-loop-check out**
Parameter Description	*neighbor-address*: Specifies the address of a peer. This address may be an IPv4 address or an IPv6 address. *peer-group-name*: Specifies the name of a peer group.
Command Mode	BGP configuration mode
Usage Guide	N/A

Configuration Example

Configuring Outbound Loop Detection for a BGP Neighbor

Scenario **Figure 5-16**	
Configuration Steps	● Enable BGP on all devices and set the AS numbers as shown in the preceding figure. ● Establish the EBGP neighborship between Device A and Device C, and between Device B and Device C. ● On Device C, enable outbound loop detection for its neighbors: Device A and Device B.

A	A# configure terminal A(config)# router bgp 65536 A(config-router)# neighbor 10.1.1.1 remote-as 65537
B	B# configure terminal B(config)# router bgp 65536 B(config-router)# neighbor 20.1.1.1 remote-as 65537
C	C# configure terminal C(config)# router bgp 65537 C(config-router)# neighbor 10.1.1.2 remote-as 65536 C(config-router)# neighbor 10.1.1.2 as-loop-check out C(config-router)# neighbor 20.1.1.2 remote-as 65536 C(config-router)# neighbor 20.1.1.2 as-loop-check out
Verification	Run the **show** command to display the BGP neighbor status.
C	C# show ip bgp neighbors 10.1.1.2 BGP neighbor is 10.1.1.2, remote AS 65536, local AS 65537, external link Using as path loop detection in announcing route BGP version 4, remote router ID 10.0.0.1 BGP state = Established, up for 00:06:37 Last read 00:06:37, hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received (old and new) Address family IPv4 Unicast: advertised and received Graceful restart: advertised and received Remote Restart timer is 120 seconds Address families preserved by peer: None …

5.4.14 Configuring Inter-VRF Multi-Path Route Import

Configuration Effect

- Ensure inter-VRF route import and ECMP multi-path for inter-imported routes.

Notes

- N/A

Configuration Steps

Configuring a BGP VRF Address Family

- (Mandatory) Perform this configuration in BGP configuration mode.

Importing Static Routes to BGP

- (Mandatory) Perform this configuration in BGP address family mode.

Importing Multi-Path Static Routes to BGP

- (Mandatory) Perform this configuration in BGP address family mode.

Configuring BGP ECMP

- (Mandatory) Perform this configuration in BGP address family mode.

Configuring Inter-VRF Import for All Path Routes

- (Mandatory) Perform this configuration in BGP address family mode.

Verification

- Run the **show** command to display route information.

Related Commands

Configuring a BGP VRF Address Family

Command	**address-family ipv4 vrf** *vrf-name*
Parameter Description	*vrf-name:* Indicates the VRF instance name.
Command Mode	BGP configuration mode
Usage Guide	N/A

Importing Static Routes to BGP

Command	**redistribute** *protocol-type* [**route-map** *map-tag*] [**metric** *metric-value*]
Parameter Description	*protocol-type*: Indicates the source protocol type of a redistributed route. **route-map** *map-tag*: Indicates the name of an associated route map. **metric** *metric-value*: Indicates the default metric value of a redistributed route. The value range is from **0** to **4,294,967,295**.
Command Mode	BGP address family mode
Usage Guide	N/A

Importing Multi-Path Static Routes to BGP

Command	**bgp sourced-paths** *protocol-type* **all**
Parameter Description	*protocol-type*: Indicates the source protocol type of a redistributed route.
Command Mode	BGP address family mode
Usage Guide	This command needs to be used together with the redistribution command to import routes with multiple next hops from other protocols to BGP.

Configuring BGP ECMP

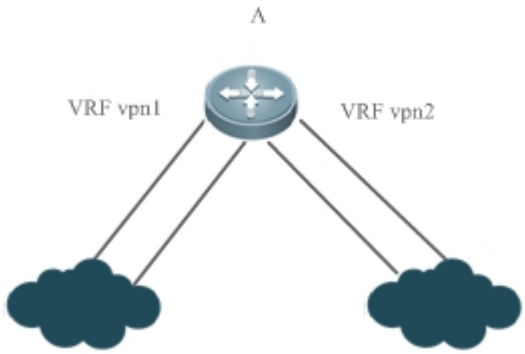
Command	**maximum-paths** { **ebgp** | **ibgp** } *number*
Parameter Description	**ebgp:** Specifies the number of equivalent paths of the EBGP multipath load balancing function. **ibgp:** Specifies the number of equivalent paths of the IBGP multipath load balancing function. *number:* Indicates the maximum number of equivalent paths. The minimum value is **1**, and the maximum value depends on the device capability. If the value is **1**, the EBGP multipath load balancing function is disabled.
Command Mode	BGP address family mode
Usage Guide	The **maximum-paths ebgp** command is also used to configure equivalence of confederation EBGP multiple paths and local inter-VRF import routes. IBGP and EBGP routes cannot form equivalent routes.

Configuring Inter-VRF Import for All Path Routes

Command	**import path selection** { **all** | **bestpath** | **multipath** }
Parameter Description	**all:** Imports all routes with next hops. **bestpath:** Imports routes with preferred next hops. By default, only routes with preferred next hops are imported. **multipath:** Imports routes with preferred and equivalent next hops.
Command Mode	BGP address family mode
Usage Guide	This command can be used to control inter-VRF route import, L3VPN remote routes import to VRF, and EVPN routes import to the IP route table.

Configuration Example

Configuring BGP Multi-Path Bypass Protection

Scenario **Figure 5-17**	 Device A connects to two networks through VRF vpn1 and vpn2, and cross-VRF access is implemented through device A.
Configuration Steps	Configure VRF. Configure VRF static routes. Configure a VRF address family. Import VRF static routes to BGP. Enable multi-path static route import to BGP. Configure BGP ECMP. Configure inter-VRF import for all path routes.
A	A# conf terminal A(config)# ip vrf vpn1 A(config-vrf)# rd 200:1 A(config-vrf)# route-target both 100:100 A(config-vrf)# exit A(config)# ip vrf vpn2 A(config-vrf)# rd 300:1 A(config-vrf)# route-target both 100:100 A(config-vrf)# exit A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# A(config-if-GigabitEthernet 0/1)# ip vrf forwarding vpn1 A(config-if-GigabitEthernet 0/1)# ip address 44.1.1.2 255.255.255.0 A(config-if-GigabitEthernet 0/1)# exit A(config)# interface GigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)# ip vrf forwarding vpn1 A(config-if-GigabitEthernet 0/2)# ip address 45.1.1.2 255.255.255.0 A(config-if-GigabitEthernet 0/2)# exit

	A(config)# interface GigabitEthernet 0/3 A(config-if-GigabitEthernet 0/3)# ip vrf forwarding vpn2 A(config-if-GigabitEthernet 0/3)# ip address 46.1.1.2 255.255.255.0 A(config-if-GigabitEthernet 0/3)# exit A(config)# interface GigabitEthernet 0/4 A(config-if-GigabitEthernet 0/4)# ip vrf forwarding vpn2 A(config-if-GigabitEthernet 0/4)# ip address 47.1.1.2 255.255.255.0 A(config-if-GigabitEthernet 0/4)# exit A(config)# ip route vrf vpn1 100.1.1.1 255.255.255.255 44.1.1.1 A(config)# ip route vrf vpn1 100.1.1.1 255.255.255.255 45.1.1.1 A(config)# ip route vrf vpn2 200.1.1.1 255.255.255.255 46.1.1.1 A(config)# ip route vrf vpn2 200.1.1.1 255.255.255.255 47.1.1.1 A(config)# router bgp 100 A(config-router)# address-family ipv4 vrf vpn1 A(config-router-af)# redistribute static A(config-router-af)# maximum-paths ebgp 32 A(config-router-af)# bgp sourced-paths static all A(config-router-af)# import path selection all A(config-router-af)# exit-address-family A(config-router)# address-family ipv4 vrf vpn2 A(config-router-af)# redistribute static A(config-router-af)# maximum-paths ebgp 32 A(config-router-af)# bgp sourced-paths static all A(config-router-af)# import path selection all A(config-router-af)# exit-address-family
Verification	Run the **show** command to display the configurations.
A	A#show ip route vrf vpn1 Routing Table: vpn1 Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2

```
            SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
            IA - Inter area, EV - BGP EVPN, A - Arp to host
            * - candidate default

Gateway of last resort is no set
C       44.1.1.0/24 is directly connected, GigabitEthernet 0/1
C       44.1.1.2/32 is local host.
C       45.1.1.0/24 is directly connected, GigabitEthernet 0/2
C       45.1.1.2/32 is local host.
S       100.1.1.1/32 [1/0] via 44.1.1.1
                           [1/0] via 45.1.1.1
B       200.1.1.1/32 [20/0] via 47.1.1.1, 02:32:01
                           [20/0] via 46.1.1.1, 02:32:01

A#show ip route vrf vpn2
Routing Table: vpn2

Codes:  C - Connected, L - Local, S - Static
            R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route
            N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
            E1 - OSPF external type 1, E2 - OSPF external type 2
            SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
            IA - Inter area, EV - BGP EVPN, A - Arp to host
            * - candidate default

Gateway of last resort is no set
C       46.1.1.0/24 is directly connected, GigabitEthernet 0/3
C       46.1.1.2/32 is local host.
C       47.1.1.0/24 is directly connected, GigabitEthernet 0/4
C       47.1.1.2/32 is local host.
B       100.1.1.1/32 [20/0] via 45.1.1.1, 03:27:07
                           [20/0] via 44.1.1.1, 03:27:07
S       200.1.1.1/32 [1/0] via 46.1.1.1
                           [1/0] via 47.1.1.1
```

5.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears BGP IPv4 unicast routes.	**clear ip bgp** [**vrf** *vrf-name*] { * | *as-number* | *peer-address* } [**soft**] [**in** | **out**] **clear bgp ipv4 unicast** [**vrf** *vrf-name*] { * | *as-number* | *peer-address* } [**soft**] [**in** | **out**] **clear ip bgp** [**vrf** *vrf-name*] **update-group** [*update-group-index* | *peer-address*] [**soft**] [**in** | **out**] **clear bgp ipv4 unicast** [**vrf** *vrf-name*] **update-group** [*update-group-index* | *peer-address*] [**soft**] [**in** | **out**]
Clears BGP IPv4 MDT routes.	**clear bgp** [**instance** *as-number*] **ipv4 mdt** { * | *as-number* | *neighbor-address* } **clear bgp ipv4 mdt** { * | *as-number* | *neighbor-address* }
Clears BGP IPv6 unicast routes.	**clear bgp ipv6 unicast** [**vrf** *vrf-name*] { * | *as-number* | *peer-address* } [**soft**] [**in** | **out**] **clear bgp ipv6 unicast** [**vrf** *vrf-name*] **update-group** [*update-group-index* | *peer-address*] [**soft**] [**in** | **out**]
Clears BGP L2VPN EVPN routes.	**clear bgp l2vpn evpn** { * | *as-number* | *neighbor-address*} [**soft**] [**in** | **out**] **clear bgp l2vpn evpn update-group** [*update-group-index* | *neighbor-address*] [**soft**] [**in** | **out**]
Clears EVPN conflict MAC	**clear evpn conflict mac** [*vni-id*]

Displaying

Description	Command
Displays BGP IPv4 unicast routes.	**show ip bgp** **show bgp ipv4 unicast**
Displays the update-group information of BGP IPv4 unicast address family.	**show ip bgp** [**vrf** *vrf-name*] **update-group** [*neighbor-address* | *update-group-index*] [**summary**] **show bgp ipv4 unicast** [**vrf** *vrf-name*] **update-group** [*neighbor-address* | *update-group-index*] [**summary**]
Displays BGP IPv4 MDT routes.	**show bgp ipv4 mdt**
Displays BGP IPv6 unicast routes.	**show bgp ipv6 unicast**
Displays the update-group information of BGP IPv6 unicast address family.	**show bgp ipv6 unicast** [**vrf** *vrf-name*] **update-group** [*neighbor-address* | *update-group-index*] [**summary**]
Displays BGP L2VPN EVPN routes.	**show bgp l2vpn evpn all**
Displays the update-group information of BGP L2VPN EVPN address family.	**show bgp l2vpn evpn all update-group** [*neighbor-address* | *update-group-index*] [**summary**]
Displays the MAC mobility or conflict.	**show evpn mac** {**conflict** | **mobility** } [*vni-id*]
Displays BGP statistics.	**show bgp statistics** [**vrf** *vrf-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Enables all BGP debugging.	**debug ip bgp all**
Debugs BGP route flapping.	**debug ip bgp dampening**
Debugs BGP event processing.	**debug ip bgp event**
Debugs BGP route filtering.	**debug ip bgp filter**
Debugs BGP status machine.	**debug ip bgp fsm**
Debugs BGP neighbor keepalive.	**debug ip bgp keepalives**
Debugs BGP core route processing.	**debug ip bgp nsm**
Debugs BGP UPDATE packets.	**debug ip bgp update**
Debugs BGP EVPN.	**debug ip bgp evpn**
Debugs BGP TRACK exchanging.	**debug ip bgp track**
Debugs BGP UPDATE-GROUP.	**debug ip bgp update-group**

6 Configuring PBR

6.1 Overview

Policy-based routing (PBR) is implemented by applying a route map including policies to interfaces and devices.

Similar to static routing, PBR is also manually configured and cannot automatically update with network changes. In addition, PBR is effective only for packets sent from local interfaces and devices. As compared with static and dynamic routing, PBR is more flexible. Static and dynamic routing can forward packets only based on destination addresses. PBR can forward packets based on source and destination addresses, packet length and input interface.

6.2 Applications

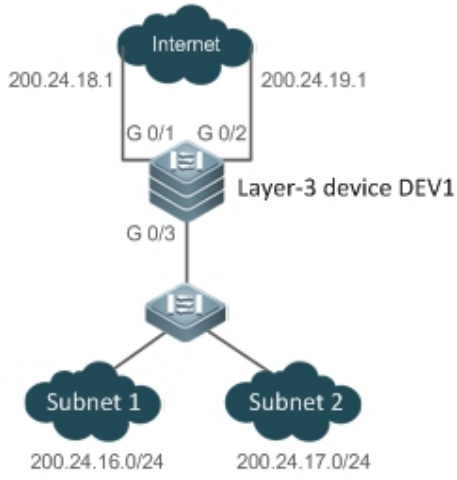
Application	Description
Selecting an ISP by Using PBR	Specify preferential output interfaces for packets from different subnets.
Implementing Traffic Classification by Using PBR	Specify QoS values for packets from different subnets.

6.2.1 Selecting an ISP by Using PBR

An existing user network often uses resources of multiple internet server providers (ISPs). PBR needs to be used since different bandwidths may be requested from different ISPs or the network resources for key users need to be protected. By controlling forwarding of certain data packets, you can make full use ISP resources as well as meet the requirements of flexible and diversified applications.

Scenario

Figure 6- 1



A LAN has two output interfaces for connecting the Internet. PBR is configured on the layer-3 device DEV1 to enable the two output interfaces to implement load sharing and mutual backup.

The specific requirements are as follows:

- Data streams from subnet 1 are sent from GE 0/1.
- Data streams from subnet 2 are sent from GE 0/2.

- If the GE 0/1 link is disconnected, the data streams on GE 0/1 are switched to GE 0/2. Vice versa.

Deployment

- Configure two different ACLs on the layer-3 device DEV1:

ACL1: source addresses belong to subnet 1.

ACL2: source addresses belong to subnet 2.

- Configure two policies in the route map on the layer-3 device DEV1:

Policy 1: sets the next hops for packets matching ACL1 to GE0/1 and GE0/2 (Based on the configuration sequence, GE0/1 takes effect first and GE0/2 works in the backup mode).

Policy 2: sets the next hops for packets matching ACL2 to GE0/2 and GE0/1 (Based on the configuration sequence, GE0/2 takes effect first and GE0/1 works in the backup mode).

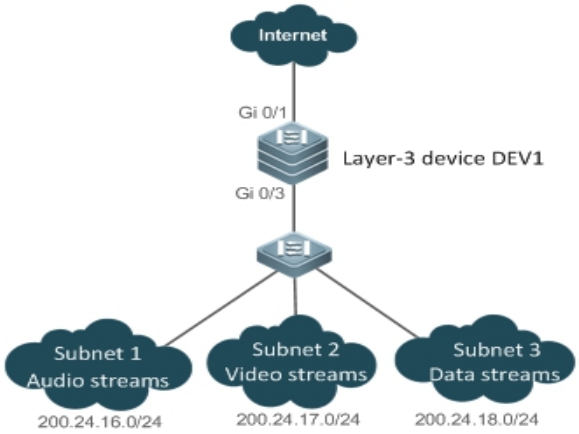
- Configure PBR on GE0/3 (by using a route map). Then, packets received on this interface are forwarded based on the policies.

6.2.2 Implementing Traffic Classification by Using PBR

Scenario

Networks of medium- and small-sized enterprises have simple structures. Different branch nodes are interconnected to the central nodes through carrier dedicated lines or the Internet VPN mode. Enterprise networks often need to implement three-in-on integration (of audio, video and data) to maximize the utilization of existing IP networks and save costs. Since all traffic is output from a single output interface, it is necessary to adjust the QoS policies for the output interface, in order to provide preferential communication quality for bandwidth- and delay-sensitive applications.

Figure 6- 2



A LAN has an output interface for connecting the Internet. PBR is configured on the layer-3 device DEV1 to change the QoS values for packets from different networks.

The specific requirements are as follows:

- For data streams from subnet 1, representing audio streams, set the DSCP value to 56.
- For data streams from subnet 2, representing video streams, set the DSCP value to 40.
- For data streams from subnet 3, representing data streams, set the DSCP value to 24.

Deployment

- Configure three different ACLs on the layer-3 device DEV1:

ACL1: source addresses belong to subnet 1.

ACL2: source addresses belong to subnet 2.

ACL3: source addresses belong to subnet 3.

- Configure three policies in the route map on the layer-3 device DEV1:

Policy 1: sets the DSCP value for packets matching ACL1 to 56.

Policy 2: sets the DSCP value for packets matching ACL2 to 40.

Policy 3: sets the DSCP value for packets matching ACL3 to 24.

- Configure PBR on GE0/3 (by using a route map). Then, the DSCP values for packets received on this interface are changed based on the policies.

6.3 Features

Feature	Description
Configuring a Policy	Before configuring PBR, configure policies in a route map.
Configuring PBR	Apply a route map including policies to interfaces and devices to implement PBR.

6.3.1 Configuring a Policy

A policy is a "match …, set…" statement, which indicates that "if certain conditions are matched, perform certain processing actions".

For detailed introduction to the policies, see the section "Route Map".

Executing Policies

In the global configuration mode, you can run the **route-map** *route-map-name* [**permit** | **deny**] [*sequence-number*] command to create a policy in a route map.

A route map may contain multiple policies. Each policy has a corresponding sequence number. A smaller sequence number means a higher priority. Policies are executed based on their sequence numbers. Once the matching condition of a policy is met, the processing action for this policy needs to be executed and the route map exits. If no matching condition of any policy is met, no processing action will be performed.

Policies have two working modes:

- permit: When the matching condition of a policy is met, perform the processing action for this policy and exit the route map.
- deny: When the matching condition of a policy is met, do not perform the processing action for this policy and exit the route map.

Matching conditions of policies

The matching conditions of a policy may contain 0, 1 or more matching rules.

- If 0 matching rule is contained, no packet will be matched.
- If one or more match rules are contained, all match rules must be matched at the same time to meet the matching conditions of the policy.

In the route map mode, run the **match** command to configure match rules. One **match** command is mapped to one match rule.

PBR supports the following **match** commands:

	Command	Description
IPv4 PBR	**match length**	The IPv4 packet length is used as the matching condition. Only one **match length** command can be configured in a policy.
	match ip address	The source IPv4 address (and the destination IPv4 address) is used as the matching condition. Multiple **match ip address** commands can be configured in a policy.
	match ip policy	The source IPv4 address (and the destination IPv4 address) and layer-3 authentication traffic redirection domain type are used as the matching conditions. Only one **match ip policy** command can be configured in a policy.
IPv6 PBR	**match ipv6 address**	The source IPv6 address (and the destination IPv6 address) is used as the matching condition. Only one **match ipv6 policy** command can be configured in a policy.

IPv4 PBR defines the source IP address (and destination IP address) ranges of packets by using the IP standard or extended ACLs. IPv6 PBR defines the source IPv6 address (and destination IPv6 address) ranges of packets by using the IPv6 extended ACLs.

On a switch, packet forwarding based on policies of IPv4 PBR interfaces supports expert-level and MAC name ACLs. Packet forwarding based on local policies does not support expert-level and MAC name ACLs.

When PBR uses an ACL that is unavailable, the route sub-map will not be matched and the next route sub-map will be matched instead. If no route sub-map is matched, a common route will be selected for forwarding. If only ACLs are configured but no ACE is configured, the PBR forwarding behavior is the same as that in a scenario where an ACL is unavailable.

On a non-switch device, packet forwarding based on policies of IPv4 PBR interfaces and local policies do not support expert-level and MAC name ACLs.

On a switch, if a route sub-map uses multiple ACLs in PBR, only the first ACL is matched.

Processing action for a policy

The processing action of a policy may contain 0, 1 or more set rules.

- If 0 set rule is contained, no processing action will be performed and the route map will directly exit.
- If one or more set rules are contained, all processing actions will be performed and the route map will exit.

If set rules have different priorities, the set rule with the highest priority will take effect.

In the route map mode, run the **set** command to configure set rules. One **set** command is mapped to one set rule.

PBR supports the following **set** commands:

	Command	Description
IPv4 PBR	**set ip tos**	Modifies the **tos** field of an IPv4 packet. This command cannot work with the **set ip dscp** command.
	set ip precedence	Modifies the **precedence** field of an IPv4 packet. This command cannot work with the **set ip dscp** command.

	Command	Description
	set ip dscp	Modifies the **dscp** field of an IPv4 packet. This command cannot work with the **set ip tos** and **set ip precedence** commands.
	set vrf	Sends IPv4 packets to a VRF for forwarding. Select routes for packets matching the match rules by using a VRF specified by **set vrf**, no matter whether the interface that receives the packets belongs to the VRF. This command cannot work with the **set interface** and **set default interface** commands.
	set ip next-hop	Configures the next hop of IPv4 packet forwarding. The next hop must be directly connected; otherwise, this command is invalid. A packet matching the match rules will be forwarded to the next hop specified by **set ip next-hop first**, no matter whether the route selected for the packet in the routing table is consistent with the next hop specified by PBR. On a switch, the output interfaces for next hops supported by PBR include the SVI, routing and layer-3 AP interfaces.
	set ip next-hop recursive	Configures the recursive next hop of IPv4 packet forwarding. The next hop can be directly connected or not directly connected. A non-directly-connected next hop will recur to a static or dynamic route in the routing table. This command supports recursion to multiple ECMP next hops of a static or dynamic route. A maximum of 32 next hops are supported. If a recursive route is a static route, only one next hop is supported for the static recursive route. The redundant backup or load balancing mode of multiple recursive next hops is also determined by the **ip policy** { **redundance** | **load-balance** } command. A packet matching the match rules will be forwarded to the recursive next hop specified by **set ip next-hop recursive** first, no matter whether the route selected for the packet in the routing table is consistent with the next hop specified by PBR. Only when a static or dynamic route has an output interface and a next-hop IP address, the policy-based recursive next hop can take effect.
	set interface	Configures the output interface of IPv4 packet forwarding. A packet matching the match rules will be forwarded from the interface specified by **set interface** first, no matter whether the route selected for the packet in the routing table is consistent with the output interface specified by PBR. This command cannot work with the **set vrf** command.
	set ip default next-hop	Configures the default next hop of IPv4 packet forwarding. A packet matching the match rules will be forwarded to the default next hop specified by this command if a route fails to be selected or the default route is selected for this packet in the routing table.

	Command	Description
	set ip default interface	Configures the default output interface of IPv4 packet forwarding. A packet matching the match rules will be forwarded from the interface specified by this command if a route fails to be selected or the default route is selected for this packet in the routing table. This command cannot work with the **set vrf** command.
	set ip policy l3-auth	Configures layer-3 authentication for Ipv4 packets. Layer-3 authentication will be enabled for packets matching the match rules. This command is effective only for packets forwarded by an interface, but not for locally initiated packets.
	set ip policy load-balance	Configures the load balancing mode for IPv4 packets. A packet matching the match rules will select an output interface based on the configured load balancing mode if the load balancing mode is enabled globally for PBR. This command is effective only for packets forwarded by an interface, but not for locally initiated packets.
	set ip policy no-ttl-decrease	Configures no decrease by 1 for the TTL field of IPv4 packets. The value of the TTL field will not be decreased by 1 at the header of an IPv4 packet matching the match rules when the packet is forwarded based on policies. This command is mainly used for traffic redirection in layer-3 authentication. This command is effective only for packets forwarded by an interface, but not for locally initiated packets.
IPv6 PBR	**set ipv6 precedence**	Modifies the **precedence** field of an IPv6 packet. IPv6 PBR does not support **set ipv6 tos** or **set ipv6 dscp**.
	set ipv6 next-hop	Configures the next hop of IPv6 packet forwarding. An IPv6 packet matching the match rules will be forwarded to the next hop specified by **set ipv6 next-hop** first, no matter whether the route selected for the IPv6 packet in the routing table is consistent with the next hop specified by PBR. The next hop must be directly connected; otherwise, this command is invalid.
	set ipv6 default next-hop	Configures the default next hop of IPv6 packet forwarding. An IPv6 packet matching the match rules will be forwarded to the default next hop specified by this command if a route fails to be selected or the default route is selected for this packet in the routing table. The next hop must be directly connected; otherwise, this command is invalid.

The priority sequence is as follows: **set ip next-hop** > **set ip next-hop recursive** > **set interface** > common route > **set ip default next-hop** > **set default interface** > default route. The preceding **set** commands can be configured at the same time but only the command with the highest priority takes effect.

The priority sequence is as follows: **set ipv6 next-hop** > common route > **set ipv6 default next-hop** > default route**.** The preceding **set** commands can be configured at the same time but only the command with the highest priority takes effect.

For switches, the **set ipv6 default next-hop** command does not take effect for IPv6 addresses whose mask length exceeds 64.

6.3.2 Configuring PBR

PBR

Apply a route map including policies to interfaces or devices to implement PBR.

- Apply a route map to an interface so that packets received by the interface are routed based on the policy.

The PBR is often used to control user packets received by a device. This command is effective only for forwarded packets, but not for locally initiated packets.

- Apply a route map to a device so that packets locally initiated are routed based on the policy.

The PBR is often used to control protocol packets exchanged between devices (such as ping packets sent locally). This command is effective only for locally initiated packets, but not for forwarded packets.

By default, PBR is not unavailable on a device and packets are forwarded based on a routing table.

On a switch, the interfaces which support PBR are L3 Ethernet interface, SVI interface and L3 AP interface.

Redundant backup or load balancing

You can set multiple next hops in a policy. Either redundant backup or load balancing can be implemented among multiple next hops. Redundant backup is implemented by default.

Redundant backup or load balancing is only effective for next hops configured in the **set ip next-hop**, **set ip next-hop recursive**, **set ip default next-hop**, **set ipv6 next-hop** and **set ipv6 default next-hop** commands, and only effective among multiple next hops in the same set rule.

- Redundant backup

Based on the configuration sequence, the first accessible next hop takes effect. When the currently effective next hop (R1) is faulty, the traffic automatically switches to the next accessible next hop (R2). When R1 becomes accessible again, the traffic automatically switches back to R1.

A newly added next hop is arranged at the last of the sequence. Assume that the original sequence of multiple next hops is R1 > R2 > R3. After R1 is deleted and added again, the sequence changes to R2 > R3 > R1.

If no next hop is accessible, packets will be discarded.

- Load balancing

When multiple accessible next hops take effect at the same time, the Weighted Cost Multiple Path (WCMP) and Equal Cost Multiple Path (ECMP) are supported. After an accessible next hop loses effect, traffic will be balanced among the other accessible next hops.

Correlation with BFD

Correlation between PBR and BFD is effective only for next hops configured by the **set ip next-hop** or **set ipv6 next-hop** command.

The **set ip next-hop** and **set ipv6 next-hop** commands carry the **verify-availability** and **bfd** [**vrf** *vrf-name*] *interface-type interface-number gateway* parameters, which can establish correlation between PBR and a BFD session and monitor the accessibility of next hops.

Correlation between PBR and BFD helps enhance the PBR's perception about network environment changes. When BFD detects that the current next hop is not accessible, the BFD will immediately notify the PBR to switch the traffic to another accessible next hop (to implement redundant backup) or all the other accessible next hops (to implement load balancing).

For the configuration and related commands for correlation between PBR and BFD, see the "BFD" section.

Correlation with Track

Correlation between PBR and Track is effective only for next hops configured by the **set ip next-hop** command.

The **set ip next-hop** command carries the **verify-availability** and **track** *track-obj-number* parameters, which can establish correlation between PBR and a Track session and monitor the accessibility of next hops.

Correlation between PBR and Track helps enhance the PBR's perception about network environment changes. When Track detects that the current next hop is not accessible, the Track will immediately notify the PBR to switch the traffic to another accessible next hop (to implement redundant backup) or all the other accessible next hops (to implement load balancing).

Only IPv4 PBR supports correlation with Track.

For the configuration and related commands for correlation between PBR and Track, see the "RNS" section.

VRF transfer

If this feature is selected for VRF based on PBR, an interface to which PBR is applied can filter received IP packets by using the match rules. If the packets are successfully matched, the interface will specify a VRF instance for route selection in the set rules. The match rules include the packet length and ACL (IP access list). Since the match rules are flexible, you can allocate different traffic to different VRF instances based on actual requirements.

Generally, packets received on a VRF interface will be forwarded from this VRF interface, and packets received on a global interface will be forwarded based on a global routing table. PBR can break this limit and enable packets to be transferred between VRF and a global route map. The specific information is as follows:

- From a global routing table to VRF: Packets received from a global interface are transferred to a specified VRF instance for forwarding.
- From a VRF instance to another VRF: instance: Packets received from a VRF interface are transferred to another VRF interface for forwarding.
- From VRF to a global routing table: Packets received from a VRF interface are transferred to the global routing table for forwarding.

Single-protocol VRF enables packets to be transferred only to VRF instances using IPv4 PBR. Multi-protocol VRF enables packets to be transferred to VRF instances using IPv4 and IPv6 PBR.

For VRF configuration and related commands, see the "VRF" section.

Only the following **set** commands enable packets to be transferred between VRFs or global routing tables.

Command	Description
set vrf	Transfers packets from a global routing table to a VRF instance, and then from the VRF instance to another VRF instance.
set ip next-hop	Carries the **vrf** *vrf-name* and **global** parameters. Configures **vrf** *vrf-name* to transfer packets from a global routing table to a VRF instance and from the VRF instance to another VRF instance. Configures **global** to transfer packets from a VRF instance to a global routing table.
set ipv6 next-hop	Carries the **vrf** *vrf-name* and **global** parameters. Configures **vrf** *vrf-name* to transfer packets from a global routing table to a VRF instance and from the VRF instance to another VRF instance. Configures **global** to transfer packets from a VRF instance to a global routing table.

Source-addressed-based PBR

Run the global configuration commands **ip policy-source in-interface** and **ipv6 policy-source in-interface** to directly generate source-address-based PBR. You do not need to configure a route map.

- If only source IPv4 or Ipv6 addresses need to be matched for packets forwarded by an interface, you can apply the command for source-address-based PBR. The procedure for configuring this command is simpler than that for interface-based PBR.

- If source-address-based PBR is applied to a specified interface, packets received on this interface will be routed based on policies according to the source addresses.

The PBR is often used to control user packets received by a device. This command is effective only for forwarded packets, but not for locally initiated packets.

Source-address-based PBR has a higher priority than interface-based PBR. If source-address-based PBR and interface-based PBR are applied to the same interface, only interface-based PBR takes effect.

By default, source-address-based routing is not available on devices and packets are forwarded based on the routing table.

Policy-based traffic redirection in layer-3 authentication

Redirect traffic based on PBR for layer-3 authentication.

- Run the **match ip policy** command to forward packets matching the layer-3 authentication traffic redirection domain type.

- Run the **set ip policy l3-auth** command to select a route for and forward successfully matched packets for layer-3 authentication.

- Run the **set ip policy load-balance** command to set the load balancing mode. To ensure that packets redirected by PBR and corresponding response packets are redirected to the same layer-3 authentication charging card, you need to apply PBR in the input and output directions of a device. In addition, you also need to run the **set ip policy load-balanc**e command to set corresponding load balancing modes to ensure that the load balancing modes in the input and output directions of the device are symmetrical.

- Run the **set ip policy no-ttl-decrease** command to ensure that the value of the TTL field in packets forwarded based on a policy applied to an interface will not be decreased by 1. During policy-based traffic redirection in layer-3 authentication, the value of the TTL field at the IPv4 packet header will be decreased by 1 when packet traffic is redirected to a layer-3 authentication charging card. However, this additional TTL overhead is unnecessary. Therefore, you need to run the **set ip policy no-ttl-decrease** command to ensure that the value of the TTL field of a PBR-based traffic redirection packet will not be decreased by 1.

Only IPv4 packets are supported in policy-based traffic redirection in layer-3 authentication.

6.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of PBR	(Mandatory) It is used to apply PBR to forward packets.	
	ip policy route-map	Applies PBR for IPv4 packets received by an interface.
	ipv6 policy route-map	Applies PBR for IPv6 packets received by an interface.
	ip local policy route-map	Applies PBR for IPv4 packets locally initiated.
	ipv6 local policy route-map	Applies PBR for IPv6 packets locally initiated.
Setting Redundant Backup or Load Balancing	(Optional) It is used to set whether PBR implements redundant backup or load balancing among multiple next hops.	

Configuration	Description and Command	
	ip policy { **redundance** | **load-balance** }	Sets whether IPv4 PBR implements redundant backup or load balancing among multiple next hops. The default setting is redundant backup.
	ipv6 policy { **redundance** | **load-balance** }	Sets whether IPv6 PBR implements redundant backup or load balancing among multiple next hops. The default setting is redundant backup.
Configuring Source-Address-Based PBR	(Optional) It is used to apply source-address-based PBR to forward packets.	
	ip policy-source in-interface	Applies source-address-based PBR for IPv4 packets received by an interface.
	Ipv6 policy-source in-interface	Applies source-address-based PBR for IPv6 packets received by an interface.

6.4.1 Configuring Basic Functions of PBR

Configuration Effect

Perform personalized routing management for user data streams by preparing flexible policies.

Perform personalized management for protocol interaction and network topologies by preparing flexible policies.

Notes

- A route map must be used when PBR is configured; therefore, you must configure a route map on a device.
- If an ACL is used when the route map is configured, you must configure the ACL on the device.

Configuration Steps

Applying PBR for IPv4 packets received by an interface

- To perform personalized routing management for IPv4 user data streams passing a device, you should perform this configuration.
- Perform this configuration on the input interface for user data streams.
- Run the **ip policy route-map** command to apply a route map to an interface. Then, PBR is executed for IPv4 packets received on this interface.

Command	**ip policy route-map** *route-map-name*
Parameter Description	*route-map-name*: Indicates the name of a route map.
Defaults	By default, PBR is unavailable on a device and packets are forwarded based on a routing table.
Command Mode	Interface configuration mode
Usage Guide	Only one **ip policy route-map** command can be configured for an interface. If multiple **ip policy route-map** commands are configured for an interface, only the last configuration takes effect. If the route map used in PBR is unavailable, the PBR does not take effect.

Applying PBR for IPv6 packets received by an interface

- To perform personalized routing management for IPv6 user data streams passing a device, you should perform this configuration.
- Perform this configuration on the input interface for user data streams.
- Run the **ipv6 policy route-map** command to apply a route map to an interface. Then, PBR is executed for IPv6 packets received on this interface.

Command	**ipv6 policy route-map** *route-map-name*
Parameter Description	*route-map-name*: Indicates the name of a route map.
Defaults	By default, PBR is unavailable on a device and packets are forwarded based on a routing table.
Command Mode	Interface configuration mode
Usage Guide	Only one **ipv6 policy route-map** command can be configured for an interface. If multiple **ipv6 policy route-map** commands are configured for an interface, only the last configuration takes effect. If the route map used in PBR is unavailable, the PBR does not take effect.

Applying PBR for IPv4 packets locally initiated

- To perform personalized management for IPv4 protocol interaction and IPv4 network topologies, you should perform this configuration.
- Run the **ip local policy route-map** command to apply a route map to a device. Then, PBR is executed for IPv4 packets locally initiated.

Command	**ip local policy route-map** *route-map-name*
Parameter Description	*route-map-name*: Indicates the name of a route map.
Defaults	By default, PBR is unavailable on a device and packets are forwarded based on a routing table.
Command Mode	Global configuration mode
Usage Guide	Only one **ip local policy route-map** command can be configured for a device. If the route map used in PBR is unavailable, the PBR does not take effect.

Applying PBR for IPv6 packets locally initiated

- To perform personalized management for IPv6 protocol interaction and IPv6 network topologies, you should perform this configuration.
- Run the **ipv6 local policy route-map** command to apply a route map to a device. Then, PBR is executed for IPv6 packets locally initiated.

Command	**ipv6 local policy route-map** *route-map-name*
Parameter Description	*route-map-name*: Indicates the name of a route map.
Defaults	By default, PBR is unavailable on a device and packets are forwarded based on a routing table.
Command Mode	Global configuration mode

Usage Guide	Only one **ipv6 local policy route-map** command can be configured for a device. If the route map used in PBR is unavailable, the PBR does not take effect.

Verification

- Check the configurations of PBR.
- Check the configurations of the route map used by PBR.
- If an ACL is used when the route map is configured, you should check the configurations of the ACL.

Checking the configurations of IPv4 PBR

Command	**show ip policy** [*route-map-name*]
Parameter Description	*route-map-name*: Indicates the name of a route map.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Check the interfaces configured with IPv4 PBR according to the output information and the name of the used route map. FS# show ip policy Banlance mode: redundance Interface Route map local RM_for_PBR_1 GigabitEthernet 0/1 RM_for_PBR_2 **Local** indicates applying policy-based routing for IPv4 packets locally initiated.

Checking the configurations of IPv6 PBR

Command	**show ipv6 policy** [*route-map-name*]
Parameter Description	*route-map-name*: Indicates the name of a route map.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Check the interfaces configured with IPv6 PBR according to the output information and the name of the used route map. FS#show ipv6 policy Banlance mode: redundance Interface Route map local RM_for_PBR_1 VLAN 1 RM_for_PBR_2 **Local** indicates applying policy-based routing for IPv6 packets locally initiated.

Checking the configurations of a route map

Command	**show route-map** [*route-map-name*]

Parameter Description	*route-map-name*: Indicates the name of a route map.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Multiple route maps may be available on a device. Focus on the route map used in PBR and check its policy settings. FS# show route-map route-map RM_FOR_PBR, permit, sequence 10 Match clauses: ip address acl1 Set clauses: ip next-hop 200.24.18.1 route-map RM_FOR_PBR, permit, sequence 20 Match clauses: ip address acl2 Set clauses: ip next-hop 200.24.19.1

Checking the configurations of an ACL

Command	**show access-lists** [*acl-id* | *acl-name*]
Parameter Description	*acl-id*: Indicates the ACL ID. *acl-name*: Indicates the ACL name.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Multiple ACLs may be available on a device. Focus on the ACL used by a route map and check its configurations. FS# show access-lists 1 ip access-list standard 1 10 permit 200.24.16.0 0.0.0.255 ip access-list standard 2 10 permit 200.24.17.0 0.0.0.255

Checking the routing information of IPv4 PBR

Command	**show ip pbr route** [**interface** *if-name* | **local**]
Parameter Description	*if-name*: Indicates an interface name. **local**: Indicates local.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Specify a local interface or device and check the routing information of IPv4 PBR.

```
FS# show ip pbr route
PBR IPv4 Route Summay : 1
Interface              : GigabitEthernet 0/1
  Sequence             : 10
  Min Length           : None
  Max Length           : None
  VRF ID               : 0
  Route Flags          :
    Route Type         : PBR
    Direct             : Permit
    Priority           : High
    Tos_Dscp           : None
    Precedence         : None
  Tos_Dscp             : 0
  Precedence           : 0
  Mode                 : redundance
  Nexthop Count        : 1
    Nexthop[0]         : 192.168.8.100
    Weight[0]          : 1
    Ifindex[0]         : 2
```

Checking the routing information of IPv6 PBR

Command	**show ipv6 pbr route** [**interface** *if-name* | **local**]
Parameter Description	*if-name*: Indicates an interface name. **local**: Indicates local.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Specify a local interface or device and check the routing information of IPv6 PBR.

```
FS# show ipv6 pbr route
PBR IPv6 Route Summary : 1
Interface              : GigabitEthernet 0/1
  Sequence             : 10
  ACL[0]               : 2900
  ACL_CLS[0]           : 5
  Min Length           : None
```

```
  Max Length          : None
  VRF ID              : 0
  Route Flags         :
    Route Type        : PBR
    Direct            : Permit
    Priority          : High
    Tos_Dscp          : None
    Precedence        : None
  Tos_Dscp            : 0
  Precedence          : 0
  Mode                : redundance
  Nexthop Count       : 1
    Nexthop[0]        : 10::2
    Weight[0]         : 1
    Ifindex[0]        : 2
```

Checking a route map used by IPv4 PBR

Command	**show ip pbr route-map** *rmap-name*
Parameter Description	*rmap-name*: Indicates the route map name.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Specify a route map and check the route map used by IPv4 PBR.

```
FS# show ip pbr route-map rm
PBR VRF: GLOBAL, ID: 0
  Forward Mode: redundance
  Forwarding: On

Route-map rm
   Route-map index: Sequence 10, permit
      Match rule:
         ACL ID :      2900, CLS: 1, Name: acl1
      Set rule:
         IPv4 nexthop: 192.168.8.100, (VRF name: , ID: 0), Weight: 0
            PBR state info ifx: 2, Connected: True, Track state: Up
```

Checking a route map used by IPv6 PBR

Command	**show ipv6 pbr route-map** *rmap-name*
Parameter Description	*rmap-name*: Indicates the route map name.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Specify a route map and check the route map used by IPv6 PBR. FS# show ipv6 pbr route-map rm6 PBR VRF: GLOBAL, ID: 0 Forward Mode: redundance Forwarding: On Route-map rm6 Route-map index: Sequence 10, permit Match rule: ACL ID : 2901, CLS: 5, Name: acl6 Set rule: IPv6 nexthop: 10::2, (VRF name: , ID: 0), Weight: 0 PBR state info ifx: 2, Connected: True, Track state: Up

Checking the statistics about packets forwarded by IPv4 PBR

Command	**show ip pbr statistics** [**interface** *if-name* | **local**]
Parameter Description	*if-name*: Indicates an interface name. **local**: Indicates local.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	FS# show ip pbr statistics IPv4 Policy-based route statistic gigabitEthernet 0/1 statistics : 10

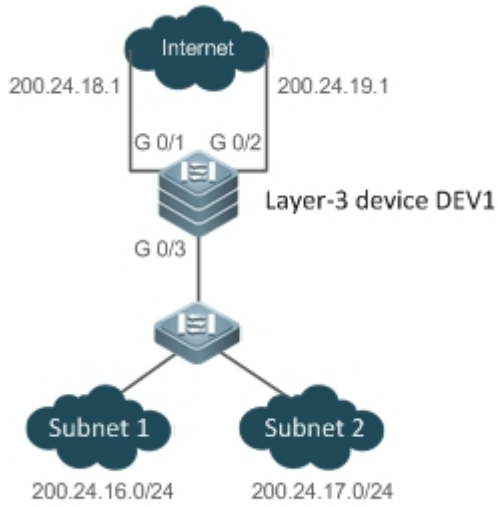
Checking the statistics about packets forwarded by IPv6 PBR

Command	**show ipv6 pbr statistics** [**interface** *if-name* | **local**]
Parameter Description	*if-name*: Indicates an interface name. **local**: Indicates local.
Command Mode	Privilege, global and interface configuration modes

Usage Guide	FS# show ipv6 pbr statistics IPv6 Policy-based route statistic gigabitEthernet 0/1 statistics : 20

Configuration Example

Configuring IPv4 PBR and selecting an output link based on source addresses of packets

Scenario **Figure 6- 3**	
	The layer-3 device DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 200.24.16.0/24 whereas the network segment where subnet 2 resides is 200.24.17.0/24. DEV 1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 200.24.18.1 and 200.24.19.1.
	This LAN has two output interfaces for connecting the Internet. The requirements are as follows: ● Data streams from subnet 1 for accessing the Internet should pass GE 0/1. ● Data streams from subnet 2 for accessing the Internet should pass GE 0/2. ● If the GE 0/1 link is disconnected, the data streams on the GE 0/1 interface are switched to the GE 0/2 interface. Vice versa.
Configuration Steps	● Configure two ACLs to match packets from subnets 1 and 2 respectively. ● Set a policy to set the next hops for packets from subnet 1 to GE0/1 and GE0/2. (Pay attention to the configuration sequence.) ● Set a policy to set the next hops for packets from subnet 2 to GE0/2 and GE0/1. (Pay attention to the configuration sequence.) ● Apply the policy to GE 0/3. ● Set PBR to implement redundant backup among multiple next hops. (The default setting is redundant backup.) During redundant backup, based on the configuration sequence, the first next hop takes effect first.

```
DEV1(config)# access-list 1 permit 200.24.16.0 0.0.0.255
DEV1(config)# access-list 2 permit 200.24.17.0 0.0.0.255
DEV1(config)# route-map RM_FOR_PBR 10
DEV1(config-route-map)# match ip address 1
DEV1(config-route-map)# set ip next-hop 200.24.18.1
DEV1(config-route-map)# set ip next-hop 200.24.19.1
DEV1(config-route-map)# exit
DEV1(config)# route-map RM_FOR_PBR 20
DEV1(config-route-map)# match ip address 2
DEV1(config-route-map)# set ip next-hop 200.24.19.1
DEV1(config-route-map)# set ip next-hop 200.24.18.1
DEV1(config-route-map)# exit
DEV1(config)# interface GigabitEthernet 0/3
DEV1(config-if-GigabitEthernet 0/3)# ip policy route-map RM_FOR_PBR
DEV1(config-if-GigabitEthernet 0/3)# exit
DEV1(config)# ip policy redundance
```

Verification

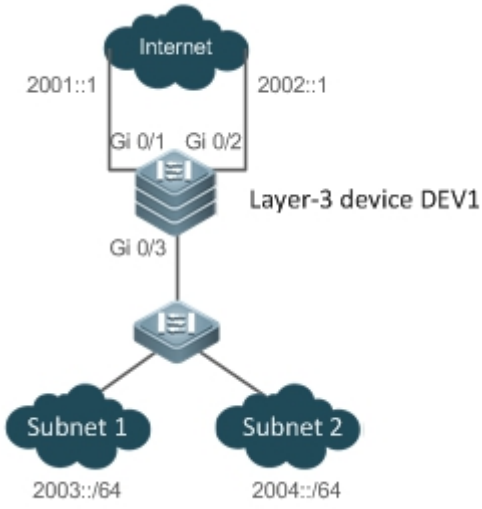
- Check the configurations of IPv4 PBR.
- Check the configurations of the route map.
- Check the configurations of an ACL.

```
DEV1# show ip policy
Interface                    Route map
GigabitEthernet 0/3          RM_FOR_PBR
```

```
DEV1# show route-map
route-map RM_FOR_PBR, permit, sequence 10
  Match clauses:
    ip address 1
  Set clauses:
ip next-hop 200.24.18.1 200.24.19.1
route-map RM_FOR_PBR, permit, sequence 20
  Match clauses:
    ip address  2
  Set clauses:
```

	ip next-hop 200.24.19.1 200.24.18.1
	DEV1# show access-lists ip access-list standard 1 10 permit 200.24.16.0 0.0.0.255 ip access-list standard 2 10 permit 200.24.17.0 0.0.0.255

Configuring IPv6 PBR and selecting an output link based on source addresses of packets

Scenario **Figure 6- 4**	
	DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 2003::/64 whereas the network segment where subnet 2 resides is 2004::/64. DEV1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 2001::1/64 and 2002::1/64.
	This LAN has two output interfaces for connecting the Internet. The requirements are as follows: ● Data streams from subnet 1 for accessing the Internet should pass GE 0/1. ● Data streams from subnet 2 for accessing the Internet should pass GE 0/2. ● If the GE 0/1 link is disconnected, the data streams on the GE 0/1 interface are switched to the GE 0/2 interface. Vice versa.
Configuration Steps	● Configure two ACLs to match packets from subnets 1 and 2 respectively. ● Set a policy to set the next hops for packets from subnet 1 to GE0/1 and GE0/2. (Pay attention to the configuration sequence.) ● Set a policy to set the next hops for packets from subnet 2 to GE0/2 and GE0/1. (Pay attention to the configuration sequence.) ● Apply the policy to GE 0/3. ● Set PBR to implement redundant backup among multiple next hops. During redundant backup, based on the configuration sequence, the first next hop takes effect first.

```
DEV1(config)# ipv6 access-list net1
DEV1(config-ipv6-acl)# permit ipv6 2003::/64 any
DEV1(config-ipv6-acl)# exit
DEV1(config)# ipv6 access-list net2
DEV1(config-ipv6-acl)# permit ipv6 2004::/64 any
DEV1(config-ipv6-acl)# exit
DEV1(config)# route-map RM_FOR_PBR 30
DEV1(config-route-map)# match ipv6 address net1
DEV1(config-route-map)# set ipv6 next-hop 2001::1
DEV1(config-route-map)# set ipv6 next-hop 2002::1
DEV1(config-route-map)# exit
DEV1(config)# route-map RM_FOR_PBR 40
DEV1(config-route-map)# match ipv6 address net2
DEV1(config-route-map)# set ipv6 next-hop 2002::1
DEV1(config-route-map)# set ipv6 next-hop 2001::1
DEV1(config-route-map)# exit
DEV1(config)# interface GigabitEthernet 0/3
DEV1(config-if-GigabitEthernet 0/3)# ipv6 policy route-map RM_FOR_PBR
DEV1(config-if-GigabitEthernet 0/3)# exit
DEV1(config)# ipv6 policy redundance
```

Verification

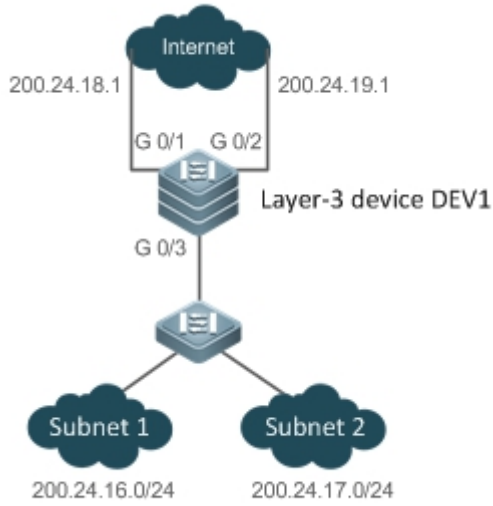
- Check the configurations of IPv6 PBR.
- Check the configurations of the route map.
- Check the configurations of an ACL.

```
DEV1# show ipv6 policy
Interface                    Route map
GigabitEthernet 0/3          RM_FOR_PBR
```

```
DEV1# show route-map
route-map RM_FOR_PBR, permit, sequence 11
  Match clauses:
    ipv6 address net1
  Set clauses:
    ipv6 next-hop 2001::1 2002::1
```

	route-map RM_FOR_PBR, permit, sequence 21 Match clauses: ipv6 address net2 Set clauses: ipv6 next-hop 2002::1 2001::1
	DEV1# show access-lists ipv6 access-list net1 10 permit ipv6 2003::/64 any (0 packets matched) ipv6 access-list net2 10 permit ipv6 2004::/64 any (0 packets matched)

Configuring correlation between IPv4 PBR and Track

Scenario **Figure 6- 5**	
	The layer-3 device DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 200.24.16.0/24 whereas the network segment where subnet 2 resides is 200.24.17.0/24. DEV 1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 200.24.18.1 and 200.24.19.1.
	● DEV1 can fast detect a faulty output link and switch to a backup link.
Configuration Steps	When configuring IPv4 PBR and selecting an output link based on source addresses of the packets, add or modify the following configurations (red fields): ● Set two Track objects and track the accessibility of the next hops of the two output interfaces. ● When configuring a policy, set the correlation between the next hops and the Track objects.
DEV1	DEV1(config)# ip access-list extended 101

	DEV1(config-ip-acl)# permit ip 200.24.16.0 0.0.0.255 any DEV1(config-ip-acl)# exit DEV1(config)# ip access-list extended 102 DEV1(config-ip-acl)# permit ip 200.24.17.0 0.0.0.255 any DEV1(config-ip-acl)# exit DEV1(config)#ip rns 1 DEV1(config-ip-rns)#icmp-echo 200.24.18.1 DEV1(config)#ip rns schedule 1 start-time now life forever DEV1(config)#track 1 rns 1 DEV1(config)#ip rns 2 DEV1(config-ip-rns)#icmp-echo 200.24.19.1 DEV1(config)#ip rns schedule 2 start-time now life forever DEV1(config)#track 2 rns 2 DEV1(config)# route-map RM_FOR_PBR 10 DEV1(config-route-map)# match ip address 101 DEV1(config-route-map)# set ip next-hop verify-availability 200.24.18.1 track 1 DEV1(config-route-map)# set ip next-hop verify-availability 200.24.19.1 track 2 DEV1(config-route-map)# exit DEV1(config)# route-map RM_FOR_PBR 20 DEV1(config-route-map)# match ip address 102 DEV1(config-route-map)# set ip next-hop verify-availability 200.24.19.1 track 2 DEV1(config-route-map)# set ip next-hop verify-availability 200.24.18.1 track 1 DEV1(config-route-map)# exit DEV1(config)# interface GigabitEthernet 0/3 DEV1(config-if-GigabitEthernet 0/3)# ip policy route-map RM_FOR_PBR DEV1(config-if-GigabitEthernet 0/3)# exit DEV1(config)# ip policy redundance
Verification	● Check whether the Track objects are up.
DEV1	DEV1#show track Track 1 Reliable Network Service 1 The state is Up 1 change, current state last: 120 secs

	Delay up 30 secs, down 50 secs Track 2 Reliable Network Service 2 The state is Up 1 change, current state last: 130 secs Delay up 30 secs, down 50 secs

Configuring IPv4 PBR and transferring global packets to a VRF for forwarding

	VRF1 and VRF2 are available on the device. Select VRFs for forwarding IPv4 packets received on GE0/3: ● Forward IPv4 packets from subnet 1 in VRF 1. ● Forward IPv4 packets from subnet 2 in VRF 2.
Configuration Steps	● Configure a single-protocol VRF (or multi-protocol VRF to enable the IPv4 address family). ● Configure ACL1: the source addresses of IPv4 packets belong to subnet 1. ● Configure ACL2: the source addresses of IPv4 packets belong to subnet 2. ● Set policy 10 in a route map: forward packets matching ACL 1 in VRF1. ● Set policy 20 in a route map: forward packets matching ACL 2 in VRF2. ● Apply the route map to GE 0/3. ● Redirect the host route and direct route on GE 0/3 to the VRF.
Single-protocol VRF	DEV1 (config)# ip vrf VRF1 DEV1 (config)# ip vrf VRF2 DEV1 (config)# access-list 1 permit 192.168.195.0 0.0.0.255 DEV1 (config)# access-list 2 permit 192.168.196.0 0.0.0.255 DEV1 (config)# route-map PBR-VRF-Selection permit 10 DEV1 (config-route-map)# match ip address 1 DEV1 (config-route-map)# set vrf VRF1 DEV1 (config-route-map)# exit DEV1 (config)# route-map PBR-VRF-Selection permit 20 DEV1 (config-route-map)# match ip address 2 DEV1 (config-route-map)# set vrf VRF2 DEV1 (config-route-map)# exit DEV1 (config)# interface GigabitEthernet 0/3 DEV1 (config-if-GigabitEthernet 0/3)# ip policy route-map PBR-VRF-Selection DEV1 (config-if-GigabitEthernet 0/3)# ip address 192.168.195.1 255.255.255.0 DEV1 (config-if-GigabitEthernet 0/3)# ip vrf receive VRF1

	DEV1 (config-if-GigabitEthernet 0/3)# ip vrf receive VRF2 DEV1 (config-if-GigabitEthernet 0/3)# exit
Multi-protocol VRF	DEV1 (config)# vrf definition VRF1 DEV1 (config-vrf)# address-family ipv4 DEV1 (config-vrf-af)# exit-address-family DEV1 (config-vrf)# exit DEV1 (config)# vrf definition VRF2 DEV1 (config-vrf)# address-family ipv4 DEV1 (config-vrf-af)# exit-address-family DEV1 (config-vrf)# exit DEV1 (config)# access-list 1 permit 192.168.195.0 0.0.0.255 DEV1 (config)# access-list 2 permit 192.168.196.0 0.0.0.255 DEV1 (config)# route-map PBR-VRF-Selection permit 10 DEV1 (config-route-map)# match ip address 1 DEV1 (config-route-map)# set vrf VRF1 DEV1 (config-route-map)# exit DEV1 (config)# route-map PBR-VRF-Selection permit 20 DEV1 (config-route-map)# match ip address 2 DEV1 (config-route-map)# set vrf VRF2 DEV1 (config-route-map)# exit DEV1 (config)# interface GigabitEthernet 0/3 DEV1 (config-if-GigabitEthernet 0/3)# ip policy route-map PBR-VRF-Selection DEV1 (config-if-GigabitEthernet 0/3)# ip address 192.168.195.1 255.255.255.0 DEV1 (config-if-GigabitEthernet 0/3)# vrf receive VRF1 DEV1 (config-if-GigabitEthernet 0/3)# vrf receive VRF2 DEV1 (config-if-GigabitEthernet 0/3)# exit
Verification	● Check the configurations of IPv4 PBR. ● Check the configurations of the route map. ● Check the configurations of the ACLs.
	DEV1# show ip policy Interface Route map

	GigabitEthernet 0/3 PBR-VRF-Selection
	DEV1# show route-map route-map PBR-VRF-Selection, permit, sequence 10 Match clauses: ip address 1 Set clauses: vrf VRF1 route-map PBR-VRF-Selection, permit, sequence 20 Match clauses: ip address 2 Set clauses: vrf VRF2
	DEV1# show access-lists ip access-list standard 1 10 permit 192.168.195.0 0.0.0.255 ip access-list standard 2 10 permit 192.168.196.0 0.0.0.255

Configuring IPv6 PBR and transferring global packets to a VRF for forwarding

	VRF1 and VRF2 are available on the device. Select a VRF for forwarding IPv6 packets received on GE0/3: ● Forward IPv6 packets from subnet 1 in VRF 1. ● Forward IPv6 packets from subnet 2 in VRF 2.
Configuration Steps	● Configure multi-protocol VRFs and enable the IPv6 address family. ● Configure ACL net1: the source addresses of IPv6 packets belong to subnet 1. ● Configure ACL net2: the source addresses of IPv6 packets belong to subnet 2. ● Set policy 10 in a route map: forward packets matching ACL 1 in VRF1. ● Set policy 20 in a route map: forward packets matching ACL 2 in VRF2. ● Apply the route map to GE 0/3. ● Redirect the host route and direct route on GE 0/3 to the VRF.
Multi-protocol VRF	DEV1(config)# vrf definition VRF1 DEV1(config-vrf)# address-family ipv6 DEV1(config-vrf-af)# exit-address-family DEV1(config-vrf)# exit

```
DEV1(config)# vrf definition VRF2
DEV1(config-vrf)# address-family ipv6
DEV1(config-vrf-af)# exit-address-family
DEV1(config-vrf)# exit
DEV1(config)# ipv6 access-list net1
DEV1(config-ipv6-acl)# permit ipv6 1000::/64 any
DEV1(config-ipv6-acl)# exit
DEV1(config)# ipv6 access-list net2
DEV1(config-ipv6-acl)# permit ipv6 2000::/64 any
DEV1(config-ipv6-acl)# exit
DEV1(config)# route-map PBR-VRF-Selection permit 10
DEV1(config-route-map)# match ipv6 address net1
DEV1(config-route-map)# set vrf VRF1
DEV1(config-route-map)# exit
DEV1(config)# route-map PBR-VRF-Selection permit 20
DEV1(config-route-map)# match ipv6 address net2
DEV1(config-route-map)# set vrf VRF2
DEV1(config-route-map)# exit
DEV1(config)# interface GigabitEthernet 0/3
DEV1(config-if-GigabitEthernet 0/3)# ipv6 policy route-map PBR-VRF-Selection
DEV1(config-if-GigabitEthernet 0/3)# vrf receive VRF1
DEV1(config-if-GigabitEthernet 0/3)# vrf receive VRF2
DEV1(config-if-GigabitEthernet 0/3)# exit
```

Verification

- Check the configurations of IPv6 PBR.
- Check the configurations of the route map.
- Check the configurations of the ACLs.

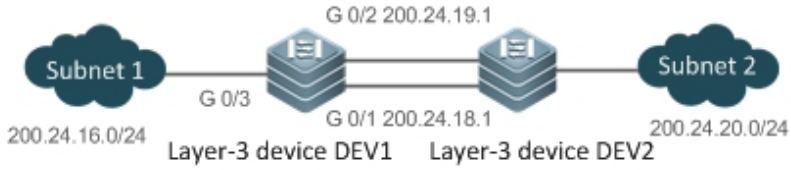
```
DEV1# show ipv6 policy
Interface                    Route map
GigabitEthernet 0/3          PBR-VRF-Selection
```

```
DEV1# show route-map
route-map PBR-VRF-Selection, permit, sequence 10
  Match clauses:
```

```
    ip address 1
    ipv6 address net1
  Set clauses:
    vrf VRF1
route-map PBR-VRF-Selection, permit, sequence 20
  Match clauses:
    ip address 2
    ipv6 address net2
  Set clauses:
    vrf VRF2
```

```
DEV1# show access-lists
ipv6 access-list net1
 10 permit ipv6 1000::/64 any
ipv6 access-list net2
 10 permit ipv6 2000::/64 any
```

Configuring IPv4 recursive PBR, selecting an output link based on source addresses of the packets, and recurring to the output link of a dynamic route

Scenario **Figure 6- 6**	
	The layer-3 device DEV 1 is connected to subnet 1 through GE0/3. The network segment where subnet 1 resides is 200.24.16.0/24. DEV 1 is connected to subnet 2 through GE0/1 and GE0/2 and their next hops are 200.24.18.1 and 200.24.19.1.
	Subnet 1 is connected to subnet 2 through two output interfaces of DEV1. The requirements are as follows: ● Configure static or dynamic routes in advance to ensure that static or dynamic routes in the network segment 200.24.20.0 are available in the routing table of DEV1. ● Data streams from subnet 1 for accessing the Internet can recur to a dynamic route whose IP address is 200.24.20.1. ● If the GE 0/1 link is disconnected, the data streams on GE 0/1 are switched to GE 0/2. Vice versa.
Configuration Steps	● Configure an ACL to match packets from subnet 1. ● Set a policy to set the recursive next hop for packets from subnet 1 to 200.24.20.1.

	● Apply the policy to GE 0/3. ● Set PBR to implement redundant backup among multiple next hops. (The default setting is redundant backup.) During redundant backup, the sequence for the next hops to take effect is related to the sequence for the static or dynamic routes to take effect.
	DEV1(config)# access-list 1 permit 200.24.16.0 0.0.0.255 DEV1(config)# route-map RM_FOR_PBR 10 DEV1(config-route-map)# match ip address 1 DEV1(config-route-map)# set ip next-hop recursive 200.24.20.1 DEV1(config-route-map)# exit DEV1(config)# interface GigabitEthernet 0/3 DEV1(config-if-GigabitEthernet 0/3)# ip policy route-map RM_FOR_PBR DEV1(config-if-GigabitEthernet 0/3)# exit DEV1(config)# ip policy redundance
Verification	● Check the configurations of IPv4 PBR. ● Check the configurations of the route map. ● Check the configurations of the ACLs.
	DEV1# show ip policy Interface Route map GigabitEthernet 0/3 RM_FOR_PBR
	DEV1# show route-map route-map RM_FOR_PBR, permit, sequence 10 Match clauses: ip address 1 Set clauses: ip next-hop recursive 200.24.20.1
	DEV1# show access-lists ip access-list standard 1 10 permit 200.24.16.0 0.0.0.255

Common Errors

- A route map is used when PBR is configured but the route map does not exist.
- An ACL is used when a route map is configured but the ACL does not exist.
- A VRF is used when a route map is configured but the VRF does not exist.

- When multi-protocol VRF is configured, the IPv4 or IPv6 address family is not enabled.
- When PBR is used for VRF transfer, the host route and direct route on the interface are not redirected to the VRF.

6.4.2 Setting Redundant Backup or Load Balancing

Configuration Effect

- Using multiple next hops in the mutual backup mode can enhance the network reliability.
- Implementing load balancing among multiple next hops can expand the network bandwidth.

Notes

- The basic functions of PBR must be configured.
- Redundant backup and load balancing are effective only for the next hops set by the following **set** commands.

Command	Description
set ip next-hop	Configures the next hop of IPv4 packets. This command carries the *weight* parameter, which is used to set the weight of the WCMP. The default value is 1.
set ip default next-hop	Configures the default next hop of IPv4 packets. This command carries the *weight* parameter, which is used to set the weight of the WCMP. The default value is 1.
set ipv6 next-hop	Configures the next hop of IPv6 packets. This command carries the *weight* parameter, which is used to set the weight of the WCMP. The default value is 1.
set ipv6 default next-hop	Configures the default next hop of IPv6 packets. This command carries the *weight* parameter, which is used to set the weight of the WCMP. The default value is 1.
set ip next-hop recursive	Configures the recursive next hop of IPv4 packets. Only one command can be configured for a route map and packets can recur to multiple next hops (up to 32 next hops) of a static or dynamic ECMP route. The redundant backup or load balancing mode for recurring to multiple next hops is also determined by the **ip policy** { **redundance** | **load-balance** } command.

Up to eight next hops can be set for WCMP whereas up to 32 next hops can be set for ECMP.

Configuration Steps

Setting whether IPv4 PBR implements redundant backup or load balancing among multiple next hops

- If load balancing needs to be implemented among multiple next hops, this configuration needs to be performed.
- If load balancing is configured at present, you also need to perform this configuration to reset redundant backup.
- This configuration is effective for all PBRs configured on a device.

Command	**ip policy** { **redundance** | **load-balance** }
Parameter Description	redundance: Indicates redundant backup. load-balance: Indicates load balancing.
Defaults	Redundant backup is configured by default.
Command Mode	Global configuration mode
Usage Guide	If redundant backup is selected, the first next hop takes effect based on the configuration sequence.

	If load balancing is selected, all next hops take effect at the same time and share traffic by weight.

Setting whether Ipv6 PBR implements redundant backup or load balancing among multiple next hops

- If load balancing needs to be implemented among multiple next hops, this configuration needs to be performed.
- If load balancing is configured at present, you also need to perform this configuration to reset redundant backup.
- This configuration is effective for all PBRs configured on a device.

Command	**ipv6 policy** { **redundance** | **load-balance** }
Parameter Description	redundance: Indicates redundant backup. load-balance: Indicates load balancing.
Defaults	Redundant backup is configured by default.
Command Mode	Global configuration mode
Usage Guide	If redundant backup is selected, the first next hop takes effect based on the configuration sequence. If load balancing is selected, all next hops take effect at the same time and share traffic by weight.

Verification

- Check whether redundant backup or load balancing is implemented among multiple next hops.

Checking whether IPv4 PBR implements redundant backup or load balancing among multiple next hops

Command	**show ip policy** [*route-map-name*]
Parameter Description	***route-map-name*: Specifies a route map.**
Command Mode	Privilege, global and interface configuration modes
Usage Guide	See the following example and focus on the red field.

```
FS# show ip policy
Banlance mode: redundance
Interface                    Route map
local                         test
GigabitEthernet 0/3        test
```

Checking whether IPv6 PBR implements redundant backup or load balancing among multiple next hops

Command	**show ipv6 policy** [*route-map-name*]
Parameter Description	***route-map-name***: Specifies a route map.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	See the following example and focus on the red field.

```
FS#show ipv6 policy
Banlance mode: redundance
Interface                    Route map
VLAN 1                       RM_for_Vlan_1
VLAN 2                       RM_for_Vlan_2
```

Configuration Example

Configuring IPv4 PBR to implement redundant backup among multiple next hops

See the preceding example: Configuring IPv4 PBR and selecting an output link based on source addresses of packets

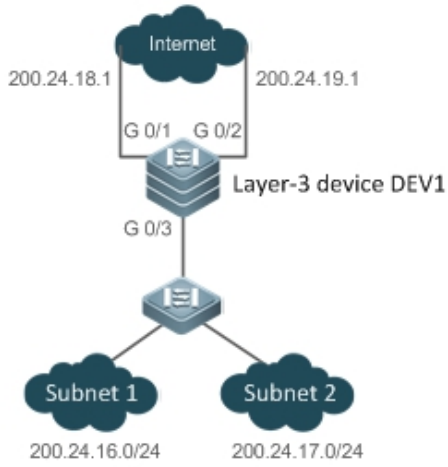
Configuring IPv6 PBR to implement redundant backup among multiple next hops

See the preceding example: Configuring IPv6 PBR and selecting an output link based on source addresses of packets

Configuring IPv4 PBR to implement load balancing among multiple next hops

Scenario

Figure 6- 7



The layer-3 device DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 200.24.16.0/24 whereas the network segment where subnet 2 resides is 200.24.17.0/24.

DEV 1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 200.24.18.1 and 200.24.19.1.

This LAN has two output interfaces for connecting the Internet. The requirements are as follows: The traffic is equally shared by GE0/1 and GE0/2.

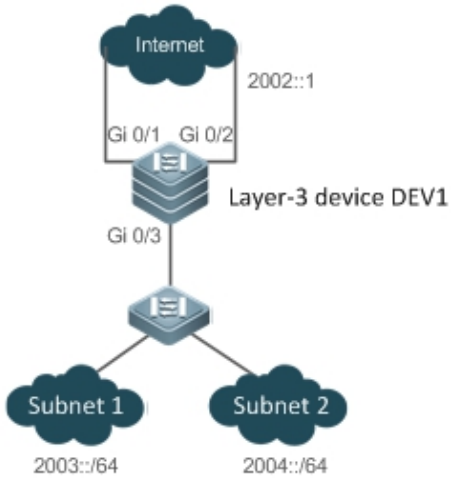
Configuration Steps

- Configure basic functions of PBR. Specify multiple next hops.
- Set the load balancing mode.

```
DEV1(config)# route-map RM_LOAD_PBR 10
DEV1(config-route-map)# set ip next-hop 200.24.18.1
DEV1(config-route-map)# set ip next-hop 200.24.19.1
```

	DEV1(config-route-map)# exit DEV1(config)# interface GigabitEthernet 0/3 DEV1(config-if-GigabitEthernet 0/3)# ip policy route-map RM_LOAD_PBR DEV1(config-if-GigabitEthernet 0/3)# exit DEV1(config)# ip policy load-balance
Verification	● Check the configurations of IPv4 PBR. ● Check the configurations of the route map.
	DEV1# show ip policy Balance mode: load-balance Interface Route map GigabitEthernet 0/3 RM_LOAD_PBR
	DEV1# show route-map route-map PBR-VRF-Selection, permit, sequence 10 Match clauses: Set clauses: ip next-hop 200.24.18.1 8 ip next-hop 200.24.19.1 8

Configuring IPv6 PBR to implement load balancing among multiple next hops

Scenario **Figure 6- 8**	
	DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 2003::/64 whereas the network segment where subnet 2 resides is 2004::/64. DEV1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 2001::1/64 and 2002::1/64.

	This LAN has two output interfaces for connecting the Internet. The requirements are as follows: The traffic is equally shared by GE0/1 and GE0/2.
Configuration Steps	● Configure basic functions of PBR. Specify multiple next hops. ● Set the load balancing mode.
	DEV1(config)# route-map RM_LOAD_PBR 20 DEV1(config-route-map)# set ipv6 next-hop 2001::1 DEV1(config-route-map)# set ipv6 next-hop 2002::1 DEV1(config-route-map)# exit DEV1(config)# interface GigabitEthernet 0/3 DEV1(config-if-GigabitEthernet 0/3)# ipv6 policy route-map RM_LOAD_PBR DEV1(config-if-GigabitEthernet 0/3)# exit DEV1(config)# ipv6 policy load-balance
Verification	● Check the configurations of IPv6 PBR. ● Check the configurations of the route map.
	DEV1# show ipv6 policy Balance mode: load-balance Interface Route map GigabitEthernet 0/3 RM_LOAD_PBR DEV1# show route-map route-map PBR-VRF-Selection, permit, sequence 10 Match clauses: Set clauses: ipv6 next-hop 2001::1 ipv6 next-hop 2002::1

6.4.3 Configuring Source-Address-Based PBR

Configuration Effect

Perform personalized routing management for IPv4 or IPv6 addresses of user data streams by preparing flexible policies.

Notes

● Source-address-based PBR has a higher priority than interface-based PBR. When they are applied to an interface at the same time, interface-based PBR takes effect whereas source-address-based PBR does not take effect.

Configuration Steps

Applying source-address-based PBR for IPv4 packets received by an interface

- To perform personalized routing management based on source IPv4 addresses for IPv4 user data streams passing a device, you should perform this configuration.
- The global configuration takes effect on the input interface of specified user data streams.
- Run the **ip policy-source in-interface** command to perform source-address-based PBR for IPv4 packets received by a specified interface.

Command	**ip policy-source in-interface** *interface-type sequence* {*source-address mask* | *source-address/mask*} {[**default**] **next-hop** *ip-address* [*weight*] [**default**] **interface** *out-interface-type*| **vrf** *vrf-name*}
Parameter Description	*interface-type*: Specifies the type of an interface to which source-address PBR is applied. *sequence*: Indicates the sequence number of a policy. A smaller sequence number means a higher priority. *source-address*: Indicates the source IPv4 address. *mask*: Indicates the mask of the source IPv4 address. *ip-address*: Indicates the next-hop IPv4 address. *weight*: Indicates the weight of a next hop. *out-interface-type*: Indicates the type of the next hop output interface. *vrf-name*: Indicates the name of a VRF instance.
Defaults	By default, source-address-based PBR is unavailable on a device and packets are forwarded based on a routing table.
Command Mode	Global configuration mode
Usage Guide	Configure multiple **ip policy-source in-interface** commands for the same interface. The sequence numbers of different source addresses are different. A smaller sequence number means a higher priority of the source-address-based PBR. When the sequence number is the same, the priorities of next hops are as follows: **vrf** *vrf-name* > **next-hop** *ip-address* > **interface** *out-interface-type* > **default next-hop** *ip-address*> **default interface** *out-interface-type* Source-address-based PBR has a higher priority than interface-based PBR. When they are applied to an interface at the same time, interface-based PBR takes effect whereas source-address-based PBR does not take effect.

Applying source-address-based PBR for IPv6 packets received by an interface

- To perform personalized routing management based on source IPv6 addresses for IPv6 user data streams passing a device, you should perform this configuration.
- The global configuration takes effect on the input interface of specified user data streams.
- Run the **ipv6 policy-source in-interface** command to perform source-address-based PBR for IPv6 packets received by a specified interface.

Command	**ipv6 policy-source in-interface** *interface-type sequence* {*source-address/prefix-length*} {[**default**] **next-hop** *ipv6-address* [*weight*] [**default**] **interface** *out-interface-type*| **vrf** *vrf-name*}
Parameter Description	*interface-type*: Specifies the type of an interface to which source-address PBR is applied. *sequence*: Indicates the sequence number of a policy. A smaller sequence number means a higher priority. *source-address*: Indicates the source IPv6 address. *prefix-length*: Indicates the prefix length of a source IPv6 address. *ipv6-address*: Indicates the next-hop IPv6 address. *weight*: Indicates the weight of a next hop.

	out-interface-type: Indicates the type of the next hop output interface. *vrf-name*: Indicates the name of a VRF instance.
Defaults	By default, source-address-based PBR is unavailable on a device and packets are forwarded based on a routing table.
Command Mode	Global configuration mode
Usage Guide	Configure multiple **ipv6 policy-source in-interface** commands for the same interface. The sequence numbers of different source addresses are different. A smaller sequence number means a higher priority of the source-address-based PBR. When the sequence number is the same, the priorities of next hops are as follows: **vrf** *vrf-name* > **next-hop** *ipv6-address* > **interface** *out-interface-type* > **default next-hop** *ipv6-address*> **default interface** *out-interface-type* Source-address-based PBR has a higher priority than interface-based PBR. When they are applied to an interface at the same time, interface-based PBR takes effect whereas source-address-based PBR does not take effect.

Verification

Checking the routing information of source-address-based IPv4 PBR

Command	**show ip pbr source-route** [**interface** *if-name*]
Parameter Description	*if-name*: Indicates an interface name.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Specify an interface and check the routing information of IPv4 source-address-based PBR.

```
FS# show ip pbr source-route
PBR IPv4 Source Route
Interface              : GigabitEthernet 0/1
  Sequence             : 10
  Source address : 10.1.1.1/24
  VRF ID               : 0
  Route Flags          :
    Route Type         : PBR
    Direct             : Permit
    Priority           : High
    Match_ipaddr       : Exist
  Mode                 : redundance
  Nexthop Count        : 1
    Nexthop[0]         : 192.168.8.100
    Weight[0]          : 1
    Ifindex[0]         : 2
```

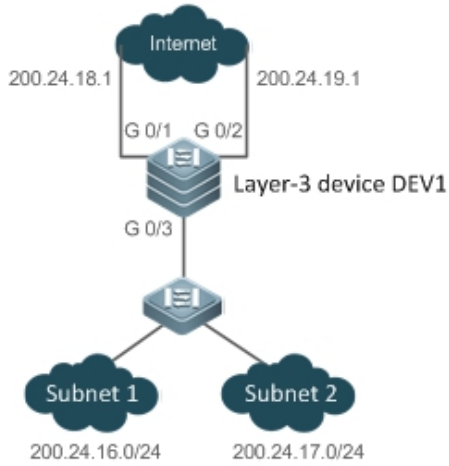
Checking the routing information of IPv6 PBR

Command	**show ipv6 pbr source-route** [**interface** *if-name*]
Parameter Description	*if-name*: Indicates an interface name.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Specify an interface and check the routing information of IPv6 PBR.

```
FS# show ipv6 pbr source-route
PBR IPv6 Source Route
Interface              : GigabitEthernet 0/1
  Sequence             : 10
  Source address : 1000::1/64
  VRF ID                 : 0
  Route Flags        :
    Route Type        : PBR
    Direct               : Permit
    Priority            : High
    Match_ipaddr    : Exist
  Mode                    : redundance
  Nexthop Count    : 1
    Nexthop[0]        : 1001::2
    Weight[0]          : 1
    Ifindex[0]        : 3
```

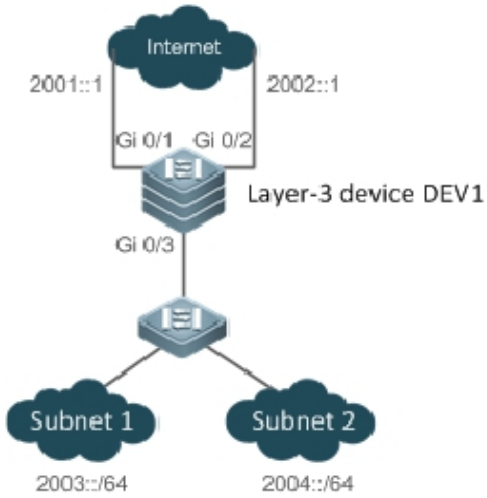
Configuration Example

Configuring IPv4 source-address-based PBR and selecting an output link based on source addresses of packets

Scenario **Figure 6- 9**	
	The layer-3 device DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 200.24.16.0/24 whereas the network segment where subnet 2 resides is 200.24.17.0/24. DEV 1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 200.24.18.1 and 200.24.19.1.
	This LAN has two output interfaces for connecting the Internet. The requirements are as follows: ● Data streams from subnet 1 for accessing the Internet should pass GE 0/1. ● Data streams from subnet 2 for accessing the Internet should pass GE 0/2. ● If the GE 0/1 link is disconnected, the data streams on GE 0/1 are switched to GE 0/2. Vice versa.
Configuration Steps	● Set source-address-based PBR and set the next hops for packets from the GE0/3 subnet 1 to GE0/1 and GE0/2. (Pay attention to the configuration sequence.) ● Set source-address-based PBR and set the next hops for packets from the GE0/3 subnet 2 to GE0/2 and GE0/1. (Pay attention to the configuration sequence.) ● Set PBR to implement redundant backup among multiple next hops. (The default setting is redundant backup.) During redundant backup, based on the configuration sequence, the first next hop takes effect first.
	DEV1(config)# ip policy-source in-interface gigabitEthernet 0/3 1 200.24.16.0/24 next-hop 200.24.18.1 200.24.19.1 DEV1(config)# ip policy-source in-interface gigabitEthernet 0/3 2 200.24.17.0/24 next-hop 200.24.19.1 200.24.18.1 DEV1(config)# ip policy redundance
Verification	● Check the routing information of source-address-based IPv4 PBR.
	DEV1# show ip pbr source-route PBR IPv4 Source Route Interface : GigabitEthernet 0/3 Sequence : 1 Source address : 200.24.16.0/24

```
  VRF ID              : 0
  Route Flags       :
    Route Type       : PBR
    Direct            : Permit
    Priority          : High
    Match_ipaddr    : Exist
  Mode                 : redundance
  Nexthop Count    : 2
    Nexthop[0]       : 200.24.18.1
    Weight[0]        : 1
    Ifindex[0]       : 1
    Nexthop[1]       : 200.24.19.1
    Weight[1]        : 1
    Ifindex[1]       : 2

Interface             : GigabitEthernet 0/3
  Sequence            : 2
  Source address : 200.24.17.0/24
  VRF ID              : 0
  Route Flags       :
    Route Type       : PBR
    Direct            : Permit
    Priority          : High
    Match_ipaddr    : Exist
  Mode                 : redundance
  Nexthop Count    : 2
    Nexthop[0]       : 200.24.19.1
    Weight[0]        : 1
    Ifindex[0]       : 2
    Nexthop[1]       : 200.24.18.1
    Weight[1]        : 1
    Ifindex[1]       : 1
```

Configuring IPv6 source-address-based PBR and selecting an output link based on source addresses of packets

Scenario **Figure 6- 10**	
	DEV 1 is connected to subnet 1 and subnet 2 through GE0/3. The network segment where subnet 1 resides is 2003::/64 whereas the network segment where subnet 2 resides is 2004::/64. DEV1 is connected to the Internet through GE0/1 and GE0/2 and their next hops are 2001::1/64 and 2002::1/64.
	This LAN has two output interfaces for connecting the Internet. The requirements are as follows: ● Data streams from subnet 1 for accessing the Internet should pass GE 0/1. ● Data streams from subnet 2 for accessing the Internet should pass GE 0/2. ● If the GE 0/1 link is faulty, the data streams on GE 0/1 are switched to GE 0/2. Vice versa.
Configuration Steps	● Set source-address-based PBR and set the next hops for packets from the GE0/3 subnet 1 to GE0/1 and GE0/2. (Pay attention to the configuration sequence.) ● Set source-address-based PBR and set the next hops for packets from the GE0/3 subnet 2 to GE0/2 and GE0/1. (Pay attention to the configuration sequence.) ● Set PBR to implement redundant backup among multiple next hops. During redundant backup, based on the configuration sequence, the first next hop takes effect first.
	DEV1(config)# ipv6 policy-source in-interface gigabitEthernet 0/3 1 2003::/64 next-hop 2001::1 2002::1 DEV1(config)# ip policy-source in-interface gigabitEthernet 0/3 2 2004::/64 next-hop 2002::1 2001::1 DEV1(config)# ipv6 policy redundance
Verification	● Check the configuration of IPv6 source-address-based PBR.
	DEV1# show ipv6 pbr source-route PBR IPv6 Source Route Interface : GigabitEthernet 0/3 Sequence : 1

```
  Source address : 2003::/64
  VRF ID                 : 0
  Route Flags        :
    Route Type        : PBR
    Direct               : Permit
    Priority           : High
    Match_ipaddr    : Exist
  Mode                     : redundance
  Nexthop Count    : 2
    Nexthop[0]        : 2001::1
    Weight[0]          : 1
    Ifindex[0]        : 1
    Nexthop[1]        : 2002::1
    Weight[1]          : 1
    Ifindex[1]        : 2

Interface               : GigabitEthernet 0/3
  Sequence             : 2
  Source address : 2004::/64
  VRF ID                 : 0
  Route Flags        :
    Route Type        : PBR
    Direct               : Permit
    Priority           : High
    Match_ipaddr    : Exist
  Mode                     : redundance
  Nexthop Count    : 2
    Nexthop[0]        : 2002::1
    Weight[0]          : 1
    Ifindex[0]        : 2
    Nexthop[1]        : 2001::1
    Weight[1]          : 1
    Ifindex[1]        : 1
```

6.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the statistics about packets forwarded by IPv4 PBR.	**clear ip pbr statistics** [**interface** *if-name* | **local**]
Clears the statistics about packets forwarded by IPv6 PBR.	**clear ipv6 pbr statistics** [**interface** *if-name* | **local**]

Displaying

Description	Command
Displays the configurations of IPv4 PBR.	**show ip policy**
Displays the configurations of IPv6 PBR.	**show ipv6 policy**
Displays the configurations of a route map.	**show route-map** [*name*]
Displays the configurations of an ACL.	**show access-list**
Displays the correlation between IPv4 PBR and BFD.	**show ip pbr bfd**
Displays the correlation between IPv6 PBR and BFD.	**show ipv6 pbr bfd**
Displays the routing information of IPv4 PBR.	**show ip pbr route** [**interface** *if-name* | **local**]
Displays the routing information of IPv6 PBR.	**show ipv6 pbr route** [**interface** *if-name* | **local**]
Displays a route map used by IPv4 PBR.	**show ip pbr route-map** *rmap-name*
Displays a route map used by IPv6 PBR.	**show ipv6 pbr route-map** *rmap-name*
Displays the routing information of IPv4 source-address-based PBR.	**show ip pbr source–route** [**interface** *if-name*]
Displays the routing information of IPv6 source-address-based PBR.	**show ipv6 pbr source-route** [**interface** *if-name*]
Displays the statistics about IPv4 PBR.	**show ip pbr statistics** [**interface** *if-name* | **local**]
Displays the statistics about IPv6 PBR.	**show ipv6 pbr statistics** [**interface** *if-name* | **local**]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs PBR errors.	**debug pbr error**
Debugs PBR events.	**debug pbr events**
Debugs multiple service cards supported by PBR.	**debug pbr ms**
Debugs PBR message communication.	**debug pbr msg**
Debugs interaction between PBR and NSM.	**debug pbr nsm**
Debugs packet forwarding of PBR.	**debug pbr packet**
Debugs PBR GR.	**debug pbr restart**

7 Configuring VRF

7.1 Overview

A Virtual Private Network (VPN) Routing and Forwarding (VRF) table is used for the forwarding of VPN packets. Each VPN corresponds to a VRF table.

A device that provides the VPN service has multiple routing tales, including a public network routing table and one or multiple VRF tables. The public-network routing table is used for the forwarding of public network packets, and the VRF tables are used for the forwarding of VPN packets. These routing tables are created to separate routes in the public network from those in VPNs and separate routes in different VPNs.

A VPN is a private dedicated network built in the public network. "Virtual" means that the VPN is logically exclusive, instead of physically exclusive.

Protocols and Standards

- RFC4364: BGP/MPLS IP Virtual Private Networks (VPNs)

7.2 Applications

Application	Description
Local Inter-VPN Access	Provide the VPN service on a routing device and enable VPNs to access each other.
VRF only on Provider Edges (PEs)	Provide the VPN service in an IP/Multiprotocol Label Switching (MPLS) network and connect one Customer Edge (CE) to one VPN.
VRF on CEs and PEs	Provide the VPN service in an IP/ MPLS network and connect one CE to multiple VPNs.

CE: An edge device in a customer network

PE: An edge device in a Service Provider (SP) network

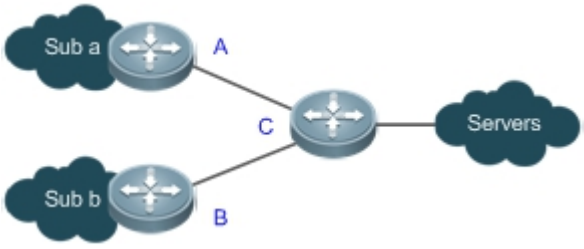
7.2.1 Local Inter-VPN Access

Scenario

Provide the VPN service on a routing device and enable VPNs to access each other.

In Figure 7- 1, Sub a runs the Routing Information Protocol (RIP), Sub b runs the Open Shortest Path First (OSPF) protocol, and Servers is a network segment directly connected to C. Provide the VPN service on C to Sub a, Sub b, and Servers, and enable Sub a and Sub b to access Servers.

Figure 7- 1



Related Configuration

- On C, create a VRF table for Sub a, bind the interface directly connected to A, and associate the VRF table with A by using RIP.

- On C, create a VRF table for Sub b, bind the interface directly connected to B, and associate the VRF table with B by using OSPF.
- On C, create a VRF table for Servers and bind the interface directly connected to Servers.
- On C, configure the route targets (RTs) of the VRF tables for Suba, Subb, and Servers. Import the routes in the VRF tables for Sub a and Sub b to the VRF table for Servers, and import the routes in the VRF table for Servers to the VRF tables for Sub a and Sub b.
- Configure the Border Gateway Protocol (BGP) on C. Introduce the RIP routes to the VRF table for Sub a, introduce the OSPF routes to the VRF table for Sub b, and introduce the direct routes to the VRF table for Servers.

7.2.2 VRF only on PEs

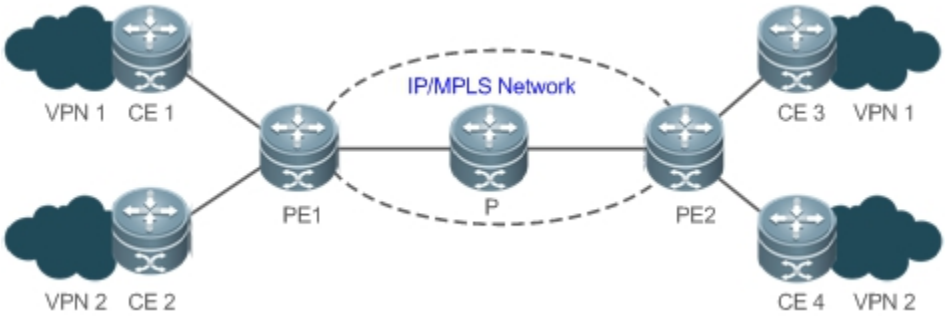
Scenario

An Internet Service Provider (ISP) provides the VPN service in an IP/MPLS backbone network.

In Figure 7- 2, VPN1 runs RIP, and VPN2 runs OSPF.

- One CE is connected to one VPN, and all routes on the CE are exclusively used by the connected VPN. Therefore, no VRF table needs to be created to separate the routes.
- On each PE, VRF tables must be created to separate the routes in VPN1, those in VPN2, and those in the public network from each other.

Figure 7- 2



Deployment

- On PE1, create a VRF table for VPN1 and bind the interface directly connected to CE1. On PE2, create a VRF table for VPN1 and bind the interface directly connected to CE3.
- On PE1, create a VRF table for VPN2 and bind the interface directly connected to CE2. On PE2, create a VRF table for VPN2 and bind the interface directly connected to CE4.
- On PE1, associate the VRF table for VPN1 with CE1 by using RIP. On PE2, associate the VRF table for VPN1 with CE3 by using RIP.
- On PE1, associate the VRF table for VPN2 with CE2 by using OSPF. On PE2, associate the VRF table for VPN2 with CE4 by using OSPF.
- Create a BGP neighbor (VPNv4 address family) between PE1 and PE2.
- In the VRF instance for VPN1 on PE1, redistribute RIP routes to BGP, and redistribute BGP routes to RIP. The configuration on PE2 is similar.
- In the VRF instance for VPN2 on PE1, redistribute OSPF routes to BGP, and redistribute BGP routes to OSPF. The configuration on PE2 is similar.

For details about the application scenario, see "Configuration Guide > MPLS > L3 VPN".

7.2.3 VRF on CEs and PEs (MCE Application)

Scenario

An ISP provides the VPN service in an IP/MPLS backbone network.

In Figure 7- 3, VPN a runs RIP, VPN b runs OSPF, and PE1 and PE2 are connected to BGP/MPLS VPNs.

- One Multi-VPN-Instance CE (MCE) is connected to multiple VPNs. VRF tables must be created to separate the routes in VPN a from those in VPN b.
- On each PE, VRF tables must be created to separate the routes in VPN a, those in VPN b, and those in the public network from each other.

Figure 7- 3

Deployment

- One MCE1, create VRF tables for VPN a and VPN b respectively, bind the interfaces directly connected to VPN a and VPN b, and bind the VLAN interface connected to PE1. The configuration on MCE2 is similar.
- On PE1, create VRF tables for VPN a and VPN b respectively, and bind the VLAN interface connected to MCE1. The configuration on PE2 is similar.
- On MCE1, associate the VRF table for VPN a with VPN a by using RIP. The configuration on MCE2 is similar.
- On MCE1, associate the VRF table for VPN b with VPN b by using OSPF. The configuration on MCE2 is similar.
- Create a BGP neighbor (VPNv4 address family) between PE1 and PE2.
- In the VRF instance for VPN a on MCE1, redistribute RIP routes to BGP, and redistribute BGP routes to RIP. The configuration on MCE2 is similar.
- In the VRF instance for VPN b on MCE1, redistribute OSPF routes to BGP, and redistribute BGP routes to OSPF. The configuration on MCE2 is similar.

For details about the application scenario, see "Configuration Guide > MPLS > L3 VPN".

7.3 Features

Overview

Feature	Description
VPN Instance	A VPN instance is used to provide the VPN service. It is typically represented by a VRF table.
VPN Route	A VPN route is used to forward VPN packets.
VPN Route Attribute	Route distinguisher (RD): Identifies the VPN to which a route belongs. RT: Indicates the route trade-off mode of VRF.

7.3.1 VPN Instance

A VPN instance is used to provide the VPN service. On a device that provides the VPN service, a VPN instance consists of the VRF table, interfaces, routing protocol processes, and configuration that belong to the same VPN. A VPN instance is typically represented by a VRF table.

Working Principle

A PE exchanges routes with a CE by using the related routing protocol in the corresponding VPN instance. A VRF table is bound to a specific interface to generate its interface set. Packets received on these interfaces will be associated with the VRF table and forwarded along corresponding routes.

Related Configuration

Single-protocol VRF tables and multiprotocol VRF tables cannot be created at the same time. Single-protocol VRF tables only support IPv4, whereas multiprotocol VRF tables support IPv4 and IPv6.

Configuring a Single-Protocol VRF Table

By default, a device has no VRF table.

Run the **ip vrf** command to create a single-protocol VRF table.

Run the **ip vrf forwarding** command to bind an interface.

Currently, single-protocol VRF tables only support IPv4.

Configuring a Multiprotocol VRF Table

By default, a device has no VRF table.

Run the **vrf definition** command to create a multiprotocol VRF table.

Run the **address-family ipv4** command to enable the IPv4 address family.

Run the **address-family ipv6** command to enable the IPv6 address family.

Run the **vrf forwarding** command to bind an interface.

Multiprotocol VRF tables support IPv4 and IPv6.

7.3.2 VPN Route

A VPN route is only used to forward VPN packets. It comes from:

- Direct route and host route on the bound interface
- Direct route and host route on the configured import interface (not bound)
- Static and dynamic routes (RIP, RIPng, OSPFv2, OSPFv3, ISIS, and BGP) in the configured VPN instance

For details about the static routes in a VPN instance, see "Configuration Guide > IP Route".

For details about RIP in a VPN instance, see "Configuration Guide > IP Route > RIP".

For details about RIPng in a VPN instance, see "Configuration Guide > IP Route > RIPng".

For details about OSPFv2 in a VPN instance, see "Configuration Guide > IP Route > OSPFv2".

For details about OSPFv3 in a VPN instance, see "Configuration Guide > IP Route > OSPFv3".

For details about ISIS in a VPN instance, see "Configuration Guide > IP Route > ISIS".

For details about BGP in a VPN instance, see "Configuration Guide > IP Route > BGP".

7.3.3 VPN Route Attribute

The BGP extended attributes include two attributes specific to VPN routes: RD and RT.

Working Principle

RD

Two routes with the same address but different RDs in two VRF tables can be advertised separately between PEs, because the routes are sent together with their RDs through multiprotocol BGP (MP-BGP).

RT

RT in essence indicates each VRF table's route trade-off and preferences. It is mainly used to control the advertising and installation policies for VPN routes. RT is divided into the import attribute and export attribute. The import attribute indicates the route of interest, and the export attribute indicates the advertised route. A PE advertises a route to other PEs based on the RT export rule in the corresponding VRF table. The peer PE checks all received routes against the RT import rule in each VRF table. If a route matches an RT export rule (the export rule contains the import rule), it will be added to the corresponding VRF table.

Related Configuration

RD

By default, no RD is configured in VRF mode.

Run the **rd** command to configure an RD.

RT

By default, no RT is configured in VRF mode or address family mode.

Run the **route-target** { **import | export | both** } command to configure an RT.

7.4 Configuration

Configuration	Description and Command	
Configuring a Single-Protocol VRF Table	Single-protocol VRF tables and multiprotocol VRF tables cannot be created at the same time. If IPv6 is supported, configure a multiprotocol VRF table; otherwise, you can configure a single-protocol VRF table or a multiprotocol VRF table. This configuration item creates a VRF table in an IPv4 network. IPv6 is not supported.	
	ip vrf *vrf-name*	Creates a VRF table.
	rd *rd_value*	Configures an RD.
	route-target { **import	export	both** } *rt_value*	Configures an RT.
	ip vrf forwarding *vrf-name*	Binds an interface and adds the direct route and host route on the interface to a VRF table.

Configuration	Description and Command	
	ip vrf receive *vrf_name*	Adds the direct route and host route on an interface to a VRF table without binding the interface.
Configuring a Multiprotocol VRF Table	Single-protocol VRF tables and multiprotocol VRF tables cannot be created at the same time. If IPv6 is supported, configure a multiprotocol VRF table. otherwise, you can configure a single-protocol VRF table or a multiprotocol VRF table. This configuration item creates a VRF table in an IPv4 or IPv6 network.	
	vrf definition *vrf-name*	Creates a VRF table.
	description *string*	Configures a VRF descriptor.
	rd *rd_value*	Configures an RD.
	route-target { **import** | **export** | **both** } *rt_value*	Configures an RT.
	address-family ipv4	Enables the IPv4 address family.
	address-family ipv6	Enables the IPv6 address family.
	vrf forwarding *vrf-name*	Binds an interface and adds the direct route and host route on the interface to a VRF table.
	vrf receive *vrf-name*	Adds the direct route and host route on an interface to a VRF table without binding the interface.

7.4.1 Configuring a Single-Protocol VRF Table

Configuration Effect

- Provide the VPN service on a device.
- With BGP assistance, flexibly control the separation and access between VPNs.
- With BGP assistance, provide the VPN service in an IP/MPLS backbone network.
- Only IPv4 is supported.

Notes

- No VRF table needs to be created if the device only forwards packets from one VPN or from the public network.
- If the device needs to forward public network packets and VPN packets or forward packets from multiple VPNs, VRF tables must be created to separate routes.
- In many cases, static or dynamic routes (RIP, OSPF, ISIS, and BGP) need to be added to VRF tables.

Configuration Steps

Creating a VRF Table

- Mandatory.
- Create a VRF table for each VPN.

Configuring an RD

- Optional.
- When routing information needs to be advertised through BGP in the backbone network, BGP may select the best route for advertising if overlapping network addresses exist in different VPNs, which will make some VPNs fail to obtain corresponding routing information. To solve this problem, you can configure RDs for routes to enable BGP to make routing decisions based on these RDs, thus ensuring that each VPN can obtain corresponding routing information.
- Run the **rd** command in single-protocol VRF mode.

Configuring an RT

- Optional.
- You can run the **route-target export** command to specify the attributes of the route to be advertised, and run the **route-target import** command to specify the attributes of the route to be received. You can also run the **route-target both** command to specify the export and import attributes.
- Run the **route-target** command in single-protocol VRF mode.

Binding an Interface and Adding the Direct Route and Host Route on the Interface to a VRF Table

- Mandatory.
- If the physical link for transmitting VPN packets is exclusively occupied by a VPN, bind the physical interface to the corresponding VRF table.
- If the physical link for transmitting VPN packets is shared by multiple VPNs, you need to create an independent logical link for each VPN, and bind the logical interface to the corresponding VRF table. A logical interface can be a subinterface or a VLAN interface.
- You must bind an interface to the corresponding single-protocol VRF table before you configure the IPv4 address of the interface. If you bind the interface after its IPv4 address is configured, the IPv4 address will be invalid (the IPv6 address of the interface is retained).
- If you bind an interface to the corresponding single-protocol VRF table and enable IPv6 on the interface, the device cannot forward the IPv6 packets received on the interface.

Adding the Direct Route and Host Route on an Interface to a VRF Table Without Binding the Interface

- Optional.
- If policy-based routing (PBR) is required for VRF table selection, run the **ip vrf receive** command on the interface to which PBR is applied, and import the direct route and host route on the interface to each VRF table available for choice.

Verification

- Check whether VRF tables are created correctly on the router.

Related Commands

Creating a VRF Table

Command	**ip vrf** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of the VRF table to be created. It cannot exceed 31 characters.
Command Mode	Global configuration mode

Usage Guide	After you run the command, the system will enter VRF mode.

Configuring an RD

Command	**rd** *rd_value*
Parameter Description	*rd_value* has the following three different parameter forms: (1) *rd_value* = **as_num: nn** **as_num** indicates the 2-byte number that identifies a public autonomous system (AS). **nn** is configurable in the range 0..4294967295. (2) *rd_value* = **ip_addr: nn** **ip_addr** must be a global IP address. **nn** is configurable in the range 0..65535. (3) *rd_value* = **as4_num: nn** **as4_num** indicates the 4-byte number that identifies a public AS. **nn** is configurable in the range 1..65535.
Command Mode	VRF configuration mode
Usage Guide	You cannot directly change the RD of an existing VRF table. You need to delete the VRF table first and then configure a new RD. A VRF table has only one RD. You cannot configure multiple RDs for one VRF table.

Configuring an RT

Command	**route-target** { **import | export | both** } *rt_value*
Parameter Description	*rt_value* has the following three different parameter forms: (1) *rt_value* = **as_num: nn** **as_num** indicates the 2-byte number that identifies a public AS. **nn** is configurable in the range 0..4294967295. (2) *rt_value* = **ip_addr: nn** **ip_addr** must be a global IP address. **nn** is configurable in the range 0..65535. (3) *rt_value* = **as4_num: nn** a**s4_num** indicates the 4-byte number that identifies a public AS. **nn** is configurable in the range 1..65535.
Command Mode	VRF configuration mode
Usage Guide	A VRF table can be configured with multiple import and export RT attributes.

Binding an Interface

Command	**ip vrf forwarding** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of a VRF table.
Command Mode	Interface configuration mode
Usage Guide	By default, an interface does not belong to any VRF table. After an interface is bound to the corresponding VRF table, the direct route and host route on the interface will be automatically added to the VRF table. You must bind an interface to the corresponding single-protocol VRF table before you configure the IPv4 address of the interface. If you bind the interface after its IPv4 address is configured, the IPv4 address will be invalid (the IPv6 address of

	the interface is retained).

Adding the Direct Route and Host Route on an Interface to a VRF Table Without Binding the Interface

Command	**ip vrf receive** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of a VRF table.
Command Mode	Interface configuration mode
Usage Guide	This command is used to add the host route and direct route on an interface to a VRF table. If you need to add the host route and direct route on an interface to multiple VRF tables, run the command multiple times. Different from the **ip vrf forwarding** command, the **ip vrf receive** command does not bind an interface to the corresponding VRF table. The interface is still a global interface and does not belong to any VRF table. The **ip vrf forwarding** and **ip vrf receive** commands are mutually exclusive on the same interface.

Displaying the VRF Information on a Device

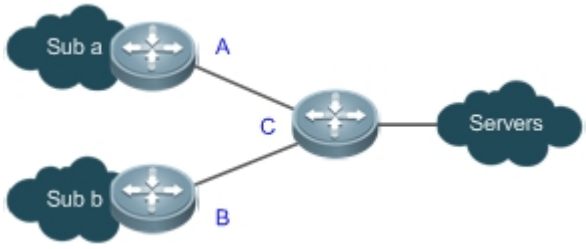
Command	**show ip vrf** [**brief** | **detail** | **interfaces**]
Parameter Description	**brief**: Displays brief information. **detail**: Displays detailed information. **interfaces**: Displays the interface binding information.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	This command is used to display the information of a specified VRF table to check whether the VRF table is bound with the correct interface.

Displaying the Routes in a VRF Table

Command	**show ip route vrf** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of a VRF table.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	This command is used to check whether a specified VRF table contains corresponding routes.

Configuration Example

Local Inter-VPN Access

Scenario **Figure 7- 4**	
	Sub a, Sub b, and Servers are three VPNs that have separate address spaces.

	Sub a runs RIP, Sub b runs OSPF, and Servers is a network segment directly connected to C.
Configuration Requirements	Routes in Sub a are separated from those in Sub b, but both Sub a and Sub b can access Servers.
Configuration Steps	● On C, create a VRF table for Sub a, bind the interface directly connected to A, and associate the VRF table with A by using RIP. ● On C, create a VRF table for Sub b, bind the interface directly connected to B, and associate the VRF table with B by using OSPF. ● On C, create a VRF table for Servers and bind the interface directly connected to Servers. ● On C, configure the RTs of the VRF tables for Sub a, Sub b, and Servers. Import the routes in the VRF tables for Sub a and Sub b to the VRF table for Servers, and import the routes in the VRF table for Servers to the VRF tables for Sub a and Sub b. ● Configure the Border Gateway Protocol (BGP) on C. Introduce the RIP routes to the VRF table for Sub a, introduce the OSPF routes to the VRF table for Sub b (enabled with an address family), and introduce the direct routes to the VRF table for Servers (enabled with an address family). Planning of interfaces and addresses:

Interface Description	Interface Name	IP Address/Mask	VRF Table
Interface on C connected to A	GE0/1	10.10.1.1/24	VRF table for Sub a
Interface on C connected to B	GE0/2	10.10.2.1/24	VRF table for Sub b
Interface on C connected to Servers	GE0/3	10.10.3.1/24	VRF table for Servers
Interface on A connected to C	GE0/1	10.10.1.2/24	-
Interface on B connected to C	GE0/2	10.10.2.2/24	-

A

```
A(config)#interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#no switchport port
A(config-if-GigabitEthernet 0/1)#ip address 10.10.1.2 255.255.255.0
A(config-if-GigabitEthernet 0/1)#exit
A(config)#router rip
A(config-router)#version 2
A(config-router)#no auto-summary
A(config-router)#network 10.10.1.0 0.0.0.255
```

B

```
B(config)#interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/2)#no switchport port
B(config-if-GigabitEthernet 0/2)#ip address 10.10.2.2 255.255.255.0
B(config-if-GigabitEthernet 0/2)#exit
B(config)#router ospf 1
B(config-router)#network 10.10.2.0 0.0.0.255 area 0
```

C	C(config)# ip vrf Suba C(config-vrf)# rd 100:1 C(config-vrf)# route-target import 100:3 C(config-vrf)# route-target export 100:1 C(config-vrf)# exit C(config)#interface GigabitEthernet 0/1 C(config-GigabitEthernet 0/1)#ip vrf forwarding Suba C(config-GigabitEthernet 0/1)#ip address 10.10.1.1 255.255.255.0 C(config-GigabitEthernet 0/1)# exit C(config)#router rip C(config-router)#address-family ipv4 vrf Suba C(config-router-af)# version 2 C(config-router-af)# no auto-summary C(config-router-af)#network 10.10.1.0 0.0.0.255 C(config-router-af)#exit
	C(config)# ip vrf Subb C(config-vrf)# rd 100:2 C(config-vrf)# route-target import 100:3 C(config-vrf)# route-target export 100:2 C(config-vrf)# exit C(config)#interface gigabitEthernet 0/2 C(config-GigabitEthernet 0/2)#ip vrf forwarding Subb C(config-GigabitEthernet 0/2)# ip address 10.10.2.1 255.255.255.0 C(config-GigabitEthernet 0/2)# exit C(config)# router ospf 2 vrf Subb C(config-router)# network 10.10.2.0 0.0.0.255 area 0 C(config-router)# exit
	C(config)# ip vrf Servers C(config-vrf)# rd 100:3 C(config-vrf)# route-target import 100:1 C(config-vrf)# route-target import 100:2 C(config-vrf)# route-target export 100:3 C(config-vrf)# exit

	C(config)# interface gigabitEthernet 0/3 C(config-GigabitEthernet 0/3)# ip vrf forwarding Servers C(config-GigabitEthernet 0/3)# ip address 10.10.3.1 255.255.255.0 C(config-GigabitEthernet 0/3)# exit
	C(config)# router bgp 200 C(config-router)# address-family ipv4 vrf vpna C(config-router-af)# redistribute rip C(config-router-af)# exit C(config-router)# address-family ipv4 vrf vpnb C(config-router-af)# redistribute ospf 1 C(config-router-af)# exit C(config-router)# address-family ipv4 vrf Servers C(config-router-af)# redistribute connected subnets C(config-router-af)# exit
Verification	● Run the **show ip vrf interface** command on C to check the interface binding information. ● Run the **show ip route vrf** command on C to check whether two VRF tables are created to separate the routes in Sub a from those in Sub b and whether both VRF tables contain the routes in Servers.
C	C# show ip vrf interfaces Interface IP-Address VRF Protocol GigabitEthernet 0/1 10.10.1.1 Suba up GigabitEthernet 0/2 10.10.2.1 Subb up GigabitEthernet 0/3 10.10.3.1 Servers up
	C# show ip route vrf Subb Routing Table: Subb Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default

	Gateway of last resort is no set O 10.2.0.0/16 [20/0] via 0.0.0.0, 00:10:46, GigabitEthernet 0/2 O 10.10.2.0/24 [20/0] via 0.0.0.0, 00:10:46, GigabitEthernet 0/2 C 10.10.2.1/32 is local host. C 10.10.3.0/24 is directly connected, GigabitEthernet 0/3 C 10.10.3.1/32 is local host.

Common Errors

- An interface is bound to a VRF table after the IP interface of the interface is configured.
- When a physical link is used to forward packets from multiple VPNs, the corresponding physical interface is bound to a VRF table.
- VPN routes are not introduced to BGP.

7.4.2 Configuring a Multiprotocol VRF Table

Configuration Effect

- Provide the VPN service on a device.
- With BGP assistance, flexibly control the separation and access between VPNs.
- With BGP assistance, provide the VPN service in an IP/MPLS backbone network.
- Support IPv4 and IPv6 through address family configuration.

Notes

- No VRF table needs to be created if the device only forwards packets from one VPN or from the public network.
- If the device needs to forward public network packets and VPN packets or forward packets from multiple VPNs, VRF tables must be created to separate routes.
- In many cases, static or dynamic routes (RIP, OSPF, ISIS, and BGP) need to be added to VRF tables.

Configuration Steps

Creating a VRF Table

- Mandatory.
- Create a VRF table for each VPN.

Configuring an Address Family

- Mandatory.
- Enable the corresponding address family for each created VRF table.

Configuring an RD

- Optional.

- When routing information needs to be advertised through BGP in the backbone network, BGP may select the best route for advertising if overlapping network addresses exist in different VPNs, which will make some VPNs fail to obtain corresponding routing information. To solve this problem, you can configure RDs for routes to enable BGP to make routing decisions based on these RDs, thus ensuring that each VPN can obtain corresponding routing information.

Configuring an RT

- Optional.
- You can run the **route-target export** command to specify the attributes of the route to be advertised, and run the **route-target import** command to specify the attributes of the route to be received. You can also run the **route-target both** command to specify the export and import attributes.
- Run the **route-target** command in multiprotocol VRF mode or multiprotocol VRF address family mode.

Binding an Interface and Adding the Direct Route and Host Route on the Interface to a VRF Table

- Mandatory.
- If the physical link for transmitting VPN packets is exclusively occupied by a VPN, bind the physical interface to the corresponding VRF table.
- If the physical link for transmitting VPN packets is shared by multiple VPNs, you need to create an independent logical link for each VPN, and bind the logical interface to the corresponding VRF table. A logical interface can be a subinterface or a VLAN interface.
- Before you bind an interface to a multiprotocol VRF table, enable an address family for the table. If you do not enable the IPv4 address family in advance, you cannot configure the IPv4 address and VRRP IPv4 address of the bound interface. If you do not enable the IPv6 address family in advance, you cannot configure the IPv6 address and VRRP IPv6 address of the bound interface.
- You must bind an interface to the corresponding multiprotocol VRF table before you configure the IPv4 or IPv6 address of the interface. If you bind the interface after its IPv4 or IPv6 address is configured, the address will be invalid.

Adding the Direct Route and Host Route on an Interface to a VRF Table Without Binding the Interface

- Optional.
- If PBR is required for VRF table selection, run the **ip vrf receive** command on the interface to which PBR is applied, and import the direct route and host route on the interface to each VRF table available for choice.

Verification

- Check whether multiprotocol VRF tables are created correctly on the router and corresponding address families are enabled.

Related Commands

Creating a VRF Table

Command	**vrf definition** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of the VRF table to be created. It cannot exceed 31 characters.
Command Mode	Global configuration mode
Usage Guide	After you run the command, the system will enter VRF mode.

Enabling the IPv4 Address Family

Command	**address-family ipv4**
Parameter Description	N/A
Command Mode	VRF mode
Usage Guide	After you run the command, the system will enter VRF IPv4 address family submode.

Enabling the IPv6 Address Family

Command	**address-family ipv6**
Parameter Description	N/A
Command Mode	VRF mode
Usage Guide	After you run the command, the system will enter VRF IPv6 address family submode.

Configuring an RD

Command	**rd** *rd_value*
Parameter Description	*rd_value* has the following three different parameter forms: (1) *rd_value* = **as_num: nn** **as_num** indicates the 2-byte number that identifies a public AS. **nn** is configurable in the range 0..4294967295. (2) *rd_value* = **ip_addr: nn** **ip_addr** must be a global IP address. **nn** is configurable in the range 0..65535. (3) *rd_value* = **as4_num: nn** **as4_num** indicates the 4-byte number that identifies a public AS. **nn** is configurable in the range 1..65535.
Command Mode	VRF configuration mode
Usage Guide	You cannot directly change the RD of an existing VRF table. You need to delete the VRF table first and then configure a new RD. A VRF table has only one RD. You cannot configure multiple RDs for one VRF table.

Configuring an RT

Command	**route-target** { **import** | **export** | **both** } *rt_value*
Parameter Description	*rt_value* has the following three different parameter forms: (1) *rt_value* = **as_num: nn** **as_num** indicates the 2-byte number that identifies a public AS. **nn** is configurable in the range 0..4294967295. (2) *rt_value* = **ip_addr: nn** **ip_addr** must be a global IP address. **nn** is configurable in the range 0..65535. (3) *rt_value* = **as4_num: nn** a**s4_num** indicates the 4-byte number that identifies a public AS. **nn** is configurable in the range 1..65535.
Command	VRF configuration mode or VRF address family submode

Mode	
Usage Guide	One VRF table can be configured with multiple import and export RT attributes.

Binding an Interface

Command	**vrf forwarding** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of a VRF table.
Command Mode	Interface configuration mode
Usage Guide	By default, an interface does not belong to any VRF table. After an interface is bound to the corresponding VRF table, the direct route and host route on the interface will be automatically added to the VRF table. Before you bind an interface to a multiprotocol VRF table, enable an address family for the table. If you do not enable the IPv4 address family in advance, you cannot configure the IPv4 address and VRRP IPv4 address of the bound interface. If you do not enable the IPv6 address family in advance, you cannot configure the IPv6 address and VRRP IPv6 address of the bound interface. You must bind an interface to a multiprotocol VRF table before you configure the IPv4, IPv6, VRRP IPv4, and VRRP IPv6 addresses of the interface; otherwise, these addresses will be invalid and the IPv6 protocol on the interface will be disabled. If the IPv4 address family is deleted from the multiprotocol VRF table, the IPv4 and VRRP IPv4 addresses of all interfaces bound to the VRF table will be deleted, and the IPv4 static routes in the VRF table or next-hop routes are also deleted. If the IPv6 address family is deleted from the multiprotocol VRF table, the IPv6 and VRRP IPv6 addresses of all interfaces bound to the VRF table will be deleted, the IPv6 protocol on the interfaces will be disabled, and the IPv6 static routes in the VRF table or next-hop routes are also deleted.

Adding the Direct Route and Host Route on an Interface to a VRF Table Without Binding the Interface

Command	**vrf receive** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of a VRF table.
Command Mode	Interface configuration mode
Usage Guide	This command is used to add the host route and direct route on an interface to a VRF table. If you need to add the host route and direct route on an interface to multiple VRF tables, run the command multiple times. Different from the **vrf forwarding** command, the **vrf receive** command does not bind an interface to the corresponding VRF table. The interface is still a global interface and does not belong to any VRF table. The **vrf forwarding** and **vrf receive** commands are mutually exclusive on the same interface.

Displaying the VRF Information on a Device

Command	**show vrf** [**brief** | **detail** | **ipv4** | **ipv6**]
Parameter Description	**brief**: Displays brief information. **detail**: Displays detailed information. **ipv4**: Displays the brief information of an IPv4 VRF table.

	Ipv6: Displays the brief information of an IPv6 VRF table.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	This command is used to display the information of a specified VRF table to check whether the VRF table is bound with the correct interface.

Displaying the Routes in a VRF Table

Command	**show ip route vrf** *vrf-name*
Parameter Description	*vrf-name*: Indicates the name of a VRF table.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	This command is used to check whether a specified VRF table contains corresponding routes.

Configuration Example

The following example only describes VRF-related configuration on A1, B1, MCE1, and PE1. The configuration on A2, B2, MCE2, and PE2 is similar.

VRF on CEs and PEs (MCE Application)

Scenario **Figure 7- 5**	
	VPN a and VPN b have independent address spaces. VPN a runs RIP and VPN b runs OSPF.
Configuration Requirements	The routes in VPN a are separated from those in VPN b. A1 and A2 can access each other, and B1 and B2 can access each other.
Configuration Steps	● Connect MCE1 and A1 through RIP. Extend RIP routes on A1. On MCE1, create a VRF table for VPN a, bind the directly connected interface, and configure RIP routes. ● Connect MCE1 and B1 through OSPF. Extend OSPF routes on B1. On MCE1, create a VRF table for VPN b, bind the directly connected interface, and configure OSPF routes. ● Connect MCE1 and PE1 through BGP. On MCE1 and PE1, create a VRF table for each VPN, bind the VLAN interface, and configure BGP routes. ● Configure the physical link between MCE1 and PE1 in Trunk mode. ● In the VRF instance for VPN a on MCE1, redistribute the RIP routes to BGP, and redistribute the BGP routes to RIP. ● In the VRF instance for VPN b on MCE1, redistribute the OSPF routes to BGP, and redistribute the BGP routes to OSPF.

Planning of interfaces and addresses:

Interface Description	Interface Name	IP Address/Mask	VRF Table
Physical interface on A1 connected to MCE1	GE0/1	10.10.1.2/24	-
Physical interface on B1 connected to MCE1	GE0/2	10.10.2.2/24	-
Physical interface on MCE1 connected to A1	GE0/1	10.10.1.1/24	VRF table for VPN a
Physical interface on MCE1 connected to B1	GE0/2	10.10.2.1/24	VRF table for VPN b
Logical interface on MCE1 connected to PE1	VLAN10	10.10.10.1/24	VRF table for VPN a
Logical interface on MCE1 connected to PE1	VLAN20	10.10.20.1/24	VRF table for VPN b
Logical interface on PE1 connected to MCE1	VLAN10	10.10.10.2/24	VRF table for VPN a
Logical interface on PE1 connected to MCE1	VLAN20	10.10.20.2/24	VRF table for VPN b

A1

```
A1(config)#interface GigabitEthernet 0/1
A1(config-if-GigabitEthernet 0/1)#no switchport port
A1(config-if-GigabitEthernet 0/1)#ip address 10.10.1.2 255.255.255.0
A1(config-if-GigabitEthernet 0/1)#exit
A1(config)#router rip
A1(config-router)#version 2
A1(config-router)#no auto-summary
A1(config-router)#network 10.10.1.0 0.0.0.255
```

B1

```
B1(config)#interface GigabitEthernet 0/2
B1(config-if-GigabitEthernet 0/1)#no switchport port
B1(config-if-GigabitEthernet 0/1)#ip address 10.10.2.2 255.255.255.0
B1(config-if-GigabitEthernet 0/1)#exit
B1(config)#router ospf 1
B1(config-router)#network 10.10.2.0 0.0.0.255 area 0
```

MCE1

#Create a VRF table for VPN a and a VRF table VPN b, and enable the IPv4 address family.

```
MCE1(config)#vrf definition vpna
MCE1(config-vrf)#address-family ipv4
```

	MCE1(config-vrf-af)#exit MCE1(config-vrf)#exit MCE1(config)#vrf definition vpnb MCE1(config-vrf)#address-family ipv4 MCE1(config-vrf-af)#exit MCE1(config-vrf)#exit
	#Bind interfaces to the VRF tables. MCE1(config)#interface GigabitEthernet 0/1 MCE1(config-if-GigabitEthernet 0/1)#no switchport port MCE1(config-if-GigabitEthernet 0/1)#vrf forwarding vpna MCE1(config-if-GigabitEthernet 0/1)#ip address 10.10.1.1 255.255.255.0 MCE1(config-if-GigabitEthernet 0/1)#exit MCE1(config)#interface GigabitEthernet 0/2 MCE1(config-if-GigabitEthernet 0/2)#no switchport port MCE1(config-if-GigabitEthernet 0/2)#vrf forwarding vpnb MCE1(config-if-GigabitEthernet 0/2)#ip address 10.10.2.1 255.255.255.0 MCE1(config-if-GigabitEthernet 0/2)#exit MCE1(config)#interface vlan 10 MCE1(config-if-VLAN 10)#vrf forwarding vpna MCE1(config-if-VLAN 10)#ip address 10.10.10.1 255.255.255.0 MCE1(config-if-VLAN 10)#exit MCE1(config)#interface vlan 20 MCE1(config-if-VLAN 20)#vrf forwarding vpnb MCE1(config-if-VLAN 20)#ip address 10.10.20.1 255.255.255.0 MCE1(config-if-VLAN 20)#exit
	#Configure the interface connected to PE1 in Trunk mode. MCE1(config)#interface GigabitEthernet 0/3 MCE1(config-if-GigabitEthernet 0/3)#switchport mode trunk MCE1(config-if-GigabitEthernet 0/3)#exit
	#Configure RIP and BGP routes in the VRF table for VPN a, and introduce routes in the two VRF tables to each other. MCE1(config)#router rip MCE1(config-router)#address-family ipv4 vrf vpna

	MCE1(config-router-af)# version 2 MCE1(config-router-af)# no auto-summary MCE1(config-router-af)#network 10.10.1.0 0.0.0.255 MCE1(config-router-af)#redistribute bgp subnets MCE1(config-router-af)#exit MCE1(config)# router bgp 100 MCE1(config-router)#address-family ipv4 vrf vpna MCE1(config-router-af)#neighbor 10.10.10.2 remote-as 200 MCE1(config-router-af)#redistribute rip MCE1(config-router-af)#exit
	#Configure OSPF and BGP routes in the VRF table for VPN b, and introduce routes in the two VRF tables to each other. MCE1(config)#router ospf 1 vrf vpnb MCE1(config-router)#network 10.10.2.0 0.0.0.255 area 0 MCE1(config-router)#redistribute bgp subnets MCE1(config-router)#exit MCE1(config)# router bgp 100 MCE1(config-router)#address-family ipv4 vrf vpnb MCE1(config-router-af)#neighbor 10.10.20.2 remote-as 200 MCE1(config-router-af)#redistribute ospf 1 MCE1(config-router-af)#exit
PE1	#Create a VRF table for VPN a and a VRF table VPN b, and enable the IPv4 address family. PE1(config)#vrf definition vpna PE1(config-vrf)#address-family ipv4 PE1(config-vrf-af)#exit PE1(config-vrf)#exit PE1(config)#vrf definition vpnb PE1(config-vrf)#address-family ipv4 PE1(config-vrf-af)#exit PE1(config-vrf)#exit
	#Bind interfaces to the VRF tables. PE1(config)#vlan 10 PE1(config-vlan)#exit PE1(config)#vlan 20

	PE1(config-vlan)#exit PE1(config)#interface vlan 10 PE1(config-if-VLAN 10)#vrf forwarding vpna PE1(config-if-VLAN 10)#ip address 10.10.10.2 255.255.255.0 PE1(config-if-VLAN 10)#exit PE1(config)#interface vlan 20 PE1(config-if-VLAN 20)#vrf forwarding vpnb PE1(config-if-VLAN 20)#ip address 10.10.20.2 255.255.255.0 PE1(config-if-VLAN 20)#exit
	#Configure the interface on PE1 connected to MCE1 in Trunk mode. PE1(config)#interface GigabitEthernet 0/3 PE1(config-if-GigabitEthernet 0/3)#switchport mode trunk PE1(config-if-GigabitEthernet 0/3)#exit
	#Configure BGP routes in the VRF table for VPN a. PE1(config)# router bgp 200 PE1(config-router)#address-family ipv4 vrf vpna PE1(config-router-af)#neighbor 10.10.10.1 remote-as 100 PE1(config-router-af)#exit
	#Configure BGP routes in the VRF table for VPN b. PE1(config)# router bgp 200 PE1(config-router)#address-family ipv4 vrf vpnb PE1(config-router-af)#neighbor 10.10.20.1 remote-as 100 PE1(config-router-af)#exit
Verification	● On A1, run the **show ip route** command to display the routes in VPN a. ● On B2, run the **show ip route** command to display the routes in VPN b. ● On MCE1, run the **show ip route vrf vpna** command to display the routes in VPN a, and run the **show ip route vrf vpnb** command to display the routes in VPN b. ● On PE1, run the **show ip route vrf vpna** command to display the routes in VPN a, and run the **show ip route vrf vpnb** command to display the routes in VPN b.

Common Errors

- A multiprotocol VRF table is configured, but no address family is enabled.
- An interface is bound to a VRF table after the IP interface of the interface is configured.
- When a physical link is used to forward packets from multiple VPNs, the corresponding physical interface is bound to a VRF table.
- VPN routes are not introduced to BGP.

7.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the routes in a specified VRF table.	**clear ip route vrf** *vrf-name*

Displaying

Description	Command
Displays the information of a single-protocol VRF table.	**show ip vrf** [**brief** | **detail** | **interfaces**]
Displays the information of a multiprotocol VRF table.	**show vrf** [**ipv4** | **ipv6** | **brief** | **detail**]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Displays the debugging information during the processes where a VRF table is created, an address family is enabled, and an interface is bound to the VRF table.	**debug vrf**
Prints the information of interface-related VRF operation debugging.	**debug vrf interface**

8 Configuring RIPng

8.1 Overview

RIP next generation (RIPng) is a unicast routing protocol that applies to IPv6 networks. RIPng-enabled routers exchange routing information to obtain routes to remote networks.

As an Interior Gateway Protocol (IGP), RIPng can run only within the autonomous system (AS) and is applicable to small-sized networks with routes no more than 16 hops.

Protocols and Standards

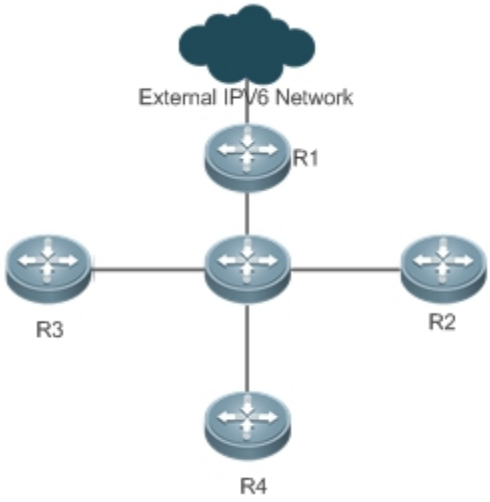
- RFC2080: Defines the RIPng.

8.2 Application

RIPng is generally used on some small-sized networks, such as office networks of small companies.

As shown in the following figure, the company builds an IPv6 network, on which all routers support IPv6. The network is small in size, but the workload is still heavy if the network is maintained manually. In this case, RIPng can be configured to adapt to topological changes of the small-sized network, which reduces the workload.

Figure 8-1



8.3 Features

Basic Concepts

IGP and EGP

IGP runs within an AS. For example, RIPng is a type of IGP.

Exterior Gateway Protocol (EGP) runs between ASs. For example, BGP is a type of EGP.

Feature

Feature	Description
RIPng and RIP	RIPng is an extension of RIPv2 on the basis of IPv6. Both are similar in functions and configurations.
Exchanging Routing Information	By exchanging routing information, RIPng-enabled devices can automatically obtain routes to a remote network and update routes in real time.
Routing Algorithm	RIPng is a protocol based on the distance-vector algorithm. It uses the vector addition method to compute the routing information.
Avoiding Route Loops	RIPng uses functions, such as split horizon and poison reverse, to avoid route loops.

8.3.1 RIPng and RIP

RIP applies to IPv4 networks. Two RIP versions are available, including RIPv1 and RIPv2.

RIPng is an extension of RIPv2 on the basis of IPv6. Both are similar in functions and configurations.

Working Principle

RIPv2

RIPv2 packets are multicast. The multicast address is 224.0.0.9, and the UDP port ID is 520. RIPv2 can identify the subnet mask.

RIPng

RIPng packets are multicast. The multicast address is FF02::9, the source address is FE80::/10, and the UDP port ID is 521. RIPng can identify the subnet mask.

This chapter describes functions and configurations of RIPng. For details about RIPv2, see "Configuring RIP".

Related Configuration

Enabling the RIPng Process

By default, the RIPng process is disabled.

Run the **ipv6 router rip** command to enable the RIPng process.

You must enable the RIPng process on a device; otherwise, all functions related to RIPng cannot take effect.

Running RIPng on an Interface

By default, RIPng does not run on an interface.

Run the **ipv6 rip enable** command to run RIPng on an interface.

After RIPng runs on an interface, RIPng packets can be exchanged on the interface and RIPng can learn routes to the network segments directly connected to the device.

Prohibiting an Interface from Sending or Receiving Packets

By default, a RIPng-enabled interface is allowed to send and receive RIPng packets.

Run the **passive-interface** command to prohibit an interface from sending RIPng packets.

8.3.2 Exchanging Routing Information

Compared with static routing, the dynamic routing protocol has a significant advantage, that is, by exchanging routing information, devices can automatically obtain routes to a remote network and update the routes in real time.

Working Principle

Initialization

After RIPng is enabled on a router, the router sends a request packet to its neighbor router, requesting for all routing information, that is, the routing table. After receiving the request message, the neighbor router returns a response packet containing the local routing table. After receiving the response packet, the router updates the local routing table, and sends an update packet to the neighbor router, informing the neighbor router of the route update information. After receiving the update packet, the neighbor router updates the local routing table, and sends the update packet to other adjacent routers. After a series of updates, all routers can obtain and retain the latest routing information.

Periodical Update

By default, periodical update is enabled for RIPng. Adjacent routers exchange complete routing information with each other every 30s (update timer), that is, the entire routing table is sent to neighbor routers.

For every non-local route, if the route is not updated within 180s (invalid timer), the metric of the route is changed to 16 (unreachable). If the route is still not updated in the next 120s (flush timer), the route is deleted from the routing table.

Default Route

In the routing table, a route to the destination network ::/0 is called default route.
The default route can be learned from a neighbor router, or sent to a neighbor router.

Route Redistribution

For RIPng, other types of routes (such as direct routes, static routes, and routes of other routing protocols) are called external routes. External routes (excluding the default route) can be redistributed to RIPng and advertised to neighbors.

Route Filtering

Filtering conditions can be configured to limit the routing information exchanged between adjacent routers. Only the routing information that meets filtering conditions can be sent or received.

Related Configuration

RIPng Timers

By default, the update timer is 30s, the invalid timer is 180s, and the flush timer is 120s.
Run the **timers basic** command to modify durations of RIPng timers.
Increasing the duration of the flush timer can reduce the route flapping. Decreasing the duration of the flush timer helps accelerate route convergence.
The durations of RIPng timers must be consistent on adjacent routers. Unless otherwise required, you are advised not to modify the RIPng timers.

Default Route

Run the **ipv6 rip default-information** command to advertise the default route to neighbors on an interface.

Route Redistribution

Run the **redistribute** command to redistribute external routes (excluding the default route) to RIPng and advertise them to neighbors.

Route Filtering

Run the **distribute-list out** command to set filtering rules to limit the routing information sent by the device.
Run the **distribute-list in** command to set filtering rules to limit the routing information received by the device.

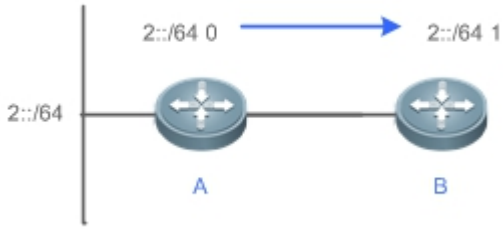
8.3.3 Routing Algorithm

RIPng is a protocol based on the distance-vector algorithm. It uses the vector addition method to compute the routing information.

Working Principle

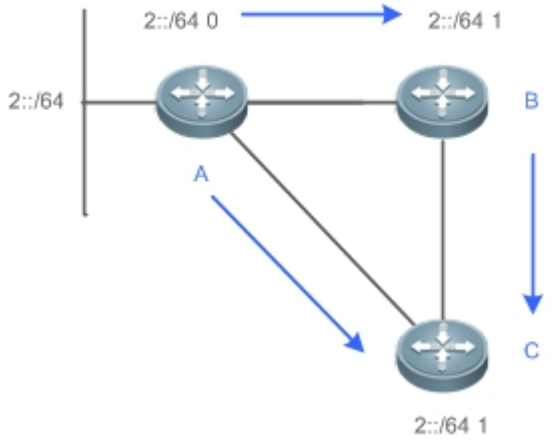
Distance-Vector Algorithm

RIPng is a protocol based on the distance-vector algorithm. The distance-vector algorithm treats a route as a vector that consists of the destination network and distance (metric). The router obtains a route from its neighbor and adds the distance vector from itself to the neighbor to the route to form its own route.
RIPng uses the hop count to evaluate the distance (metric) to the destination network. By default, the hop count from a router to its directly connected network is 0, the hop count from a router to a network that can be reached through a router is 1, and so on. That is, the metric is equal to the number of routers from the local network to the destination network. To restrict the convergence time, RIPng stipulates that the metric must be an integer between 0 and 15. If the metric is equal to or greater than 16, the destination network or host is unreachable. For this reason, RIPng cannot be applied to a large-scale network.
As shown in the following figure, Router A is connected to the network 2::/64. Router B obtains the route (2::/64, 0) from Router A and adds the metric 1 to the route to obtain its own route (2::/64, 1), and the next hop points to Router A.
Figure 8-2



Selecting the Optimum Route

RIPng selects an optimum route based on the following principle: If multiple routes to the same destination network is available, a router preferentially selects the route with the smallest metric.
As shown in the following figure, Router A is connected to the network 2::/64. Router C obtains the route (2::/64, 0) from Router A and the route (2::/64, 1) from Router B. Router C will select the route that is obtained from Router A and add metric 1 to this route to form its own route (2::/64, 1), and the next hop points to Router A.
Figure 8-3



When routes coming from different sources exist on a router, the route with the smaller distance is preferentially selected.

Route Source	Default Distance
Directly-connected network	0
Static route	1
OSPF route	110
RIPng route	120
Unreachable route	255

Related Configuration

Modifying the Distance

By default, the distance of a RIPng route is 120.

Run the **distance** command to modify the distance of a RIPng route.

Modifying the Metric

For a RIPng route that is proactively discovered by a device, the default metric is equal to the number of hops from the local network to the destination network. The metric offset of the interface is 1.

For a RIPng router that is manually configured (default route or redistributed route), the default metric is 1.

Run the **ipv6 rip metric-offset** command to modify the metric offset of the interface.

Run the **default-metric** command to modify the default metric of an external route (redistributed route).

Run the **redistribute** command to modify the metric of an external route (redistributed route) when advertising this route.

Run the **ipv6 rip default-information** command to modify the metric of a default route when advertising the default route.

8.3.4 Avoiding Route Loops

RIPng uses functions, such as split horizon and poison reverse, to avoid route loops.

Working Principle

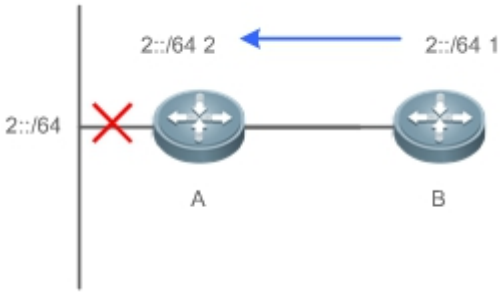
Route Loop

A RIPng route loop occurs due to inherent defects of the distance-vector algorithm.

As shown in the following figure, Router A is connected to the network 2::/64, and sends an update packet every 30s. Router B receives the route to 2::/64 from Router A every 30s. If Router A is disconnected from 2::/64, the route to 2::/64 will be deleted from the routing

table on Router A. Next time, the update packet sent by Router A no longer contains this route. As Router B does not receive an update packet related to 2::/64, Router B determines that the route to 2::/64 is valid within 180s and uses the update packet to send this route to Router A. As the route to 2::/64 does not exist on Router A, the route learned from Router B is added to the routing table. Router B determines that data can reach 2::/64 through Router A, and Router A determines that data can reach 2::/64 through Router B. In this way, a route loop is formed.

Figure 8-4

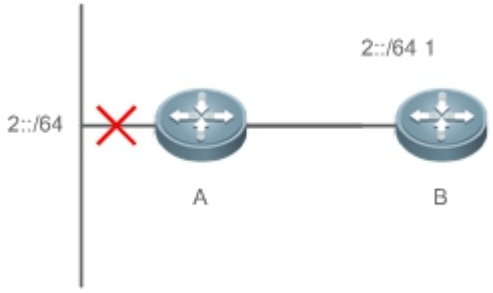


Split Horizon

Split horizon can prevent route loops. After split horizon is enabled, a route received on this interface will not be sent out from this interface.

As shown in the following figure, after split horizon is enabled on Router B, Router B will not send the route to 2::/64 back to Router A. Router B will learn 180s later that 2::/64 is not reachable.

Figure 8-5



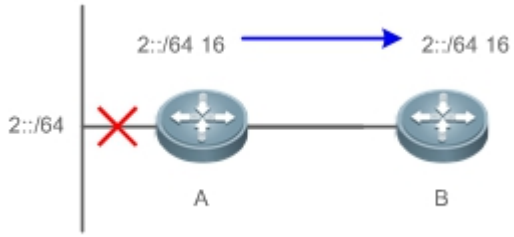
Poison Reverse

Poison reverse can also prevent route loops. Compared with slit horizon, poison reverse is more reliable, but brings more protocol packets, which makes network congestion more severe.

After poison reverse is enabled on an interface, a route received from this interface will be sent out from this interface again, but the metric of this router will be changed to 16 (unreachable).

As shown in the following figure, after poison reverse is enabled on Router A, if Route A detects a disconnection from 2::/64, Router A will not delete the route to 2::/64. Instead, Router A changes the number of hops to 16, and advertises the route through the update packet. On receiving the update packet, Router B learns that 2::/64 is not reachable.

Figure 8-6



Related Configuration

Split Horizon

By default, split horizon is enabled.

Run the **no split-horizon** command to disable split horizon.

Poison Reverse

By default, poison reverse is disabled.

Run the **split-horizon poisoned-reverse** command to enable poison reverse. (After poison reverse is enabled, split horizon is automatically disabled.)

8.4 Configuration

Configuration	Related Commands	
Configuring RIPng Basic Functions	(Mandatory) It is used to build a RIPng routing domain.	
	ipv6 router rip	Enables a RIPng routing process and enters routing process configuration mode.
	ipv6 rip enable	Runs RIPng on an interface.
	split-horizon	Enables split horizon or poison reverse.
	passive-interface	Configures a passive interface.
Advertising the Default Route or External Routes	Optional.	
	ipv6 rip default-information	Advertise the default route to neighbors on an interface.
	redistribute	Redistributes routes and advertising external routes to neighbors.
Setting Route Filtering Rules	Optional.	
	distribute-list in	Filters the received RIPng routing information.
	distribute-list out	Filters the sent RIPng routing information.
Modifying Route Selection Parameters	Optional.	
	distance	Modifies the administrative distance of a RIPng route.
	ipv6 rip metric-offset	Modifies the metric offset on an interface.
	default-metric	Configure the default metric for route redistribution.

Configuration	Related Commands	
Modifying Timers	Optional.	
	timers	Modifies the update timer, invalid timer, and flush timer of RIPng.

8.4.1 Configuring RIPng Basic Functions

Configuration Effect

- Build a RIPng routing domain on the network.
- Routers in the domain obtain routes to a remote network through RIPng.

Notes

- IPv6 addresses must be configured.
- IPv6 unicast routes must be enabled.

Configuration Steps

Enabling a RIPng Routing Process

- Mandatory.
- Unless otherwise required, perform this configuration on every router in the RIPng routing domain.

Running RIPng on an Interface

- Mandatory.
- Unless otherwise required, perform this configuration on every interconnected interface of routers in the RIPng routing domain.

Enabling Split Horizon or Poison Reverse

- By default, split horizon is enabled and poison reverse is disabled.
- Unless otherwise required, enable split horizon on every interface connected to the broadcast network, such as the Ethernet. (Retain the default setting.)
- Unless otherwise required, enable split horizon on every interface connected to the point-to-point (P2P) network, such as the PPP and HDLC. (Retain the default setting.)
- It is recommended that split horizon and poison reverse be disabled on an interface connected to a non-broadcast multi-access network, such as FR and X.25; otherwise, some devices cannot learn the complete routing information.
- If the secondary IP address is configured for an interface connected to a non-broadcast, it is recommended that split horizon and poison reverse be disabled.

Configuring a Passive Interface

- This configuration is recommended.
- Use the passive interface to set the boundary of the RIPng routing domain. The network segment of the passive interface belongs to the RIPng routing domain, but RIPng packets cannot be sent over the passive interface.
- If RIPng routes need to be exchanged on an interface (such as the router interconnect interface) in the RIPng routing domain, this interface cannot be configured as a passive interface.

Verification

- Check the routing table on a router to verify that the route to a remote network can be obtained through RIPng.

Related Commands

Enabling a RIPng Routing Process

Command	**ipv6 router rip**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	This command is used to create a RIPng routing process and enter routing process configuration mode.

Running RIPng on an Interface

Command	**ipv6 rip enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The configuration for running the RIPng on an interface is different from that of RIPv2. In RIPv2, the **network** command is configured in routing process configuration mode to define an IP address range. If the IP address of an interface belongs to this IP address range, RIP automatically runs on this interface.

Enabling Split Horizon

Command	**split-horizon** [**poisoned-reverse**]
Parameter Description	**poisoned-reverse**: Indicates that the split horizon function contains the poison reverse function.
Command Mode	Routing process configuration mode
Usage Guide	Run the **show ipv6 rip** command to check whether split horizon is enabled. The configuration is different from that of RIPv2. In RIPv2, the split horizon function is configured in interface configuration mode.

Configuring a Passive Interface

Command	**passive-interface** { **default** | *interface-type interface-num* }
Parameter Description	**default:** Indicates all interfaces. ***interface-type interface-num*:** Specifies an interface.
Command Mode	Routing process configuration mode
Usage Guide	First, run the **passive-interface default** command to configure all interfaces as passive interfaces. Then, run the **no passive-interface** *interface-type interface-num* command so that the interfaces used for interconnection between routers in the domain are not passive interface.

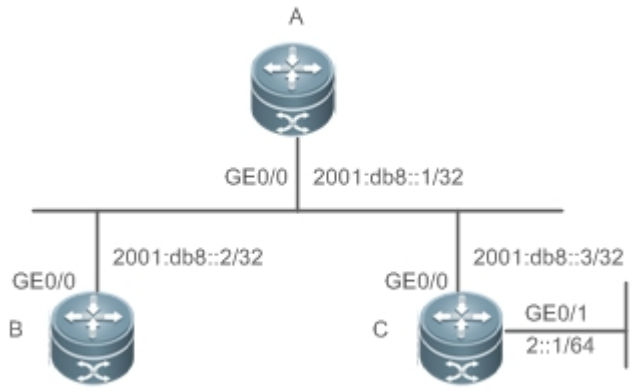
Displaying the IP Routing Table

Command	**show ipv6 route**

Parameter Description	N/A
Command Mode	Privileged EXEC mode or global configuration mode
Usage Guide	Check whether the routing table contains any route to a remote network that is learned through RIPng.

Configuration Example

Building a RIPng Routing Domain

Scenario Figure 8- 7	
Configuration Steps	● Configure IPv6 addresses on all routers. ● Enable RIPng on all routers.
A	A# configure terminal Enter configuration commands, one per line. End with CNTL/Z. A(config)# ipv6 router rip A(config-router)# exit A(config)# interface GigabitEthernet 0/0 A(config-if-GigabitEthernet 0/0)# ipv6 address 2001:db8::1/32 A(config-if-GigabitEthernet 0/0)# ipv6 rip enable
B	B# configure terminal Enter configuration commands, one per line. End with CNTL/Z. B(config)# ipv6 router rip B(config-router)# exit B(config)# interface GigabitEthernet 0/0 B(config-if-GigabitEthernet 0/0)# ipv6 address 2001:db8::2/32 B(config-if-GigabitEthernet 0/0)# ipv6 rip enable
C	C# configure terminal

	Enter configuration commands, one per line. End with CNTL/Z. C(config)# ipv6 router rip C(config-router)# exit C(config)# interface GigabitEthernet 0/0 C(config-if-GigabitEthernet 0/0)# C(config-if-GigabitEthernet 0/0)# ipv6 address 2001:db8::3/32 C(config-if-GigabitEthernet 0/0)# ipv6 rip enable C(config)# interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# ipv6 address 2::1/64 C(config-if-GigabitEthernet 0/1)# ipv6 rip enable
Verification	Check the routing tables on Router A, Router B, and Router C. The routing tables should contain routes to a remote network that are learned through RIPng.
A	A# show ipv6 route IPv6 routing table name - Default - 6 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area R 2::/64 [120/2] via FE80::2D0:F8FF:FEFB:D521, GigabitEthernet 0/0 C 2001:DB8::/32 via GigabitEthernet 0/0, directly connected L 2001:DB8::1/128 via GigabitEthernet 0/0, local host C FE80::/10 via ::1, Null0 C FE80::/64 via GigabitEthernet 0/0, directly connected L FE80::2D0:F8FF:FEFB:E7CE/128 via GigabitEthernet 0/0, local host
B	B# show ipv6 route IPv6 routing table name - Default - 6 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route

	N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area R 2::/64 [120/2] via FE80::2D0:F8FF:FEFB:D521, GigabitEthernet 0/0 C 2001:DB8::/32 via GigabitEthernet 0/0, directly connected L 2001:DB8::2/128 via GigabitEthernet 0/0, local host C FE80::/64 via GigabitEthernet 0/0, directly connected L FE80::2D0:F8FF:FEFB:C9BA/128 via GigabitEthernet 0/0, local host
C	FS# show ipv6 route IPv6 routing table name - Default - 9 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS, V - Overflow route N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area C 2::/64 via GigabitEthernet 0/1, directly connected L 2::2/128 via GigabitEthernet 0/1, local host C 2001:DB8::/32 via GigabitEthernet 0/0, directly connected L 2001:DB8::3/128 via GigabitEthernet 0/0, local host C FE80::/10 via ::1, Null0 C FE80::/64 via GigabitEthernet 0/0, directly connected L FE80::2D0:F8FF:FEFB:D521/128 via GigabitEthernet 0/0, local host C FE80::/64 via GigabitEthernet 0/1, directly connected L FE80::2D0:F8FF:FEFB:D521/128 via GigabitEthernet 0/1, local host

Common Errors

- The IPv6 address is not configured on an interface.
- The interface used for interconnection between devices is configured as a passive interface.

8.4.2 Advertising the Default Route or External Routes

Configuration Effect

- In the RIPng domain, introduce a unicast route of another AS so that the unicast routing service to this AS can be provided for users in the RIPng domain.
- In the RIPng domain, inject a default route to another AS so that the unicast routing service to this AS can be provided for users in the RIPng domain.

Notes

- The RIPng basic functions must be configured.

Configuration Steps

Configuring External Route Redistribution

- Optional.
- Perform this configuration if external routes of the RIPng domain should be introduced to the AS border router (ASBR).

Generating a Default Route

- Optional.
- Perform this configuration if the default route should be introduced to an ASBR so that other routers in the RIPng domain access other AS domains through this ASBR by default.

Verification

- Run the **show ipv6 route rip** command on a non-ASBR to check whether the external routes of the domain and default route have been loaded.

Related Commands

Advertising the Default Route to Neighbors on an Interface

Command	**ipv6 rip default-information** { **only**|**originate** } [**metric** *metric-value*]
Parameter Description	**only**: Advertises only IPv6 default route. **originate**: Advertises the IPv6 default route and other routes. **metric** *metric-value*: Indicates the metric of the default route. The value ranges from 1 to 15. The default value is 1.
Command Mode	Interface configuration mode
Usage Guide	After this command is configured on the interface, an IPv6 default route is advertised to the external devices through this interface, but the route itself is not added to the route forwarding table or the device and the RIPng route database. To prevent occurrence of a route loop, once this command is configured on an interface, RIPng refuses to receive the default route updates advertised by neighbors.

Redistributing Routes and Advertising External Routes to Neighbors

Command	**redistribute** { **connected** | **ospf** *process-id* | **static** } [**metric** *metric-value* | **route-map** *route-map-name*]
Parameter	**Connected**: Indicates redistribution from direct routes.

Description	**ospf** *process-id*: Indicates redistribution from OSPF. *process-id* indicates the OSPF process ID. The value ranges from 1 to 65535. **static**: Indicates redistribution from static routes. **metric** *metric-value*: Sets the metric of the route redistributed to the RIPng domain. **route-map** *route-map-name*: Sets the redistribution filtering rules.
Command Mode	Routing process configuration mode
Usage Guide	During route redistribution, it is not necessary to convert the metric of one routing protocol to the metric of another routing protocol because different routing protocols use completely different metric measurement methods. RIP measures the metric based on the hop count, and OSPF measures the metric based on the bandwidth. Therefore, the computed metrics cannot be compared with each other.

Configuration Example

Scenario	
Configuration Steps	● Configure the interface IPv6 addresses on all routers. (Omitted) ● Configure the RIPng basic functions on all routers. (Omitted) ● On Router B, configure redistribution of static routes. ● On the GE0/1 interface of Router A, configure advertisement of the default route.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ipv6 rip default-information originate
B	B# configure terminal B(config)# ipv6 router rip B(config-router)# redistribute static
Verification	● Check the routing tables on Router A and Router B, and confirm that Router A can learn the route 3001:10:10::/64, and Router B can learn the default route ::/0.

A	A# show ipv6 route rip IPv6 routing table name - Default - 17 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area R 3001:10:10::/64 [120/2] via FE80::2D0:F8FF:FE22:334A, GigabitEthernet 0/1
B	B# show ipv6 route rip IPv6 routing table name - Default - 17 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area R ::/0 [120/2] via FE80::21A:A9FF:FE41:5B06, GigabitEthernet 0/1

8.4.3 Setting Route Filtering Rules

Configuration Effect

- Routes that do not meet filtering criteria cannot be loaded to the routing table, or advertised to neighbors. In this way, users within the network can be prevented from accessing specified destination networks.

Notes

- The RIPng basic functions must be configured.

Configuration Steps

Filtering the Received RIP Routing Information

- To refuse receiving some specified routes, you can configure the route distribution control list to process all the received route update packets. If no interface is specified, route update packets received on all interfaces will be processed.

Filtering the Sent RIP Routing Information

- If this command does not contain any optional parameter, route update advertisement control takes effect on all interfaces. If the command contains the interface parameter, route update advertisement control takes effect only on the specified interface. If the command contains other routing process parameters, route update advertisement control takes effect only on the specified routing process.

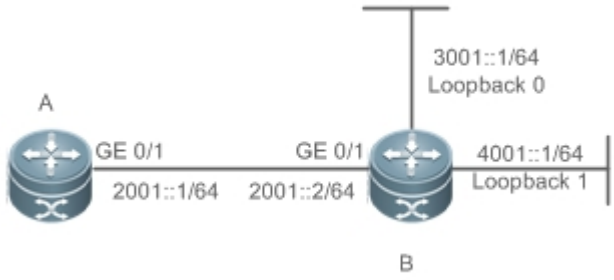
Verification

- Run the **show ipv6 route rip** command to check that the routes that have been filtered out are not loaded to the routing table.

Related Commands

Command	**distribute-list prefix-list** *prefix-list-name* { **in** | **out** } [*interface-type interface-name*]
Parameter Description	**prefix-list** *prefix-list-name:* Indicates the name of the prefix list, which is used to filter routes. **in** | **out:** Specifies update routes (received or sent routes) that are filtered. *interface-type interface-name:* Indicates that the distribution list is applied to the specified interface.
Command Mode	Routing process configuration mode
Usage Guide	N/A

Configuration Example

Scenario	
Configuration Steps	- Configure the interface IPv6 addresses on all routers. (Omitted) - Configure the RIPng basic functions on all routers. (Omitted) - On router A, configure route filtering.
A	A# configure terminal A(config)# ipv6 prefix-list hello permit 4001::/64 A(config)# ipv6 router rip A(config-router)# distribute-list prefix-list hello in
Verification	- Check that Router A can learn only the route to 4001::/64.

A	A# show ipv6 route rip IPv6 routing table name - Default - 17 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area R 4001::/64 [120/2] via FE80::2D0:F8FF:FE22:334A, GigabitEthernet 0/1

8.4.5 Modifying Route Selection Parameters

Configuration Effect

- Change the RIPng routes to enable the traffic pass through specified nodes or avoid passing through specified nodes.
- Change the sequence that a router selects various types of routes so as to change the priorities of RIPng routes.

Notes

- The RIPng basic functions must be configured.

Configuration Steps

Modifying the Administrative Distance of a RIPng Route

- Optional.
- Perform this configuration if you wish to change the priorities of RIPng routes on a router that runs multiple unicast routing protocols.

Modifying the Metric Offset on an Interface

- Optional.
- Unless otherwise required, perform this configuration on a router where the metrics of routes need to be adjusted.

Configuring the Default Metric of an External Route Redistributed to RIPng

- Optional.
- Unless otherwise required, perform this configuration on an ASBR to which external routes are introduced.

Verification

- Run the **show ipv6 rip** command to display the administrative distance of RIPng routes.
- Run the **show ipv6 rip data** command to display the metrics of external routes redistributed to RIPng.

Related Commands

Modifying the Administrative Distance of a RIPng Route

Command	**distance** *distance*
Parameter Description	*distance*: Sets the administrative distance of a RIPng route. The value is an integer ranging from 1 to 254.
Command Mode	Routing process configuration mode
Usage Guide	Run this command to set the administrative distance of a RIPng route.

Modifying the Metric Offset on an Interface

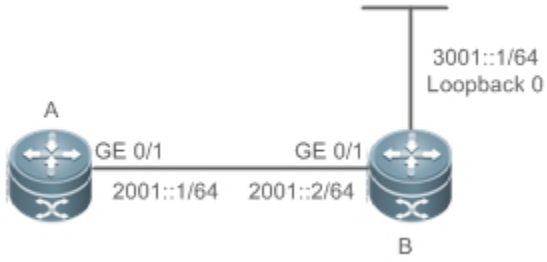
Command	**ipv6 rip metric-offset** *value*
Parameter Description	*value*: Indicates the interface metric offset. The value ranges from 1 to 16.
Command Mode	Routing process configuration mode
Usage Guide	Before a route is added to the routing table, the metric of the route must be added with the metric offset set on the interface. You can control the use of a route by setting the interface metric offset.

Configuring the Default Metric of an External Route Redistributed to RIPng

Command	**default-metric** *metric*
Parameter Description	*metric*: Indicates the default metric. The valid value ranges from 1 to 16. If the value is equal to or greater than 16, the FSOS determines that this route is unreachable.
Command Mode	Global configuration mode
Usage Guide	If the metric is not specified during redistribution of a routing protocol process, RIPng uses the metric defined by the **default-metric** command. If the metric is specified, the metric defined by the **default-metric** command is overwritten by the specified metric. If this command is not configured, the value of **default-metric** is 1.

Configuration Example

Modifying the Administrative Distance of a RIPng Route

Scenario	
Configuration Steps	● Configure the interface IPv6 addresses on all routers. (Omitted) ● Configure the RIPng basic functions on all routers. (Omitted) ● On Router A, set the administrative distance of a RIPng route to 160.
	A# configure terminal A(config)# ipv6 router rip A(config-router)# distance 160

Verification	● On Router A, check whether the administrative distance of a RIPng route is 160.
	A# show ipv6 route rip | in 3001::/64 R 3001::/64 [160/2] via FE80::2D0:F8FF:FE22:334A, GigabitEthernet 0/1

8.4.6 Modifying Timers

Configuration Effect

- Change the duration of RIPng timers to accelerate or slow down the change of the protocol state or occurrence of an event.

Notes

- The RIPng basic functions must be configured.
- Modifying the protocol control parameters may result in protocol running failures. Therefore, you are advised not to modify the timers.

Configuration Steps

Modifying the Update Timer, Invalid Timer, and Flush Timer

- Mandatory.
- Unless otherwise required, perform this configuration on a router where RIPng timers need to be modified.

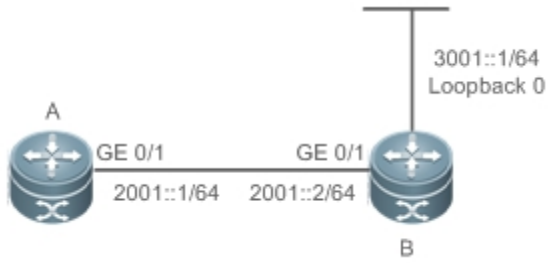
Verification

- Run the **show ipv6 rip** command to display settings of timers.

Related Commands

Command	**timers** *update invalid flush*
Parameter Description	*Update:* Indicates the route update time in second. It defines the interval at which the device sends the route update packet. Each time an update packet is received, the invalid timer and flush timer are reset. By default, a route update packet is sent every 30s. *Invalid:* Indicates the route invalid time in second, counted from the last time when a valid update packet is received. It defines the time after which the route in the routing list becomes invalid because the route is not updated. The duration of the invalid timer must be at least three times the duration of the update timer. If no update packet is received before the invalid timer expires, the corresponding route enters the invalid state. If the update packet is received before the invalid timer expires, the timer is reset. The default duration of the invalid timer is 180s. *Flush:* Indicates the route flushing time in second, counted from the time when the RIPng route enters the invalid state. When the flush timer expires, the route in the invalid state will be deleted from the routing table. The default duration of the flush timer is 120s.
Command Mode	Routing process configuration mode
Usage Guide	By default, the update timer is 30s, the invalid timer is 180s, and the flush timer is 120s.

Configuration Example

Scenario	
Configuration Steps	● Configure the interface IPv6 addresses on all routers. (Omitted) ● Configure the RIPng basic functions on all routers. (Omitted) ● On Router A, configure the update timer, invalid timer, and flush timer.
B	B# configure terminal B(config)# ipv6 router rip B(config-router)# timers 10 30 90
Verification	● On Router B, check the settings of RIPng timers.
B	B# show ipv6 rip Routing Protocol is "RIPng" Sending updates every 10 seconds with +/-50%, next due in 12 seconds Timeout after 30 seconds, garbage collect after 90 seconds Outgoing update filter list for all interface is: not set Incoming update filter list for all interface is: not set Default redistribution metric is 1 Default distance is 120 Redistribution: Redistributing protocol connected Default version control: send version 1, receive version 1 Interface Send Recv GigabitEthernet 0/1 1 1 Routing Information Sources: Gateway: fe80::2d0:f8ff:fe22:334a Distance: 120 Last Update: 00:00:02 Bad Packets: 0 Bad Routes: 0

Common Errors

- Settings of RIPng timers on devices connected to the same network are inconsistent. Consequently, routes cannot be learned properly.

8.5 Monitoring

Displaying

Description	Command
Displays information about the RIPng process.	**show ipv6 rip**
Displays the RIPng routing table.	**show ipv6 rip database**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs RIPng.	**debug ipv6 rip** [interface *interface-type interface-num* | nsm | restart

9 Managing Routes

9.1 Overview

The network service module (NSM) manages the routing table, consolidates routes sent by various routing protocols, and selects and sends preferred routes to the routing table. Routes discovered by various routing protocols are stored in the routing table. These routes are generally classified by source into three types:

- Direct route: It is the route discovered by a link-layer protocol and is also called interface route.
- Static route: It is manually configured by the network administrator. A static route is easy to configure and less demanding on the system, and therefore applicable to a small-sized network that is stable and has a simple topology. However, when the network topology changes, the static route must be manually reconfigured and cannot automatically adapt to the topological changes.
- Dynamic route: It is the route discovered by a dynamic routing protocol.

9.2 Applications

Application	Description
Basic Functions of the Static Route	Manually configure a route.
Floating Static Route	Configure a standby route in the multipath scenario.
Load Balancing Static Route	Configure load balancing static routes in the multipath scenario.

9.2.1 Basic Functions of the Static Route

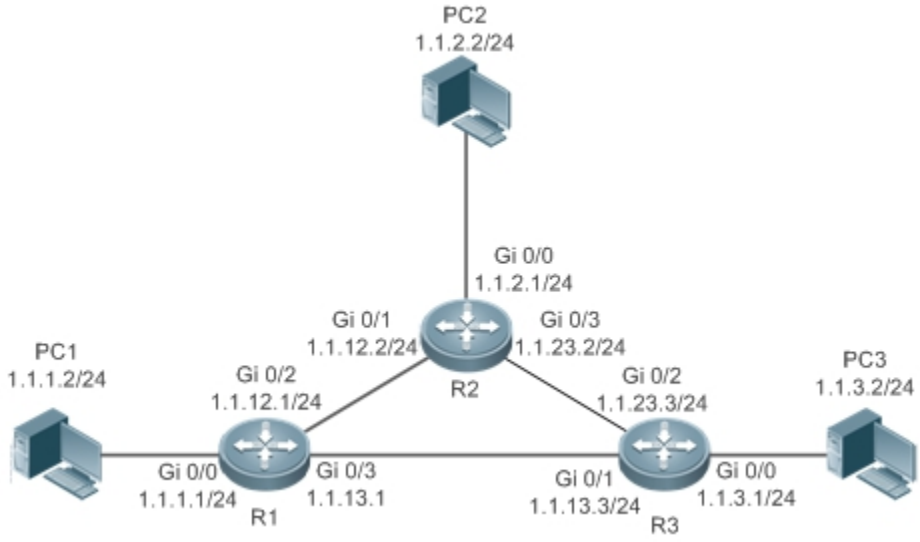
Scenario

On a network with a simple topology, you can configure only static routes to implement network interworking. Appropriate configuration and use of static routes can improve the network performance and guarantee the bandwidth for important network applications.

As shown in Figure 5- 1, to implement interworking between PC 1, PC 2, and PC 3, you can configure static routes on R 1, R 2, and R 3.

- On R 1, configure a route to the network segment of PC 2 through R 2, and a route to the network segment of PC 3 through R 3.
- On R 2, configure a route to the network segment of PC 1 through R 1, and a route to the network segment of PC 3 through R 3.
- On R 3, configure a route to the network segment of PC 1 through R 1, and a route to the network segment of PC 2 through R 2.

Figure 5- 1



Deployment

- Configure the address and subnet mask of each interface.
- Configure static routes on R 1, R 2, and R 3.

9.2.2 Floating Static Route

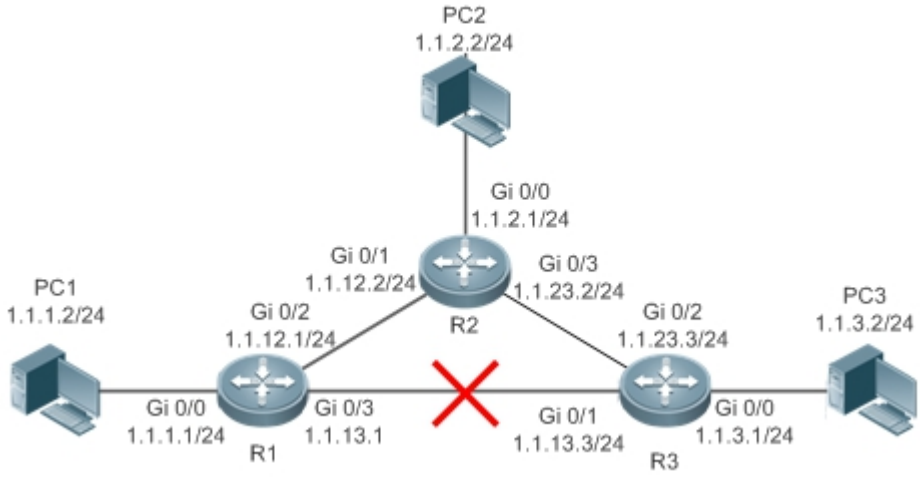
Scenario

If no dynamic routing protocol is configured, you can configure floating static routes to implement dynamic switching of routes to prevent communication interruption caused by the network connection failures.

As shown in Figure 5- 2, to prevent communication interruption caused by a line failure between R 1 and R 3, you can configure a floating static route respectively on R 1 and R 3. Normally, packets are forwarded on a path with a small administrative distance. If a link on this path is down, the route is automatically switched to the path with a large administrative distance.

- On R1, configure two routes to the network segment of PC 3, including a route through R 3 (default distance = 1) and a route through R 2 (default distance = 2).
- On R 3, configure two routes to the network segment of PC 1, including a route through R 1 (default distance = 1) and a route through R 2 (default distance = 2).

Figure 5- 2



Deployment

- Configure the address and subnet mask of each interface.
- Configure static routes on R 1, R 2, and R 3.

9.2.3 Load Balancing Static Route

Scenario

If there are multiple paths to the same destination, you can configure load balancing routes. Unlike floating routes, the administrative distances of load balancingroutes are the same. Packets are distributed among these routes based on the balanced forwarding policy.

As shown in Figure 5- 3, load balancing routes are configured respectively on R 1 and R 3 so that packets sent to the network segment of PC 3 or PC 1 are balanced between two routes, including a route through R 2 and a route through R 4.

- On R 1, configure two routes to the network segment of PC 3, including a route through R 2 and a route through R 4.
- On R 3, configure two routes to the network segment of PC 1, including a route through R 2 and a route through R 4.

Figure 5- 3

Remarks	On the switch, the load is balanced based on the source IP address by default. Run the **aggregateport load-balance** command to configure the load balancing mode of ECMP route.

Deployment

- Configure the address and subnet mask of each interface.
- Configure static routes on R 1, R 2, R 3, and R 4.
- Configure the load balancing policy on R 1 and R 3.

9.3 Features

Feature	Description
Route Computation	Generate a valid route on a device.
Optimal Route Selection	Select an optimal route to forward packets.
Default Route	Forward all packets and help reduce the size of a routing table.
Route Reliability	Quickly detect a route failure and recover communication.

9.3.1 Route Computation

Routing Function

Routing functions are classified into IPv4 and IPv6 routing functions. If the routing functions are disabled, a device is equivalent to a host and cannot forward routes.

Dynamic Route

A dynamic routing protocol learns remote routes and dynamically updates routes by exchanging routes with neighbors. If a neighbor is the next hop of a route and this neighbor fails, the route fails as well.

Static Route

On a network with a simple topology, you can configure only static routes to implement network interworking. Appropriate configuration and use of static routes can improve the network performance and guarantee the bandwidth for important network applications.

Whether a static route is active is computed based on the status of the local interface. When the exit interface of a static route is located at layer 3 (L3) and is in Up status (the link status is Up and the IP address is configured), this route is active and can be used for packet forwarding.

9.3.2 Optimal Route Selection

Administrative Distance

When multiple routing protocols generate routes to the same destination, the priorities of these routes can be determined based on the administrative distance. A smaller administrative distance indicates a higher priority.

Equal-Cost Route

If multiple routes to the same destination have different next hops but the same administrative distance, these routes are mutually equal-cost routes. Packets are distributed among these routes to implement load balancing based on the balanced forwarding policy.

On a specific device, the total number of equal-cost routes is limited. Routes beyond the limit do not participate in packet forwarding.

Floating Route

If multiple routes to the same destination have different next hops and different administrative distances, these routes are mutually floating routes. The route with the smallest administrative distance will be first selected for packet forwarding. If this route fails, a route with a larger administrative distance is further selected for forwarding, thus preventing communication interruption caused by a network line failure.

9.3.3 Default Route

In the forwarding routing table, the route with the destination network segment 0.0.0.0 and the subnet mask 0.0.0.0 is the default route. Packets that cannot be forwarded by other routes will be forwarded by the default route. The default route can be statically configured or generated by a dynamic routing protocol.

Default Gateway

On a L2 switch, the **ip default gateway** command is configured to generate a default route.

Static Default Route

On a L3 switch, a static route with the network segment 0.0.0.0 and the subnet mask 0.0.0.0 is configured to generate the default route.

Default Network

The default network is configured to generate a default route. If the **ip default-network** command is configured to specify a network (a classful network, such as a Class A, B, or C network), and this network exists in the routing table, the router will use this network as the default network and the next hop of this network is the default gateway. As the network specified by the **ip default-network** command is a classful one, if this command is used to identify a subnet in a classful network, the router automatically generates a static route of the classful network instead of any default route.

9.3.4 Route Reliability

When a device on a network is faulty, some routes become unreachable, resulting in traffic interruption. If connectivity of the next hop can be detected in real time, the route can be re-computed when a fault occurs, or traffic can be switched over to the standby route.

9.4 Configuration

Configuration Item	Description and Command	
Configuring a Static Route	(Mandatory) It is used to configure a static route entry.	
	ip route	Configures an IPv4 static route.
	ipv6 route	Configures an IPv6 static route.
Configuring a Default Route	(Optional) It is used to configure the default gateway.	
	ip default gateway	Configures an IPv4 default gateway on a L2 device.
	ipv6 default gateway	Configures an IPv6 default gateway on a L2 device.
	ip route 0.0.0.0 0.0.0.0 *gateway*	Configures an IPv4 default gateway on a L3 device.
	ipv6 route ::/0 *ipv6-gateway*	Configures an IPv6 default gateway on a L3 device.

Configuration Item	Description and Command	
	ip default network	Configures an IPv4 default network on a L3 device.
Configuring Route Limitations	(Optional) It is used to limit the number of equal-cost routes and number of static routes, or disable routing.	
	maximum-paths	Configures the maximum number of equal-cost routes.
	ip static route-limit	Configures the maximum number of IPv4 static routes.
	ipv6 static route-limit	Configures the maximum number of IPv6 static routes.
	no ip routing	Disables IPv4 routing.
	noipv6 unicast-routing	Disables IPv6 routing.

9.4.1 Configuring a Static Route

Configuration Effect

- Generate a static route in the routing table. Use the static route to forward packets to a remote network.

Notes

- Static routes cannot be configured on a L2 switch.
- If the **no ip routing** command is configured on a L3 switch, you cannot configure IPv4 static routes on this switch, and existing IPv4 static routes will also be deleted. Before the device is restarted, reconfiguring the **ip routing** command can recover the deleted IPv4 static routes. After the device is restarted, deleted IPv4 static routes cannot be recovered.
- If the **no ipv6 unicast- routing** command is configured on a L3 switch, you cannot configure IPv6 static routes on this switch, and existing IPv6 static routes will also be deleted. Before the device is restarted, reconfiguring the **ipv6 unicast- routing** command can recover the deleted IPv6 static routes. After the device is restarted, deleted IPv6 static routes cannot be recovered.

Configuration Steps

Configuring a Static IPv4 Route

Configure the following command on an IPv4-enabled router.

Command	**ip route** *network net-mask* { *ip-address* | *interface* [*ip-address*] } [*distance*] [**tag** *tag*] [**permanent |**] [**weight** *number*] [**description** *description-text*] [**disabled** | **enabled**]	
Parameter Description	*network*	Indicates the address of the destination network.
	net-mask	Indicates the mask of the destination network.
	ip-address	(Optional) Indicates the next-hop address of the static route. You must specify at least one of *ip-address* and *interface*, or both of them. If *ip-address* is not specified, a static direct route is configured.
	interface	(Optional) Indicates the next-hop exit interface of the static route. You must specify at least one of *ip-address* and *interface*, or both of them. If *interface* is not specified, a recursive static direct route is configured. The exit interface is obtained by the next hop in the routing table.

	distance	(Optional) Indicates the administrative distance of the static route. The administrative distance is 1 by default.
	tag	(Optional) Indicates the tag of the static route. The tag is 0 by default.
	permanent	(Optional) Indicates the flag of the permanent route. The static route is not a permanent route by default.
	weight *number*	(Optional) Indicates the weight of the static route. The weight is 1 by default.
	description*description-text*	(Optional) Indicates the description of the static route. By default, no description is configured. *description-text* is a string of one to 60 characters.
	disabled/enabled	(Optional) Indicates the enable flag of the static route. The flag is enabled by default.
Defaults	By default, no static route is configured.	
Command Mode	Global configuration mode	
Usage Guide	The simplest configuration of this command is **ip route** *networknet-maskip-address.*	

Configuring an IPv6 Static Route

Configure the following command on an IPv6-enabled router.

Command	**ipv6 route***ipv6-prefix/prefix-length* { *ipv6-address* | *interface* [*ipv6-address*] } [*distance*] [**weight***number*] [**description***description-text*]	
Parameter Description	*ipv6-prefix*	Indicates the IPv6 prefix, which must comply with the address expression specified in RFC4291.
	prefix-length	Indicates the length of the IPv6 prefix. Note that a slash (/) must be added in front of the length.
	ipv6-address	(Optional) Indicates the next-hop address of the static route. You must specify at least one of *ipv6-address* and *interface,* or both of them. If *ipv6-address* is not specified, a static direct route is configured.
	interface	(Optional) Indicates the next-hop exit interface of the static route. You must specify at least one of *ipv6-address* and *interface,* or both of them. If *interface* is not specified, a recursive static direct route is configured. The exit interface is obtained by the next hop in the routing table.
	distance	(Optional) Indicates the administrative distance of the static route. The administrative distance is 1 by default.
	weight *number*	(Optional) Indicates the weight of the static route, which must be specified when you configure equal-cost routes. The weight ranges from 1 to 8. When the weights of all equal-costroutes of a route are summed up, the sum cannot exceed the maximum number of equal-cost routes that can be configured for the route. Weighting of equal-cost routes of a route indicates the traffic ratio of these routes. The weight is 1 by default.
	description*description-text*	(Optional) Indicates the description of the static route. By default, no description is configured. *description-text* is a string of one to 60 characters.
Defaults	By default, no static route is configured.	
Command Mode	Global configuration mode	
Usage Guide	The simplest configuration of this command is **ipv6 route***ipv6-prefix / prefix-lengthipv6-address.*	

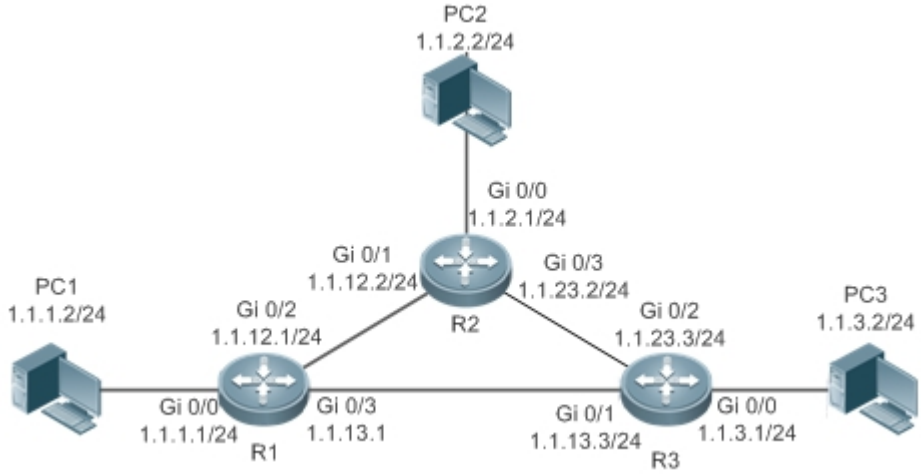
Verification

- Run the **show ip route** command to display the IPv4 routing table and check whether the configured IPv4 static route takes effect.

- Run the **show ipv6 route** command to display the IPv6 routing table and check whether the configured IPv6 static route takes effect.

Configuration Example

Configuring Static Routes to Implement Interworking of the IPv4 Network

Scenario **Figure 5- 4**	
Configuration Steps	- Configure interface addresses on each device.
R1	R1#configure terminal R1(config)#interface gigabitEthernet 0/0 R1(config-if-GigabitEthernet 0/0)# ip address 1.1.1.1 255.255.255.0 R1(config-if-GigabitEthernet 0/0)# exit R1(config)#interface gigabitEthernet 0/2 R1(config-if-GigabitEthernet 0/2)# ip address 1.1.12.1 255.255.255.0 R1(config-if-GigabitEthernet 0/0)# exit R1(config)#interface gigabitEthernet 0/3 R1(config-if-GigabitEthernet 0/3)# ip address 1.1.13.1 255.255.255.0
R2	R2#configure terminal R2(config)#interface gigabitEthernet 0/0 R2(config-if-GigabitEthernet 0/0)# ip address 1.1.2.1 255.255.255.0 R2(config-if-GigabitEthernet 0/0)# exit R2(config)#interface gigabitEthernet 0/1 R2(config-if-GigabitEthernet 0/1)# ip address 1.1.12.2 255.255.255.0 R2(config-if-GigabitEthernet 0/0)# exit R2(config)#interface gigabitEthernet 0/3 R2(config-if-GigabitEthernet 0/3)# ip address 1.1.23.2 255.255.255.0

R3	R3#configure terminal R3(config)#interface gigabitEthernet 0/0 R3(config-if-GigabitEthernet 0/0)# ip address 1.1.3.1 255.255.255.0 R3(config-if-GigabitEthernet 0/0)# exit R3(config)#interface gigabitEthernet 0/1 R3(config-if-GigabitEthernet 0/1)# ip address 1.1.13.3 255.255.255.0 R3(config-if-GigabitEthernet 0/0)# exit R3(config)#interface gigabitEthernet 0/2 R3(config-if-GigabitEthernet 0/2)# ip address 1.1.23.3 255.255.255.0
	● Configure static routes on each device.
R1	R1#configure terminal R1(config)#ip route 1.1.2.0 255.255.255.0 GigabitEthernet 0/2 1.1.12.2 R1(config)# ip route 1.1.3.0 255.255.255.0 GigabitEthernet 0/3 1.1.13.3
R2	R2#configure terminal R2(config)#ip route 1.1.1.0 255.255.255.0 GigabitEthernet 0/1 1.1.12.1 R2(config)# ip route 1.1.3.0 255.255.255.0 GigabitEthernet 0/3 1.1.23.3
R3	R3#configure terminal R3(config)#ip route 1.1.2.0 255.255.255.0 GigabitEthernet 0/2 1.1.23.2 R3(config)# ip route 1.1.1.0 255.255.255.0 GigabitEthernet 0/1 1.1.13.1
Verification	● Display the routing table.
R1	R1# show ip route Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area, * - candidate default Gateway of last resort is no set C 1.1.1.0/24 is directly connected, GigabitEthernet 0/0 C 1.1.1.1/32 is local host. S 1.1.2.0/24 [1/0] via 1.1.12.2, GigabitEthernet 0/2

	S 1.1.3.0/24 [1/0] via 1.1.13.3, GigabitEthernet 0/2 C 1.1.12.0/24 is directly connected, GigabitEthernet 0/2 C 1.1.12.1/32 is local host. C 1.1.13.0/24 is directly connected, GigabitEthernet 0/3 C 1.1.13.1/32 is local host.
R2	R2# show ip route Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area, * - candidate default Gateway of last resort is no set S 1.1.1.0/24 [1/0] via 1.1.12.1, GigabitEthernet 0/0 C 1.1.2.0/24 is directly connected, GigabitEthernet 0/0 C 1.1.2.1/32 is local host. S 1.1.3.0/24 [1/0] via 1.1.23.3, GigabitEthernet 0/3 C 1.1.12.0/24 is directly connected, GigabitEthernet 0/1 C 1.1.12.2/32 is local host. C 1.1.23.0/24 is directly connected, GigabitEthernet 0/3 C 1.1.23.2/32 is local host.
R3	R3# show ip route Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area, * - candidate default Gateway of last resort is no set S 1.1.1.0/24 [1/0] via 1.1.13.1, GigabitEthernet 0/2 S 1.1.2.0/24 [1/0] via 1.1.23.2, GigabitEthernet 0/2

	C 1.1.3.0/24 is directly connected, GigabitEthernet 0/0 C 1.1.3.1/32 is local host. C 1.1.13.0/24 is directly connected, GigabitEthernet 0/1 C 1.1.13.3/32 is local host. C 1.1.23.0/24 is directly connected, GigabitEthernet 0/2 C 1.1.23.3/32 is local host.

Configuring Static Routes to Implement Interworking of the IPv6 Network

Scenario **Figure 5- 5**	
Configuration Steps	● Configure interface addresses on each device.
R1	R1#configure terminal R1(config)#interface gigabitEthernet 0/0 R1(config-if-GigabitEthernet 0/0)# ipv6 address 1111:1111::1/64 R1(config-if-GigabitEthernet 0/0)# exit R1(config)#interface gigabitEthernet 0/1 R1(config-if-GigabitEthernet 0/1)# ipv6 address 1111:1212::1/64
R2	R2#configure terminal R2(config)#interface gigabitEthernet 0/0 R2(config-if-GigabitEthernet 0/0)#ipv6 address 1111:2323::1/64 R2(config-if-GigabitEthernet 0/0)# exit R2(config)#interface gigabitEthernet 0/1 R2(config-if-GigabitEthernet 0/1)# ipv6 address 1111:1212::2/64
	● Configure static routes on each device.
R1	R1#configure terminal R1(config)# ipv6 route 1111:2323::0/64 gigabitEthernet 0/1
R2	R2#configure terminal R2(config)#ipv6 route 1111:1111::0/64 gigabitEthernet 0/1
Verification	● Display the routing table.

R1	R1# show ipv6 route IPv6 routing table name - Default - 10 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area C 1111:1111::/64 via GigabitEthernet 0/0, directly connected L 1111:1111::1/128 via GigabitEthernet 0/0, local host C 1111:1212::/64 via GigabitEthernet 0/1, directly connected L 1111:1212::1/128 via GigabitEthernet 0/1, local host S 1111:2323::/64 [1/0] via GigabitEthernet 0/1, directly connected C FE80::/10 via ::1, Null0 C FE80::/64 via GigabitEthernet 0/0, directly connected L FE80::2D0:F8FF:FEFB:C092/128 via GigabitEthernet 0/0, local host C FE80::/64 via GigabitEthernet 0/1, directly connected L FE80::2D0:F8FF:FEFB:C092/128 via GigabitEthernet 0/1, local host
R2	R2# show ipv6 route IPv6 routing table name - Default - 10 entries Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area C 1111:2323::/64 via GigabitEthernet 0/0, directly connected L 1111:2323::1/128 via GigabitEthernet 0/0, local host C 1111:1212::/64 via GigabitEthernet 0/1, directly connected

	L 1111:1212::1/128 via GigabitEthernet 0/1, local host S 1111:1111::/64 [1/0] via GigabitEthernet 0/1, directly connected C FE80::/10 via ::1, Null0 C FE80::/64 via GigabitEthernet 0/0, directly connected L FE80::2D0:F8FF:FEFB:C092/128 via GigabitEthernet 0/0, local host C FE80::/64 via GigabitEthernet 0/1, directly connected L FE80::2D0:F8FF:FEFB:C092/128 via GigabitEthernet 0/1, local host

Common Errors

- The link on the interface is not up.
- No IP address is configured for the interface.

9.4.2 Configuring a Default Route

Configuration Effect

- Generate a default route in the routing table. The default route is used to forward packets that cannot be forwarded by other routes.

Notes

- On a L2 switch, run the **ip default gateway** or **ipv6 default gateway** command to configure the default gateway.
- On a L3 switch, run the **ip route** 0.0.0.0 0.0.0.0 *gateway*or **ipv6 route** ::/0 *ipv6-gateway*command to configure the default gateway.
- If the **no ip routing** or **no ipv6 unicast- routing** command is configured on a L3 switch, you can run the **ip default gateway** or **ipv6 default gateway** command to configure the default gateway.

Configuration Steps

Configuring the IPv4 Gateway on a L2 Switch

Command	**ip default-gateway***gateway*
Parameter Description	gateway: indicates the IPv4 gateway address.
Defaults	By default, no static default route is configured.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the IPv6 Gateway on a L2 Switch

Command	**ipv6 default-gateway***gateway*
Parameter Description	gateway: indicates the IPv4 gateway address.
Defaults	By default, no static default route is configured.
Command	Global configuration mode

Mode	
Usage Guide	N/A

Configuring the IPv4 Default Gateway on a L3 Switch

Command	**ip route 0.0.0.00.0.0.0**{*ip-address* | *interface* [*ip-address*]} [*distance*] [**tag** *tag*] [**permanent**] [**weight** *number*] [**description***description-text*] [**disabled** | **enabled**]	
Parameter Description	**0.0.0.0**	Indicates the address of the destination network.
	0.0.0.0	Indicates the mask of the destination network.
	ip-address	(Optional) Indicates the next-hop address of the static route. You must specify at least one of *ip-address* and *interface,* or both of them. If *ip-address* is not specified, a static direct route is configured.
	interface	(Optional) Indicates the next-hop exit interface of the static route. You must specify at least one of *ip-address* and *interface,* or both of them. If *interface* is not specified, a recursive static direct route is configured. The exit interface is obtained by the next hop in the routing table.
	distance	(Optional) Indicates the administrative distance of the static route. The administrative distance is 1 by default.
	tag	(Optional) Indicates the tag of the static route. The tag is 0 by default.
	permanent	(Optional) Indicates the flag of the permanent route. The static route is not a permanent route by default.
	weight *number*	(Optional) Indicates the weight of the static route. The weight is 1 by default.
	Description *description-text*	(Optional) Indicates the description of the static route. By default, no description is configured. *description-text* is a string of one to 60 characters.
	disabled /enabled	(Optional) Indicates the enable flag of the static route. The flag is enabled by default.
Defaults	By default, no static default route is configured.	
Command Mode	Global configuration mode	
Usage Guide	The simplest configuration of this command is **ip route0.0.0.0 0.0.0.0** *ip-address.*	

Configuring the IPv6 Default Gateway on a L3 Switch

Command	**ipv6 route::/0** { *ipv6-address* | *interface* [*ipv6-address*] } [*distance*] [**weight***number*] [**description***description-text*]	
Parameter Description	**::**	Indicates the IPv6 prefix, which must comply with the address expression specified in RFC4291.
	0	Indicates the length of the IPv6 prefix. Note that a slash (/) must be added in front of the length.
	Ipv6-address	(Optional) Indicates the next-hop address of the static route. You must specify at least one of *ipv6-address* and *interface,* or both of them. If *ipv6-address* is not specified, a static direct route is configured.
	interface	(Optional) Indicates the next-hop exit interface of the static route. You must specify at least one of *ipv6-address* and *interface,* or both of them. If *interface* is not specified, a recursive static direct route is configured. The exit interface is obtained by the next hop in the routing table.
	distance	(Optional) Indicates the administrative distance of the static route. The administrative distance is 1 by default.
	weight *number*	(Optional) Indicates the weight of the static route, which must be specified when you configure equal-cost routes. The weight ranges from 1 to 8. When the weights of all equal-cost routes of a route are summed up, the sum cannot exceed the maximum number of equal-cost routes that

		can be configured for the route. Weighting of equal-cost routes of a route indicates the traffic ratio of these routes. The weight is 1 by default.
	Description *description-text*	(Optional) Indicates the description of the static route. By default, no description is configured. *description-text* is a string of one to 60 characters.
Defaults	By default, no static default route is configured.	
Command Mode	Global configuration mode	
Usage Guide	The simplest configuration of this command is **ipv6 route** ::/0 *ipv6-gateway*.	

Configuring the IPv4 Default Network on a L3 Switch

Command	**ip default-network** *network*	
Parameter Description	*network*	Indicates the address of the network. (The network must be a Class A, B, or C network.)
Defaults	By default, no default network is configured.	
Command Mode	Global configuration mode	
Usage Guide	If the network specified by the **ip default-network** command exists, a default route is generated and the next hop to this network is the default gateway. If the network specified by the **ip default-network** command does not exist, the default route is not generated.	

Verification

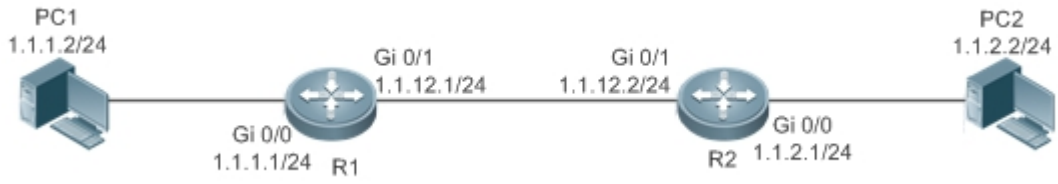
- On a L2 switch (or a L3 switch where routing is disabled), run the **show ip redirects** or **show ipv6 redirects** command to display the default gateway.
- On a L3 switch where routing is enabled, run the **show ip route** or **show ipv6 route** command to display the default route.

Configuration Example

Configuring IPv4 Default Routes on L3 Switches to Implement Network Interworking

Scenario

Figure 5- 6



Configuration Steps

- Configure IP addresses on L3 devices.

R1

```
R1#configure terminal
R1(config)#interface gigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)# ip address 1.1.1.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/1
```

	R1(config-if-GigabitEthernet 0/1)# ip address 1.1.12.1 255.255.255.0 R1(config-if-GigabitEthernet 0/0)# exit
R2	R2#configure terminal R2(config)#interface gigabitEthernet 0/0 R2(config-if-GigabitEthernet 0/0)# ip address 1.1.2.1 255.255.255.0 R2(config-if-GigabitEthernet 0/0)# exit R2(config)#interface gigabitEthernet 0/1 R2(config-if-GigabitEthernet 0/1)# ip address 1.1.12.2 255.255.255.0 R2(config-if-GigabitEthernet 0/0)# exit
R1	● Configure an IPv6 default gateway on R 1. R1#configure terminal R1(config)#ip route 0.0.0.0 0.0.0.0 GigabitEthernet 0/1 1.1.12.2
R2	R2#configure terminal R2(config)#ip route 0.0.0.0 0.0.0.0 GigabitEthernet 0/1 1.1.12.1
Verification	● Display the routing table.
R1	R1# show ip route Codes: C - Connected, L - Local, S - Static R - RIP, O - OSPF, B - BGP, I - IS-IS N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 IA - Inter area, * - candidate default Gateway of last resort is 1.1.12.2 S* 0.0.0.0/0 [1/0] via 1.1.12.2, GigabitEthernet 0/1 C 1.1.1.0/24 is directly connected, GigabitEthernet 0/0 C 1.1.1.1/32 is local host. C 1.1.12.0/24 is directly connected, GigabitEthernet 0/1 C 1.1.12.1/32 is local host.

9.4.3 Configuring Route Limitations

Configuration Effect

- Limit the number of equal-cost routes and number of static routes, or disable routing.

Notes

Route limitations cannot be configured on a L2 switch.

Configuration Steps

Configuring the Maximum Number of Equal-Cost Routes

Command	**maximum-paths** *number*	
Parameter Description	*number*	Indicates the maximum number of equal-cost routes. The value ranges from 1 to 64.
Defaults	The default value varies with the device model.	
Command Mode	Global configuration mode	
Usage Guide	Run this command to configure the maximum number of next hops in the equal-cost route. In load balancing mode, the number of routes on which traffic is balanced does not exceed the configured number of equal-cost routes.	

Configuring the Maximum Number of IPv4 Static Routes

Command	**ip static route-limit** *number*	
Parameter Description	*number*	Indicates the upper limit of routes. The value ranges from 1 to 10,000.
Defaults	By default, a maximum of 1,024 IP static routes can be configured.	
Command Mode	Global configuration mode	
Usage Guide	Run this command to configure the maximum number of IPv4 static routes. If the maximum number of IPv4 static routes is reached, no more IPv4 static route can be configured.	

Configuring the Maximum Number of IPv6 Static Routes

Command	**ipv6 static route-limit***number*	
Parameter Description	*number*	Indicates the upper limit of routes. The value ranges from 1 to 10,000.
Defaults	By default, a maximum of 1,000 IPv6 static routes can be configured.	
Command Mode	Global configuration mode	
Usage Guide	Run this command to configure the maximum number of IPv6 static routes. If the maximum number of IPv6 static routes is reached, no more IPv6 static route can be configured.	

Disabling IPv4 Routing

Command	**no ip routing**
Parameter Description	N/A

Defaults	By default, IPv4 routing is enabled.
Command Mode	Global configuration mode
Usage Guide	Run this command to disable IPv4 routing. If the device functions only as a bridge or a voice over IP (VoIP) gateway, the device does not need to use the IPv4 routing function of the FSOS software. In this case, you can disable the IPv4 routing function of the FSOS software.

Disabling IPv6 Routing

Command	**no ipv6 unicast-routing**
Parameter Description	N/A
Defaults	By default, IPv6 routing is enabled.
Command Mode	Global configuration mode
Usage Guide	Run this command to disable IPv6 routing. If the device functions only as a bridge or a VoIP gateway, the device does not need to use the IPv6 routing function of the FSOS software. In this case, you can disable the IPv6 routing function of the FSOS software.

9.5 Monitoring

Displaying

Description	Command
Displays the IPv4 routing table.	**show ip route**
Displays the IPv6 routing table.	**show ipv6route**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs IPv4 route management.	**debug nsm kernel ucast- v4**
Debugs IPv6 route management.	**debug nsm kernel ucast-v6**
Debugs default network management.	**debug nsm kernel default-network**
Debugs internal events of route management.	**debug nsm events**
Debugs sending of route management and routing protocol messages.	**debug nsm packet send**
Debugs receiving of route management and routing protocol messages.	**debug nsm packet recv**

10 Configuring Keys

10.1 Overview

Keys are a kind of parameters that are used in algorithms for conversion from plain text to cipher text or from cipher text to plain text.

Plain text and cipher text authentication are supported for packet authentication in a routing protocol, during which keys need to be used.

At present, keys are used only for RIP packet authentication.

10.2 Applications

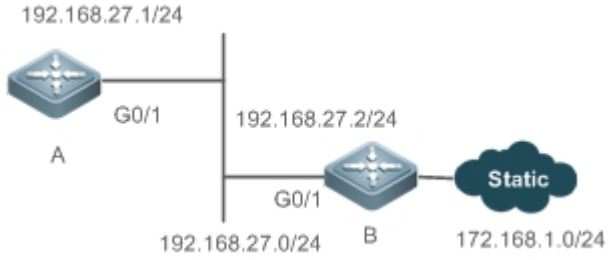
Application	Description
RIP Authentication	RIP uses keys for packet authentication.

10.2.1 RIP Authentication

Scenario

Network devices run RIP and use the MD5 authentication mode to increase the protocol security.

Figure 6- 1



Deployment

- Configure a key chain on A. Configure RIP to enable packet authentication and use the key chain.
- Configure a key chain on B. Configure RIP to enable packet authentication and use the key chain.

10.3 Features

Overview

Feature	Description
Key Chain	Provide a tool for authentication in a routing protocol.

10.3.1 Key Chain

Working Principle

A key chain may contain multiple different keys. Each key contains the following attributes:

- Key ID: Identifies a key. In the current key chain, keys and IDs are mapped in the one-to-one manner.

- Authentication string: Indicates a set of key characters used for verifying the consistency of authentication strings in a routing protocol.
- Lifetime: Specifies the lifetime of the current key for sending or receiving packets. Different authentication keys can be used in different periods.

Related Configuration

Creating a Key Chain and a Key

In the global configuration mode, run the **key chain** *key-chain-name* command to define a key chain and enter the key chain configuration mode.

In the key chain configuration mode, run the **key** *key-id* command to define a key and enter the key chain key configuration mode.

Configuring an Authentication String

In the key chain key configuration mode, run the **key-string** [**0**|**7**] *text* command to specify an authentication string.

- A plain text authentication string is configured by default. The value **0** indicates that a plain text authentication key is configured.
- The value **7** indicates that a cipher text authentication string is configured.
- The encryption authentication service is disabled by default. You can run the **service password-encryption** command to enable the encryption service to forcibly convert plain text authentication into cipher text.

Configuring Lifetime

In the key chain key configuration mode, you can configure the lifetime of a key chain in the receiving and sending directions.

- **accept-lifetime** *start-time* { **infinite** | *end-time* | **duration** *seconds* }: Configures the lifetime of a key chain in the receiving direction.
- **send-lifetime** *start-time* { **infinite** | *end-time* | **duration** *seconds* }: Configures the lifetime of a key chain in the sending direction.

10.4 Configuration

Configuration	Description and Command	
Configuring a Key Chain	(Mandatory) It is used to create a key.	
	key chain	Creates a key chain.
	key	Configures a key ID.
	key-string	Configures a key string.
	accept-lifetime	Configures the lifetime in the receiving direction.
	send-lifetime	Configures the lifetime in the sending direction.

10.4.1 Configuring a Key Chain

Configuration Effect

- Define a key chain to be used by a routing protocol.

Notes

- A key chain can take effect only after it is associated with a routing protocol.

Configuration Steps

Creating a Key Chain

- This configuration is mandatory if a key chain needs to be used.
- If there is no special requirement, you should perform this configuration on all routers for which routing protocol authentication needs to be performed.

Configuring a Key ID

- This configuration is mandatory if a key chain needs to be used.
- If there is no special requirement, you should perform this configuration on all routers for which routing protocol authentication needs to be performed.

Configuring a Key String

- This configuration is mandatory if a key chain needs to be used.
- If there is no special requirement, you should perform this configuration on all routers for which routing protocol authentication needs to be performed.

Configure the Lifetime in the Receiving Direction

- Optional.
- If the lifetime in the sending direction is not configured, the key chain will be always effective.

Configure the Lifetime in the Sending Direction

- Optional.
- If the lifetime in the sending direction is not configured, the key chain will be always effective.

Verification

- Use keys in a routing protocol and observe the neighborship established by the routing protocol. If the keys are inconsistent, the neighborship fails to be established.

Related Commands

Configuring a Key Chain

Command	**key chain** *key-chain-name*
Parameter Description	*key-chain-name*: Indicates the name of a key chain.
Command Mode	Global configuration mode
Usage Guide	To make a key chain take effect, you must configure at least one key.

Configuring a Key ID

Command	**key** *key-id*
Parameter Description	*key-id*: Indicates the authentication key ID in a key chain, ranging from 0 to 2,147,483,647.
Command	Key chain configuration mode.

Mode	
Usage Guide	-

Configuring a Key Authentication String

Command	**key-string** [**0**|**7**] *text*
Parameter Description	**0:** Specifies that the key is displayed in plain text. **7:** Specifies that the key is displayed in cipher text. ***text*:** Specifies the authentication string characters.
Command Mode	Key chain key configuration mode.
Usage Guide	-

Configuring the Lifetime in the Sending Direction

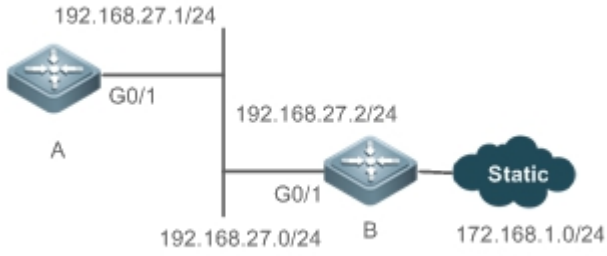
Command	**send-lifetime** *start-time* {**infinite** | *end-time* | **duration** *seconds*}
Parameter Description	***start-time*:** Indicates the start time of the lifetime. **infinite:** Indicates that the key is always effective. ***end-time*:** Indicates the end time of the lifetime, which must be later than start-time. **duration *seconds*:** Specifies the duration from the start time to the end time, ranging from 1 to 2,147,483,646.
Command Mode	Key chain key configuration mode.
Usage Guide	Run this command to define the lifetime of the key in the sending direction.

Configuring the Lifetime in the Receiving Direction

Command	**accept-lifetime** *start-time* **{infinite |** *end-time* **| duration** *seconds***}**
Parameter Description	***start-time*:** Indicates the start time of the lifetime. **infinite:** Indicates that the key is always effective. ***end-time*:** Indicates the end time of the lifetime, which must be later than start-time. **duration *seconds*:** Specifies the duration from the start time to the end time, ranging from 1 to 2,147,483,646.
Command Mode	Key chain key configuration mode.
Usage Guide	Run this command to define the lifetime of the key in the receiving direction.

Configuration Example

Configuring a Key Chain and Using the Key Chain in RIP Packet Authentication

Scenario **Figure 6- 2**	

Configuration Steps	● Configure a key on all routers. ● Configure RIP on all routers. ● Enable RIP authentication on all routers.
A	A>enable A#configure terminal A(config)#key chain ripchain A(config-keychain)#key 1 A(config-keychain-key)#key-string Hello A(config-keychain-key)#accept-lifetime 16:30:00 Oct 1 2013 duration 43200 A(config-keychain-key)#send-lifetime 16:30:00 Oct 1 2013 duration 43200 A(config-keychain-key)#exit A(config-keychain)#key 2 A(config-keychain-key)#key-string World A(config-keychain-key)#accept-lifetime 04:00:00 Oct 2 2013 infinite A(config-keychain-key)#send-lifetime 04:00:00 Oct 2 2013 infinite A(config-keychain-key)#exit A(config)#interface gigabitEthernet 0/1 A(config-if)#ip address 192.168.27.1 255.255.255.0 A(config-if)#ip rip authentication key-chain ripchain A(config-if)#ip rip authentication mode md5 A(config-if)#exit A(config)#router rip A(config-router)#version 2 A(config-router)#network 192.168.27.0
B	B>enable B#configure terminal B(config)#key chain ripchain B(config-keychain)#key 1 B(config-keychain-key)#key-string Hello B(config-keychain-key)#accept-lifetime 16:30:00 Oct 1 2013 duration 43200 B(config-keychain-key)#send-lifetime 16:30:00 Oct 1 2013 duration 43200 B(config-keychain-key)#exit B(config-keychain)#key 2 B(config-keychain-key)#key-string World

	B(config-keychain-key)#accept-lifetime 04:00:00 Oct 2 2013 infinite B(config-keychain-key)#send-lifetime 04:00:00 Oct 2 2013 infinite B(config-keychain-key)#exit B(config)#interface gigabitEthernet 0/1 B(config-if)#ip address 192.168.27.2 255.255.255.0 B(config-if)#ip rip authentication key-chain ripchain B(config-if)#ip rip authentication mode md5 B(config-if)#exit B(config)#router rip B(config-router)#version 2 B(config-router)#network 192.168.27.0 B(config-router)#redistribute static
Verification	Run the **show ip route rip** command to check whether router A can receive an RIP route from router B.
A	A(config)#show ip route rip R 172.168.0.0/16 [120/1] via 192.168.27.2, 00:05:16, GigabitEthernet 0/1

Common Errors

- A key is not correctly associated with a routing protocol, which causes that authentication does not take effect.
- The keys configured on multiple routers are not consistent, which causes authentication failure.

10.5 Monitoring

Displaying

Description	Command
Displays the configurations of a key chain.	**show key chain** [*key-chain-name*]

11 Configuring Routing Policies

11.1 Overview

Routing policies are a policy set for changing the packet forwarding path or routing information and are often implemented by a filtering list and a route map. Routing policies are flexibly and widely applied in the following methods:

- Use a filtering list in a routing protocol to filter or modify routing information.
- Use a route map in a routing protocol to filter or modify routing information. Where, the route map can further use a filtering list.
- Use a route map in policy-based routing (PBR) to control packet forwarding or modify packet fields.

11.2 Applications

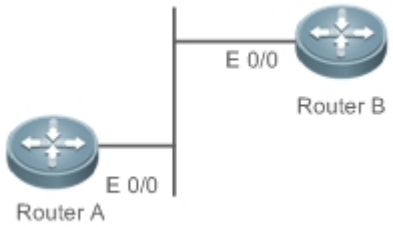
Application	Description
Route Filtering	Use a filtering list in a routing protocol to filter the routing information sent or received by the protocol.
Route Re-distribution	Use a route map in a routing protocol to filter or modify routing information and re-distribute RIP routes to OSPF. Only RIP routes with 4 hops can be re-distributed.

11.2.1 Route Filtering

By default, a routing protocol advertises and learns all routing information. When a filtering list is used, the routing protocol advertises only required routes or receives only required routing information.

Scenario

Figure 7- 1



As shown in Figure 7- 1, router A has routes to 3 networks: 10.0.0.0, 20.0.0.0 and 30.0.0.0.
Configure a filtering list on the routers to achieve the following purposes:

- Filter the sent routing information on router A to filter routes that router A does not need to send.
- Filter the received routing information on router B to filter routes that router B does not need to learn.

Deployment

- Filter the sent routing information 30.0.0.0 on router A.
- Filter the received routing information 20.0.0.0 on router B to ensure that router B learns only routing information 10.0.0.0.

11.2.2 Route Re-distribution

By default, route re-distribution will re-distribute all routing information in a routing protocol to another routing protocol. All routing attributes will also be inherited. You can use a route map to perform conditional control for re-distribution between two routing protocols, including:

- Specify the range for re-distributing routes and re-distribute only routing information that meets certain rules.
- Set the attributes of routes generated by re-distribution.

Scenario

Figure 7- 2

As shown in Figure 7- 2, configure route re-distribution on the devices to achieve the following purposes:

- Re-distribute only RIP routes with 4 hops to OSPF.
- In the OSPF routing domain, the initial metric of this route is 40, the route type is the external route type-1 and the route tag value is set to 40.

Deployment

- Configure a route with 4 hops in the route map rip_to_ospf: match, and set the initial metric of this route to 40, the route type to the external route type-1 and the route tag value to 40.
- Configure route re-distribution to re-distribute RIP routes to OSPF and use the route map rip_to_ospf.

11.3 Features

Overview

Feature	Description
Filtering List	Define a group of lists based on a route attribute, which can be used by a routing protocol for route filtering.
Route Map	A policy defines "if certain conditions are matched, you can perform certain processing actions".

11.3.1 Filtering List

Filtering lists are a group of lists defined based on a routing attribute and are a tool for filtering routing policies. Independent filtering lists are meaningless and can be used to filter routes only when they are applied in a routing protocol.

Working Principle

Based on different routing attributes, filtering lists are classified into the following types:

- **Access Control List (ACL)**

ACLs comprise IPv4 and IPv6 ACLs. When defining ACLs, you can specify IPv4/IPv6 addresses and masks to match the destination network segment or next-hop addresses of routing information.

For description about ACLs, see the *ACL Configuration Guide*.

Address Prefix List (prefix-list)

Similar to ACLs, prefix-lists, including IPv4 prefix-lists and IPv6 prefix-lists, are used to match destination network segments of routing information during route filtering.

Related Configuration

Creating an ACL

By default, no ACL is configured and no policy is set.

In the global configuration mode, run the **ip access-list** { **extended** | **standard** } { *id* | *name* } command to create an IPv4 ACL.

You can set multiple policies in an ACL, sorted by their sequence numbers. Policies have two working modes: permit and deny.

Creating a Prefix-List

By default, no prefix-list is configured and no entry is set.

In the global configuration mode, run the **ip prefix-list** *prefix-list-name* [**seq** *seq-number*] { **deny** | **permit** } *ip-prefix* [**ge** *minimum-prefix-length*] [**le** *maximum-prefix-length*] command to create an IPv4 prefix-list and add a prefix entry to the list.

You can set multiple entries in the prefix-list, sorted by their sequence numbers. Entries have two working modes: permit and deny.

Run the **ip prefix-list** *prefix-list-name* **description** *descripton-text* command to add description to the prefix-list.

Run the **ip prefix-list sequence-number** command to enable the sorting function for the prefix-list.

Creating an Extcommunity-List

By default, no excommunity-list is configured and no entry is set.

In the global configuration mode, run the **ip extcommunity-list** {*standard-list* **| standard** *list-name* } { **permit | deny** } [**rt** *value*] [**soo** *value*] command to create a standard extcommunity list and add an entry to the list.

Run the **ip extcommunity-list** {*expanded-list* | **expanded** *list-name* } { **permit | deny** } [*regular-expression*] command to create an extcommunity list and add an entry to the list.

You can also run the **ip extcommunity-list** {*expanded-list* | **expanded** *list-name*| *standard-list* | **standard** *list-name* } command to create an extcommunity list and enter the configuration mode of **ip extcommunity-list** to add entries.

You can set multiple entries in the extcommunity-list. Entries have two working modes: permit and deny.

11.3.2 Route Map

A policy is a "match …, set…" statement, which indicates that "if certain conditions are matched, you can perform some processing actions".

Working Principle

Executing policies

A route map may contain multiple policies. Each policy has a corresponding sequence number. A smaller sequence number means a higher priority. Policies are executed based on their sequence numbers. Once the matching condition of a policy is met, the processing action for this policy needs to be performed and the route map exits. If no matching condition of any policy is met, no processing action will be performed.

Working Modes Of Policies

Policies have two working modes:

- permit: When the matching condition of a policy is met, the processing action for this policy will be performed and the route map will exit.
- deny: When the matching condition of a policy is met, the processing action for this policy will not be performed and the route map will exit.

Matching Conditions Of Policies

The matching condition of a policy may contain 0, 1 or more match rules.

- If the matching condition contains 0 match rule, no packet will be matched.
- If the matching condition contains one or more match rules, all rules must be matched.

Processing Action for a Policy

The processing action of a policy may contain 0, 1 or more set rules.

- If the processing action contains 0 set rule, no processing action will be performed and the route map will directly exit.
- If the processing action contains one or more set rules, all processing actions will be performed and then the route map will exit.

If set rules have different priorities, the set rule with the highest priority will take effect.

Related Configuration

Creating a Route Map (Policy)

By default, no route map is configured and no policy is set.

In the global configuration mode, you can run the **route-map** *route-map-name* [**permit** | **deny**] [*sequence-number*] command to create a route map and add a policy to the route map.

You can set multiple policies in a route map. Each policy uses different sequence numbers.

Setting Matching Conditions of a Policy

By default, no match rule is set (that is, the matching condition of a policy contains 0 match rule).

In the route map mode, run the **match** command to set match rules. One **match** command is mapped to one match rule.

FSOS provides abundant **match** commands for setting flexible matching conditions.

Command	Description

Command	Description
match interface	Uses the output interface of a route as the matching condition.
match ip address	Uses the destination IPv4 address of a route as the matching condition.
match ip next-hop	Uses the next-hop IPv4 address of a route as the matching condition.
match ip route-source	Uses the source IPv4 address of a route as the matching condition.
match ipv6 address	Uses the destination IPv6 address of a route as the matching condition.
match ipv6 next-hop	Uses the next-hop IPv6 address of a route as the matching condition.
match ipv6 route-source	Uses the source IPv6 address of a route as the matching condition.
match metric	Uses the metric of a route as the matching condition.
match tag	Uses the tag value of a route as the matching condition.

Setting the Processing Actions of a Policy

By default, no set rule is configured (that is, the processing action of a policy contains 0 set rule).

In the route map mode, run the **set** command to configure set rules. One **set** command is mapped to one set rule.

FSOS provides abundant **set** commands for setting flexible processing actions.

Command	Description
set ip default nexthop	Specifies the default next hop of a route. This command has a lower priority than a common route and a higher priority than **set default interface**.
set ip dscp	Modifies the **dscp** field of an IP packet.
set ip nexthop	Specifies the next hop of a route. This command has a higher priority than **set interface**.
set ip precedence	Modifies the **precedence** field of an IP packet.
set ip tos	Modifies the **tos** field of an IP packet.
set level	Sets the destination area type to which a route will be directed.
set metric	Modifies the metric value of a route.
set metric-type	Sets the metric type of a route.
set next-hop	Sets the next-hop IP address of a route.
set tag	Sets the tag value of a route.

11.4 Configuration

Configuration	Description and Command	
Configuring a Route Map	(Optional) It is used to define a policy.	
	route-map	Creates a policy (route map).
	match	Sets the matching conditions of the policy.
	set	Sets the processing actions of the policy.
Configuring a Filtering List	(Optional) It is used to define a filtering list.	
	ip as-path	Defines AS path filtering rules.
	ip community-list	Defines a community list.

Configuration	Description and Command	
	ip prefix-list	Creates a prefix-list.
	ip prefix-list description	Adds description to a prefix-list.
	ip prefix-list sequence-number	Enables the sorting function for a prefix-list.
	Ipv6 prefix-list	Creates an IPv6 prefix-list.
	ipv6 prefix-list description	Adds description to an IPv6 prefix-list.
	ipv6 prefix-list sequence-number	Enables the sorting function for an IPv6 prefix-list.

11.4.1 Configuring a Route Map

Configuration Effect

- Define a set of routing policies to be used by routing protocols or PBR.

Notes

- If a **match** command uses an ACL to define packet matching conditions, the ACL must be configured.
- The following **match** commands cannot be configured at the same time:

The Following match Commands	Cannot Be Configured with the Following match Commands At the Same Time
match ip address	match ip prefix-list
match ipv6 address	match ipv6 prefix-list
match ip next-hop	match ip next-hop prefix-list
match ipv6 next-hop	match ipv6 next-hop prefix-list
match ip route-source	match ip route-source prefix-list
match ipv6 route-source	match ipv6 route-source prefix-list

- The following **set** commands cannot be configured at the same time:

The Following set Commands	Cannot Be Configured with the Following set Commands At the Same Time
set ip dscp	set ip tos
set ip dscp	set ip precedence

Configuration Steps

Creating a Policy (Route Map)

- Mandatory.
- Perform this configuration on a device to which a policy needs to be applied.

Setting Matching Conditions of a Policy

- Optional.
- If no match rule is configured, no packet will be matched.
- If multiple match rules are configured, all the match rules must be matched.

- Perform this configuration on a device to which a policy needs to be applied.

Setting the Processing Actions of a Policy

- Optional.
- If no set rule is configured, no processing action will be performed.
- If multiple set rules are configured, all set rules must be executed (if the set rules have different priorities, the set rule with the highest priority takes effect).
- Perform this configuration on a device to which a policy needs to be applied.

Verification

- Check the configurations of the route map.

Related Commands

Creating a Policy (Route Map)

Command	**route-map** *route-map-name* [{ **permit** | **deny** } *sequence*]
Parameter Description	*route-map-name*: Indicates the name of a route map, comprising not more than 32 characters. permit: Specifies the working mode of this policy as permit, which is the default mode. deny: Specifies the working mode of this policy as deny. The default mode is permit. *sequence*: Specifies the sequence number of this policy. A smaller value means a higher priority. The default value is 10.
Command Mode	Global configuration mode
Usage Guide	If this route map is unavailable, this command will create a route map and add a policy to the route map. If this route map is available, this command will add a policy to the route map.
Command	**match interface** *interface-type interface-number* […*interface-type interface-number*]
Parameter Description	*interface-type interface-number*: Indicates the interface type and interface number.
Command Mode	Route map configuration mode
Usage Guide	This match rule is used to match the next-hop output interface of a route or a packet.

Command	**match ip address** { *access-list-number* [*access-list-number...* | *access-list-name...*] | *access-list-name* [*access-list-number...*| *access-list-name*] | **prefix-list** *prefix-list-name* [*prefix-list-name...*] }
Parameter Description	*access-list-number*: Indicates the access list number. For a standard access list, the value ranges are 1 to 99 and 1300 to 1999. For an extended access list, the value ranges are 100 to 199 and 2000 to 2699. *access-list-name*: Indicates the access list name. **prefix-list** *prefix-list-name*: Indicates the name of a prefix-list to be matched.
Command Mode	Route map configuration mode
Usage Guide	This match rule matches the destination IPv4 address of a packet or route by using an ACL or a prefix-list.

	An ACL and a prefix-list cannot be configured at the same time.

Command	**match ip next-hop** { *access-list-number* [*access-list-number...* | *access-list-name...*] | *access-list-name* [*access-list-number...* | *access-list-name*] | **prefix-list** *prefix-list-name* [*prefix-list-name...*] }
Parameter Description	*access-list-number*: Indicates the access list number. For a standard access list, the value ranges are 1 to 99 and 1300 to 1999. For an extended access list, the value ranges are 100 to 199 and 2000 to 2699. *access-list-name*: Indicates the access list name. **prefix-list** *prefix-list-name*: Indicates the name of a prefix-list to be matched.
Command Mode	Route map configuration mode
Usage Guide	This match rule matches the next-hop IPv4 address of a route by using an ACL or a prefix-list. An ACL and a prefix-list cannot be configured at the same time.

Command	**match ip route-source** { *access-list-number* [*access-list-number...* | *access-list-name...*] | *access-list-name* [*access-list-number...* | *access-list-name*] | **prefix-list** *prefix-list-name* [*prefix-list-name...*] }
Parameter Description	*access-list-number*: Indicates the access list number. For a standard access list, the value ranges are 1 to 99 and 1300 to 1999. For an extended access list, the value ranges are 100 to 199 and 2000 to 2699. *access-list-name*: Indicates the access list name. **prefix-list** *prefix-list-name*: Indicates the name of a prefix-list to be matched.
Command Mode	Route map configuration mode
Usage Guide	This match rule matches the source IPv4 address of a route by using an ACL or a prefix-list. An ACL and a prefix-list cannot be configured at the same time.

Command	**match ipv6 address** { *access-list-name* | **prefix-list** *prefix-list-name* }
Parameter Description	*access-list-name*: Indicates the access list name. **prefix-list** *prefix-list-name*: Indicates the name of an IPv6 prefix-list to be matched.
Command Mode	Route map configuration mode
Usage Guide	This match rule matches the destination IPv6 address of a packet or route by using an ACL or a prefix-list. An ACL and a prefix list cannot be configured at the same time.

Command	**match ipv6 next-hop** { *access-list-name* | **prefix-list** *prefix-list-name* }
Parameter Description	*access-list-name*: Indicates the access list name. **prefix-list** *prefix-list-name*: Indicates the name of an IPv6 prefix-list to be matched.
Command Mode	Route map configuration mode
Usage Guide	This match rule matches the next-hop IPv6 address of a route by using an ACL or a prefix-list. An ACL and a prefix-list cannot be configured at the same time.

Command	**match ipv6 route-source** { *access-list-name* | **prefix-list** *prefix-list-name* }
Parameter Description	*access-list-name*: Indicates the access list name. **prefix-list** *prefix-list-name*: Indicates the name of an IPv6 prefix-list to be matched.
Command Mode	Route map configuration mode
Usage Guide	This match rule matches the source IPv6 address of a route by using an ACL or a prefix-list. An ACL and a prefix-list cannot be configured at the same time.

Command	**match metric** *metric*
Parameter Description	*metric*: Indicates the metric value of a route, ranging from 0 to 4,294,967,295.
Command Mode	Route map configuration mode
Usage Guide	This match rule is used to match the metric value of a route.

Command	**match tag** *tag* […*tag*]
Parameter Description	*tag*: Indicates the tag value of a route.
Command Mode	Route map configuration mode
Usage Guide	This match rule is used to match the tag value of a route.

Setting the Processing Actions of a Policy

Command	**set ip default next-hop** *ip-address* [*weight*] […*ip-address* [*weight*]]
Parameter Description	*ip-address*: Indicates the next-hop IP address. *weight*: Indicates the weight of this next hop.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to specify the default next hop of a route.

Command	**set ip dscp** *dscp_value*
Parameter Description	*dscp_value*: Sets the DSCP value in the IP header of an IP packet.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to modify the dscp field of an IP packet.

Command	**set ip next-hop** *ip-address* [*weight*] […*ip-address* [*weight*]]
Parameter Description	*ip-address*: Indicates the next-hop IP address. *weight*: Indicates the weight of this next hop.

Command Mode	Route map configuration mode
Usage Guide	This set rule is used to specify the next hop of a route.
Command	**set ip precedence** { *number* | **critical** | **flash** | **flash-override** | **immediate** | **internet** | **network** | **priority** | **routine** }
Parameter Description	*number*: Indicates the priority of the IP header with a number, ranging from 0 to 7. 7: critical 6: flash 5: flash-override 4: immediate 3: internet 2: network 1: priority 0: routine **critical** | **flash** | **flash-override** | **immediate** | **internet** | **network** | **priority** | **routine**: priority of an IP header.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to modify the **precedence** field of an IP packet header.

Command	**set ip tos** { *number* | **max-reliability | max-throughput | min-delay | min-monetary-cost | normal** }
Parameter Description	*number*: Indicates the TOS value of an IP header with a number, ranging from 0 to 15. 2: **max-reliability** 4: **max-throughput** 8: **min-delay** 1: **min-monetary-cost** 0: **normal** **max-reliability | max-throughput | min-delay | min-monetary-cost | normal**: priority of an IP header.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to modify the tos field of an IP packet.

Command	**set ipv6 default next-hop** *global-ipv6-address* [*weight*] [*global-ipv6-address* [*weight*] ...]
Parameter Description	*global-ipv6-address*: Indicates the next-hop IPv6 address for packet forwarding. The next-hop router must be a neighbor router. *weight*: Indicates the weight in the load balancing mode, ranging from 1 to 8. A larger value means larger packet traffic to be shared by the next hop.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to specify the default next hop IPv6 address of a route.

Command	**set ipv6 next-hop** *global-ipv6-address* [*weight*] [*global-ipv6-address* [*weight*] ...]
Parameter	

Description	*global-ipv6-address*: Indicates the next-hop IPv6 address for packet forwarding. The next-hop router must be a neighbor router. *weight*: Indicates the weight in the load balancing mode, ranging from 1 to 8. A larger value means larger packet traffic to be shared by the next hop.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to specify the next hop IPv6 address of a route.

Command	**set ipv6 precedence** { *number* | **critical** | **flash** | **flash-override** | **immediate** | **internet** | **network** | **priority** | **routine** }
Parameter Description	*number*: Indicates the priority of the IP header with a number, ranging from 0 to 7. 7: critical 6: flash 5: flash-override 4: immediate 3: internet 2: network 1: priority 0: routine **critical** | **flash** | **flash-override** | **immediate** | **internet** | **network** | **priority** | **routine**: priority of an IP header.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to set the priority of an IPv6 packet header.

Command	**set level** { **stub-area** | **backbone** }
Parameter Description	**stub-area**: Indicates that the re-distribution route is advertised to OSPF Stub Area. **backbone**: Indicates that the re-distribution route is advertised to the OSPF backbone area.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to set the destination area type to which a route will be redirected.

Command	**set metric** [+ *metric-value* | **-** *metric-value* | *metric-value*]
Parameter Description	+: Increases (based on the metric value of the original route). -: Decreases (based on the metric value of the original route). *metric-value*: Sets the metric value of a re-distribution route. A larger value means a lower priority.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to modify the metric value of a route.

Command	**set metric-type** *type*
Parameter	*type*: Sets the type of a re-distribution route. The default type of an OSPF re-distribution route is type-2.

Description	
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to set the metric type.

Command	**set next-hop** *ip-address*
Parameter Description	*ip-address*: Indicates the next-hop IP address.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to set the next-hop IP address.

Command	**set tag** *tag*
Parameter Description	*tag*: Sets the tag of a re-distribution route.
Command Mode	Route map configuration mode
Usage Guide	This set rule is used to set the tag value of a route.

Displaying the Configurations of a Route Map

Command	**show route-map** [*name*]
Parameter Description	***name***: Specifies a route map.
Command Mode	Privilege, global and interface configuration modes
Usage Guide	Run the **show route-map** command to display the configurations of a route map. If an ACL is used when a route map is configured, you can run the **show access-list** command to display the configurations of the ACL.

Configuration Example

Using a Route Map in Route Re-distribution to Filter and Modify Routing Information

Scenario **Figure 7- 3**	As shown in Figure 7- 3, a device is connected to both an OSPF routing domain and RIP routing domain.
	● Re-distribute only RIP routes with 4 hops to OSPF. In the OSPF route domain, if the route type is the external route type-1, set the tag value of the route to 40. ● Re-distribute only OSPF routes with the tag value 10 to RIP. In the RIP route domain, set the initial metric value of this route to 10.

Configuration Steps	● Configure the route map redrip: Match a route with 4 hours, set the initial metric value of the route to 40, set the route type to the external route type-1, and set the tag value of the route to 40. ● Configure the route map redospf: match a route with the tag value 10 and set the initial metric value of the route to 10. ● Configure re-distribution of the RIP route to OSPF and apply the route map redrip. ● Configure re-distribution of the OSPF route to RIP and apply the route map redospf.
	FS(config)# route-map redrip permit 10 FS(config-route-map)# match metric 4 FS(config-route-map)# set metric-type type-1 FS(config-route-map)# set tag 40 FS(config-route-map)# exit FS(config)# route-map redospf permit 10 FS(config-route-map)# match tag 10 FS(config-route-map)# set metric 10 FS(config-route-map)# exit FS(config)# router ospf 1 FS(config-router)# redistribute rip subnets route-map redrip FS(config-router)# exit FS(config)# router rip FS(config-router)# redistribute ospf 1 route-map redospf FS(config-router)# exit
Verification	● Check the configurations of the route map to verify the policy rules. ● Check the OSPF routing information library to verify that the rules matching the policy rules are re-distributed.
	FS# show route-map route-map redrip, permit, sequence 10 Match clauses: metric 4 Set clauses: metric 40 metric-type type-1 tag 40 route-map redospf, permit, sequence 10

	Match clauses: tag 10 Set clauses: metric 10
	FS# show ip ospf database external OSPF Router with ID (192.100.1.9) (Process ID 1) AS External Link States LS age: 5 Options: 0x2 (-|-|-|-|-|-|E|-) LS Type: AS-external-LSA Link State ID: 192.168.199.0 (External Network Number) Advertising Router: 192.100.1.9 LS Seq Number: 80000001 Checksum: 0x554d Length: 36 Network Mask: /24 Metric Type: 1 TOS: 0 Metric: 4 Forward Address: 0.0.0.0 External Route Tag: 40

11.4.2 Configuring a Filtering List

Configuration Effect

- Define a set of route filtering rules to be used by routing protocols.

Notes

- A configured filtering list can take effect only after it is associated with a routing protocol.

Configuration Steps

Configuring a Prefix-List

- To filter address prefixes, you should perform this configuration.
- If there is no special requirement, you should perform this configuration on a route for which filtering based on a prefix-list needs to be performed.

Configuring a Community List

- To filter community attributes, you should perform this configuration.
- If there is no special requirement, you should perform this configuration on a route for which community attributes need to be filtered.

Verification

- Check whether the filtering list is correctly configured.
- Check the routing table to verify that routes can be correctly filtered.

Related Commands

Defining a Community List

Command	**ip community-list** { { **standard** | **expanded** } *community-list-name* | *community-list-number* } { **permit** | **deny** } [*community-number..*]
Parameter Description	**standard**: Indicates a standard community list. **expanded**: Indicates an extended community list. *community-list-name*: Indicates the community list name, comprising not more than 80 characters. *community-list-number*: Indicates the community list number. For a standard community list, the value ranges from 1 to 99. For an extended community list, the value ranges from 100 to 199. **permit:** Permits access. **deny:** Denies access. *community-number*: Indicates the community attribute value.
Command Mode	Global configuration mode
Usage Guide	

Defining an Extcommunity List

Command	**ip extcommunity-list** {*expanded-list* | **expanded** *list-name* } { **permit | deny** } [*regular-expression*]
Parameter Description	*expand-list*: Indicates an extended extcommunity list, ranging from 100 to 199. One extcommunity list may contain multiple rules. *standard-list*: Indicates a standard extcommunity list, ranging from 1 to 99. One extcommunity list may contain multiple rules. *expanded list-name*: Indicates the name of an extended extcommunity, comprising not more than 32 characters. When using this parameter, you enter the extcommunity list configuration mode. *standard list-name*: Indicates the name of a standard extcommunity list, comprising not more than 32 characters. When

	using this parameter, you enter the extcommunity list configuration mode. **permit:** Defines an extcommunity rule for permitting. **deny:** Defines an extcommunity rule for denying. *regular-expression*: (optional) Defines a matching template that is used to match an extcommunity. *sequence-number*: (Optional) Defines the sequence number of a rule, ranging from 1 to 2,147,483,647. If no sequence number is specified, the sequence number automatically increases by 10 when a rule is added by default. The initial number is 10. **rt:** (Optional) Sets the RT attribute value. This command can be used only for the standard extcommunity configuration, but not for the extended extcommunity configuration. **soo:** (Optional) Sets the SOO attribute value. This command can be used only for the standard extcommunity configuration, but not for the extended extcommunity configuration. *value*: Indicates the value of an extended community (extend_community_value).
Command Mode	Global configuration mode and ip extcommunity-list configuration mode
Usage Guide	-

Creating a Prefix-List

Command	**ip prefix-list** *prefix-list-name* [**seq** *seq-number*] { **deny** | **permit** } *ip-prefix* [**ge** *minimum-prefix-length*] [**le** *maximum-prefix-length*]
Parameter Description	*prefix-list-name*: Indicates the prefix-list name. *seq-number*: Assigns a sequence number to an prefix-list entry, ranging from 1 to 2,147,483,647. If this command does not contain the sequence number, the system will assign a default sequence number to the prefix-list entry. The default sequence number of the first entry is 5. Subsequently, the default sequence number of each entry not assigned with a value is the first multiple of 5 greater than the previous sequence number. **deny:** Denies access when certain conditions are matched. **permit:** Permits access when certain conditions are matched. *ip-prefix*: Configures the IP address and mask, ranging from 0 to 32 digits. *minimum-prefix-length*: Specifies the minimum range (namely, the start length of a range). *maximum-prefix-length*: Specifies the maximum range (namely, the end length of a range).
Command Mode	Global configuration mode
Usage Guide	-

Adding Description to a Prefix-List

Command	**ip prefix-list** *prefix-list-name* **description** *descripton-text*
Parameter Description	***prefix-list-name***: Indicates the prefix-list name. *descripton-text*: Describes the prefix-list.
Command Mode	Global configuration mode
Usage Guide	-

Enabling the Sorting Function for a Prefix-List

Command	**ip prefix-list sequence-number**
Parameter Description	**-**
Command Mode	Global configuration mode
Usage Guide	-

Creating an IPv6 Prefix-List

Command	**ipv6 prefix-list** *prefix-list-name* [**seq** *seq-number*] { **deny** | **permit** } *ipv6-prefix* [**ge** *minimum-prefix-length*] [**le** *maximum-prefix-length*]
Parameter Description	*prefix-list-name*: Indicates the prefix-list name. *seq-number*: Assigns a sequence number to an prefix-list entry, ranging from 1 to 2,147,483,647. If this command does not contain the sequence number, the system will assign a default sequence number to the prefix-list entry. The default sequence number of the first entry is 5. Subsequently, the default sequence number of each entry not assigned with a value is the first multiple of 5 greater than the previous sequence number. **deny:** Denies access when certain conditions are matched. **permit:** Permits access when certain conditions are matched. ***i****pv6-prefix*: Configures the IP address and mask, ranging from 0 to 128 digits. *minimum-prefix-length*: Specifies the minimum range (namely, the start length of a range). *maximum-prefix-length*: Specifies the maximum range (namely, the end length of a range).
Command Mode	Global configuration mode
Usage Guide	-

Adding Description to an IPv6 Prefix List

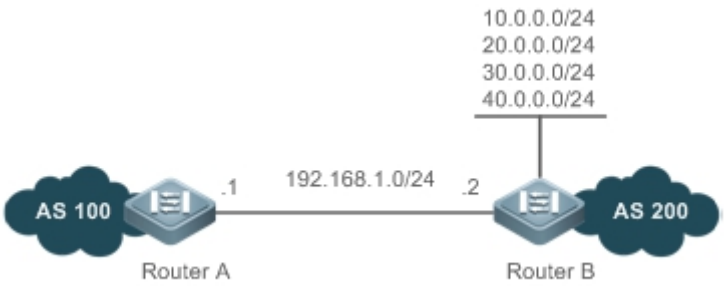
Command	**ipv6 prefix-list** *prefix-list-name* **description** *descripton-text*
Parameter Description	*prefix-list-name*: Indicates the prefix list name. *descripton-text*: Describes the prefix list.
Command Mode	Global configuration mode
Usage Guide	-

Enabling the Sorting Function for an IPv6 Prefix-List

Command	**ipv6 prefix-list sequence-number**
Parameter Description	**-**
Command Mode	Global configuration mode
Usage Guide	-

Configuration Example

Configuring a Community List

Scenario **Figure 7-4**	
Configuration Steps	● Define a standard community list to match the community attribute 100: 20. ● Advertise a route with the community attribute on B. ● Associate the community list on A to filter routes received on B.
A	A(config)# ip community-list standard test permit 100:20 A(config)# route-map COM A(config-route-map)# match community test A(config-route-map)# exit
B	B(config)# route-map comm1 B(config-route-map)# set community 100:20 200:20 B(config-route-map)# route-map comm2 B(config-route-map)# set community 100:20 B(config-route-map)# route-map comm3 B(config-route-map)# set community 200:20 B(config-route-map)# exit
Verification	● Run the **show** command to display the community list.

Common Errors

- A filtering list is configured but is not correctly applied in a routing protocol, which causes that the filtering list cannot take effect.

11.5 Monitoring

Displaying

Description	Command
Displays the configurations of a route map.	**show route-map** [*route-map-name*]
Displays the configurations of an ACL.	**show access-lists** [*id* | *name*]
Displays the configurations of an IPv4 prefix-list.	**show ip prefix-list** [*prefix-name*]
Displays the configurations of an IPv6 prefix-list.	**show ipv6 prefix-list** [*prefix-name*]
Displays the configurations of an AS-path list.	**show ip as-path-access-list** [*num*]
Displays the configurations of a community list.	**show ip community-list** [*community-list-number* | *community-list-name*]
Displays the configurations of an excommunity list.	**show ip excommunity-list** [*excommunity-list-number* | *excommunity-list-name*]

Chapter5 MPLS Configuration

1. Configuring MPLS
2. Configuring BGP/MPLS IP VPN
3. Configuring MPLS GR
4. Configuring MPLS BFD

1 Configuring MPLS

The routers and router icons described in this chapter refer to general routers as well as Layer-3 switches running routing protocols.

1.1 Overview

Multiprotocol Label Switching (MPLS) supports multiple network layer protocols, for example, Internet Protocol (IP), IP version 6 (Ipv6), and Internetwork Packet Exchange (IPX). It is compatible with multiple link layer technologies, for example, asynchronous transfer mode (ATM), frame relay, Ethernet, and Point to Point Protocol (PPP). Labels are attached to packets, and packets are forwarded based on labels. MPLS uses connectionless control plane and connection-oriented data plane, which adds connection-oriented attributes to connectionless IP networks. MPLS was originally intended to make routers forward data faster. With the development of hardware technologies and network processors, this advantage has faded. However, as MPLS integrates the Layer-2 switching technology with Layer-3 routing technology, it outperforms other technologies in solving problems of virtual private networks (VPNs) and traffic engineering (TE). MPLS VPNs get increasingly popular among carriers in regard to enterprise interconnection and new services, becoming an important means by which carriers provide value-added services on IP networks. MPLS TE also has become an important means of managing network traffic, reducing congestion, and ensuring the QoS of IP networks. Therefore, MPLS is attracting more and more attention and increasingly applied to VPNs and TE.

Protocols and Standards

RFC3032: MPLS Label Stack Encoding

RFC3036: LDP Specification

RFC4182: Removing a Restriction on the use of MPLS Explicit NULL

RFC 5283: LDP Extension for Inter-Area Label Switched Paths (LSPs)

1.2 Applications

Application	Description
Providing the MPLS VPN Service	The geographically different branches of a private network are connected through label switched paths (LSPs) to form a unified network.

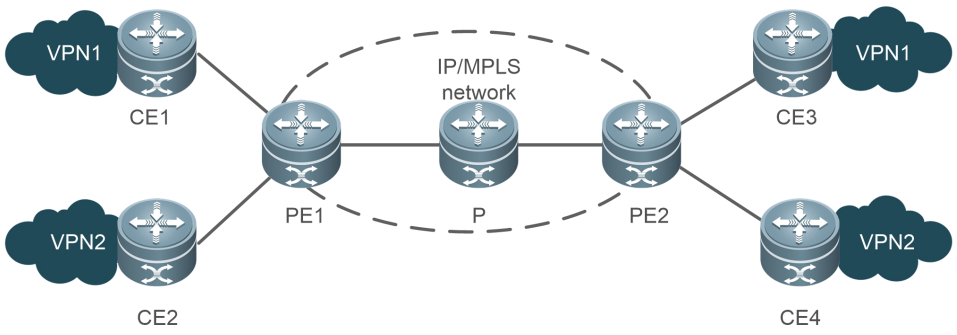
12.2.1 Providing the MPLS VPN Service

Scenario

MPLS Layer-3 VPN can be implemented through Multiprotocol Label Switching (MPLS).

Figure 1- 1 shows the application scenario of MPLS Layer-3 VPN.

Figure 1- 1 Application scenario of MPLS Layer-3 VPN



Deployment

- Public tunnels are established in a backbone network through the Label Distribution Protocol (LDP).
- Multiple VPN routing and forwarding (VRF) tables are created to maintain VPN routing information.
- Multiprotocol Extensions for BGP (MP-BGP) and Border Gateway Protocol (BGP) extended attributes are configured to transmit VPN routing information and carry VPN attributes and labels.

1.3 Features

Basic Concepts

MPLS Node

An MPLS node can identify the MPLS signaling protocol (control protocol), run one or more Layer-3 routing protocols (including static routing protocols), and forwards packets based on MPLS labels. Generally, an MPLS node is capable of forwarding original Layer-3 packets (for example, IP packets).

Forwarding Equivalence Class (FEC)

FEC refers to the packets forwarded equivalently, for example, packets sharing the same destination address or prefix. Packets can be classified in different ways in different application scenarios. In an IP unicast routing scenario, packets can be classified by address prefix; that is, a route corresponds to an FEC. Packets of the same FEC will be processed equivalently on an MPLS network.

Label

A label is a short identifier of a fixed length that takes effect locally. Labels are transmitted only between two adjacent label switching routers (LSRs) and thereby are valid only between the LSRs. One label categorizes one FEC. When arriving at the MPLS network ingress, packets will be classified into different FECs based on rules. Labels are encapsulated in packets, and the packets are forwarded based on labels on the MPLS network.

LSR

As the core of an MPLS network, LSRs switch and distribute labels. As described in RFC3031, an LSR is also an MPLS node capable of forwarding original Layer-3 packets (for example, IP packets or IPv6 packets), which means LSRs are capable of forwarding IP packets on IP networks.

Label Switching Edge Router (LER)

On the MPLS network edge, an LER classifies data flows entering the network into different FECs and requests labels. When data flows exit the network, the LER pops the labels to restore original packets. Therefore, an LER classifies data flows, maps labels, and removes labels.

LSP

Data flows of an FEC are given definite labels on different nodes. Labels are switched on all nodes during data forwarding. Data flows pass through an LSP, which is a collection of LSRs. An LSP can be considered as a tunnel that passes through the core MPLS network.

Next Hop Label Forwarding Entry (NHLFE)

An NHLFE is used to forward the data on the next hop of MPLS packets. An NHLFE includes the following content:

- The next hop of packets.
- Link layer encapsulation used to forward packets.
- Coding mode used to forward packet label stacks.
- Operations on packet label stacks.

Incoming Label Map (ILM)

An ILM table is a label forwarding table. It maps each incoming label to a series of NHLFEs (multiple NHLFEs indicate multiple paths). ILM is used when an LSR forwards labeled MPLS packets.

FEC-to-NHLFE (FTN)

Different from ILM, FTN maps each FEC to a series of NHLFEs (multiple NHLFEs indicate multiple paths). FTN is used when an LER receives unlabeled packets. The LER encapsulates labels into the packets and then forwards them.

LDP Peers

The LDP instances on two label switching routers (LSRs) that exchange label binding messages are called LDP peers.

LDP Sessions

An LDP session is established through the Transmission Control Protocol (TCP) between two LDP peers to exchange label binding messages.

Overview

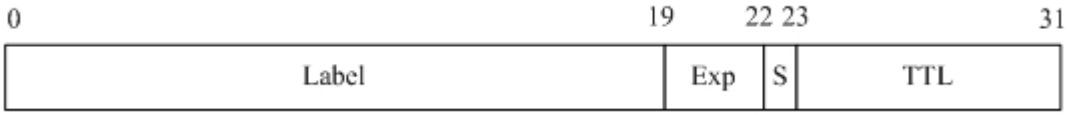
Feature	Description
MPLS Labels and Forwarding	This section describes the principles of labeling and label forwarding.
Static MPLS LSP	This section describes the process of establishing a static LSP through MPLS.
LDP and Session	This section describes the process of establishing an LDP session.
LDP Session Protection	This section describes the LDP session protection feature.
LDP LSP	This section describes the process of establishing an LSP through LDP.
Inter-Domain LDP LSP	This section describes the inter-domain LDP LSP feature.

1.3.1 MPLS Labels and Forwarding

Working Principle

Label Structure

Figure 1- 2 Coding structure of MPLS labels



As shown in the figure, a label contains four domains. The following describes the domains.

- Label domain

The label domain stores label values and contains 20 bits. A label value is the index value of the label forwarding table. The Internet Engineering Task Force (IETF) specifies 0 to 15 as retained label values and predefines them.

Retained label value	Definition
0	It indicates IPv4 explicit null labels. According to RFC3032, Labels 0 must appear on the stack bottom; routers pop the labels and then forward packets based on IP addresses. According to RFC4182, routers pop Labels 0 and then forward packets based on what follows the labels: if labels follow, routers forward packets based on labels; if IPv4 packets follow, routers forward them based on IP addresses.
1	It indicates router alert labels, which must not appear on the stack bottom. Packets with router alert labels must be processed by the local software module and forwarded based on the labels following router alert labels. However, before the packets are forwarded, router alert labels must be pushed into the stack again, which is similar to the **Router Alert** option of an IP packet. After this option is selected, the LSR will check the MPLS packets every hop.
2	It indicates IPv6 explicit null labels. According to RFC3032, Labels 2 must appear on the stack bottom; routers pop the labels and then forward packets based on IP addresses. According to RFC4182, routers pop Labels 2 and then forward packets based on what follows the labels: if labels follow, routers forward packets based on labels; if IPv6 packets follow, routers forward them based on IP addresses.
3	It indicates implicit null labels, which can be distributed through the label distribution protocol but never appear in the label stack of MPLS packets. When an LSR switches labels of MPLS packets, Labels 3 on the stack top will be popped rather than replaced. These labels will be used for "penultimate-hop popping (PHP)".
4 to 15	They will be retained by the IETF for future use.

- Exp domain

The Exp domain stores MPLS QoS information and contains three bits.

- S flag

It is a flag on the stack bottom and contains one bit. If there are multiple labels, the S flag of labels on the stack bottom will be set to 0 and others 1.

- Time To Live (TTL)

The TTL contains 8 bits, similar to that of IP packet headers. It can be copied from the TTL domain of IP packet headers (or IPv6 hop limit) when IP packets are labeled for the first time. The TTL value of outer (top) labels is decreased by 1 when labels are switched. The MPLS packets running on an ATM link are coded in a different way because they don't have a TTL domain. For more details, see RFC3032.

Label Stack

One MPLS packet may have multiple labels, which constitute a label stack. Labels close to the link layer header are called stack-top labels, and labels close to the IP header are called stack-bottom labels. LSRs always switch labels based on stack-top labels. Each label contains 32 bits. The label stack allows one MPLS packet to have multi-layer labels, which allows MPLS to support the layered network system as well as LSP tunnels.

Operations on Labels

The following describes the basic operations on labels by an MPLS node:

- Push (adding labels)

On the ingress LER, the node inserts labels between link layer header and network layer header; or transit LSRs add labels to the top of the label stack.

- Pop (popping labels)

On the egress LER, the node removes all labels to restore packets to IP packets; or transit LSRs remove stack-top labels to reduce label stack layers.

- Swap (switching labels)

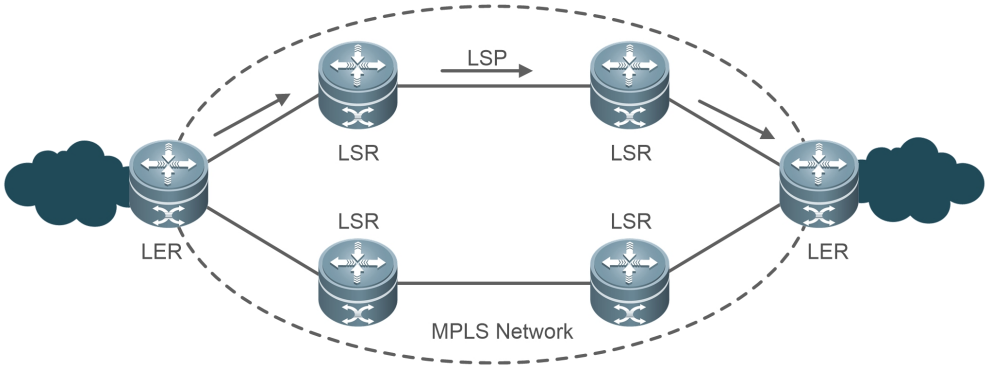
The node replaces stack-top labels based on the ILM table during packet forwarding.

Label Distribution Protocol

As a new network system, MPLS has its signaling protocol or "routing protocol". In the MPLS system, two LSRs must reach an agreement upon the meanings of labels attached to the data transmitted between them through LDP. An LSR notifies label binding to the other through LDP. MPLS has more than one LDP. MPLS uses independent distribution protocols, for example, LDP defined by LETF in RFC3036, and extends existing protocols through piggybacking to support label distribution, for example, MP-BGP and Resource Reservation Protocol (RSVP). In a simple topology, you can manually assign labels to LSRs and establish static LSPs. You can choose label distribution protocols based on scenarios.

MPLS Network

Figure 1- 3

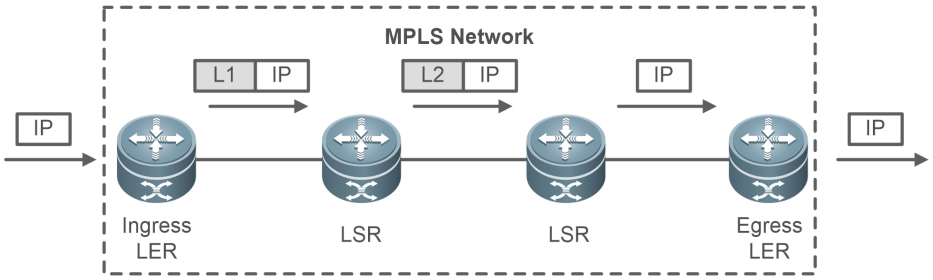


An MPLS network consists of two basic elements: LSR and LER. As the core of an MPLS network, LSRs run the MPLS signaling protocol LDP and forward labeled packets. When packets enter the MPLS network, LERs classify them into different FECs, label them, encapsulate them into MPLS packets, and forward them. When packets exit the network, LERs remove the labels to restore original packets. On an MPLS network, labeled packets will be forwarded through the LSP established through LDP.

The MPLS system consists of forwarding unit (data plane) and control unit (control plane). The forwarding unit searches the label forwarding information base and then forwards packets. The control unit creates a label forwarding information base between connected MPLS device nodes and maintains the base. Each MPLS node must run one or more routing protocols (including static routing protocols) to exchange routing information with other nodes. From the perspective of the control plane, each MPLS node is an IP router. Similar to traditional IP routers, MPLS nodes use unicast routing protocols (including static routing protocols) to establish and maintain routing tables. However, traditional routers use routing tables to establish forwarding tables, while MPLS nodes use routing tables to exchange label binding messages with adjacent nodes for destination subnets. A protocol used to exchange label binding messages is called LDP.

MPLS Forwarding

Figure 1- 4 Process of forwarding penultimate-hop popping MPLS packets



The following uses the example of traditional IP routing to describe MPLS forwarding.

- All LSRs (including LERs) use traditional routing protocols, for example, Open Shortest Path First (OSPF) and Intermediate System to Intermediate System (IS-IS), to establish IP routing tables.
- LSPs are established through LDP based on IP routing tables.
- The ingress LER receives an IP packet, analyzes the IP packet header and maps it into an FEC, attaches L1 to the packet, and forwards the labeled packet to the next-hop LSR through the LSP.
- The LSR receives the labeled packet and finds the LSP based on the label on the stack top, replaces the label, and forwards the packet to the next-hop LSR through the LSP.
- Transit LSRs repeat the previous step.

- The penultimate-hop LSR receives the labeled packet and checks the label forwarding table. If the outgoing label is an implicit null label (3), the LSR pops the label and forwards the original IP packet to the LSR of the last hop. If the label is an explicit null label, the LSR replaces the outermost label with a Label 0 and forwards the packet to the LSR of the last hop.
- If the penultimate-hop LSR pops the label, the egress LER of the last hop will receive an original IP packet and forward the packet based on the IP routing table. If the penultimate-hop LSR does not pop off the label because it is an explicit null label, the last-hop label edge router (LER) will receive the packet with the label value of 0 and pop off the label before forwarding the packet through a lookup into the IP routing table.

MPLS TTL Mode

After MPLS encapsulation and decapsulation, IP (or MPLS) packets can be processed in two TTL modes.

- TTL replication mode: It is a default mode. When Push operation is performed, the TTL of the pushed label is copied from the TTL domain of the existing IP or MPLS packet header to the TTL domain of the pushed label. When Pop operation is performed, the inner IP or MPLS packet TTL is copied back from the TTL domain of outer labels. When TTL replication is enabled, if the inner header TTL is smaller than the outer header TTL, the inner header TTL will not be copied or changed.
- TTL non-replication mode: The TTL will not be copied. **When Push operation is performed, the TTL of the pushed label is set to 255**.**When Pop operation is performed, the inner IP or MPLS packet TTL will not change.**

In TTL replication mode, LSRs that pass through the MPLS domain can be tracked by using Tracert on the Customer Edge (CE). In TTL non-replication mode, LSPs will be processed as one hop.

Related Configuration

Enabling MPLS Forwarding Globally

- By default, MPLS forwarding is disabled globally.
- Run the **mpls ip** command in global configuration mode to enable devices to support MPLS forwarding. After MPLS forwarding is enabled, the devices implement label switching preferentially; when label switching fails, the devices switch to IP forwarding.

Enabling Label Forwarding on an Interface

- By default, label forwarding is disabled on interfaces.
- Run the **label-switching** command to enable MPLS forwarding on the specified interface.When you enable label forwarding on an interface, adjust the interface's maximum transmission unit (MTU) based on the service type; otherwise, the transmission of large packets may be affected.

Configuring the MPLS MTU of an Interface

- The MPLS MTU determines whether an MPLS packet needs to be fragmented before being sent. The length defined by the MPLS MTU includes the total length of the MPLS encapsulation layer and encapsulated layer (such as IP).
- By default, the MPLS MTU of an interface is the same as its MTU.
- Run the **mpls mtu** command to configure the MPLS MTU of an interface.

Configuring MPLS Packet Fragmentation

- You can configure the processing mode for MPLS packets whose sizes exceed the MPLS MTU.
- By default, IP fragmentation is implemented. The IP packet fragments are encapsulated with the original label and forwarded along the LSP associated with this label.
- Run the **mpls ip fragment** command to configure MPLS packet fragmentation.

Configuring the ICMP Error Packet Processing Mode

- Configure the processing mode for ICMP error packets that are generated during MPLS packet forwarding. ICMP error packets include those generated when MPLS time to live (TTL) times out.
- By default, ICMP error packets are forwarded along the LSP associated with the label stack.
- Run the **mpls ip icmp-error pop labels** command to configure packet processing based on the label quantity. If the number of labels on the stack is smaller than or equal to the specified label quantity, the ICMP error packet will be forwarded through a lookup into the IP routing table that contains the forwarding equivalence class (FEC) of the topmost label. If the number of labels on the stack is greater than the specified label quantity, the ICMP error packet will be forwarded along the LSP associated with the label stack.

Configuring MPLS TTL Replication

- By default, TTL replication is applied to the packets that the local device sends and forwards.
- Run the **mpls ip ttl propagate** { **public** | **vpn** } command to control TTL replication on the local device when sending and forwarding packets.

1.3.2 Static MPLS LSP

You can enable MPLS forwarding through static configuration without LDP.

As LDP is not required for static configuration, IPv4 routes are not required. Even if no IPv4 route is available, static LSPs will take effect if the physical network is accessible.

Working Principle

Configuring Static FTN on the Ingress LSR

At the LSP ingress, you need to establish FTN entries for FECs; that is, you need to bind FECs with labels.

Configuring Static ILM on Transit LSRs

On transit LSRs, you need to configure ILM forwarding entries to map incoming labels to outgoing labels.

Configuring Static ILM on the Penultimate-hop LSR

As PHP will be enabled on the penultimate hop, ILM entries on the penultimate hop must be different from those on transit LSRs; that is, the outgoing label of ILM on the penultimate-hop LSR must be an implicit null label (3).

For information on PHP, see related materials.

Related Configuration

Configuring Static FTN on the Ingress LSR

- In global configuration mode, run the **mpls static ftn** command to configure static FTN.
- The ingress LSR must be configured with FEC to NHLFE (FTN) to forward IP traffic to an LSP.
- If you configure FTN with the outgoing label value of 3, you must configure IP routes with the same prefix to make FTN effective.

Configuring Static ILM on Transit LSRs

- In global configuration mode, run the **mpls static ilm in-label** command to configure static ILM.
- Transit LSRs must be configured with incoming label map (ILM) to forward MPLS packets.
- It is recommended that the outgoing label value of 3 be configured on the penultimate-hop LSR so that the egress LSR can directly receive IP packets; otherwise, you need to configure ILM on the egress LSR.

1.3.3 LDP and Session

LDP is mainly used to exchange label binding messages between IP unicast routes and Martini Layer-2 VPN.

This section describes how to establish LSPs on unicast routes.

LDP peer discovery and LDP session establishment are the prerequisites of LSP establishment.

Working Principle

LDP Router ID

An LDP Router ID uniquely identifies an LSR in a domain. It is in the IPv4 address format. By default, the LDP Router ID is the same as the Router ID (LSR ID) and must be globally unique.

LDP Peer Discovery and Maintenance

LDP periodically sends User Datagram Protocol (UDP) Hello packets to advertise its state and listens to Hello packets to discover LDP peers.

If LDP discovers an LDP peer but has not received a Hello packet from this peer for a period of time, LDP determines that the LDP peer is invalid. This period is called the Hello packet keep-alive period.

LDP Session Establishment and Maintenance

When LDP receives a Hello packet from an LDP peer, it uses the transport address advertised by the Hello packet to set up a TCP connection and initializes message exchange through the connection to negotiate LDP session parameters, in order to establish an LDP session.

After the LDP session is established, the two LDP peers periodically send keepalive packets to monitor the TCP connection. If a peer has not received a keepalive packet from the other peer for a period of time, it determines that the TCP connection is invalid and closes the LDP session. This period is called the keepalive packet keep-alive period.

Related Configuration

Enabling LDP Globally

- By default, LDP is disabled globally.
- Run the **mpls router ldp** [*vrf-name*] command in global configuration mode to enable LDP for a VRF instance and enter LDP configuration mode.

Enabling LDP on an Interface

- By default, LDP is disabled on interfaces.
- Run the **mpls ip** command in interface configuration mode to enable LDP on an interface.

Configuring an LDP Router ID

- By default, the LDP Router ID is the same as the Router ID.
- Run the **ldp router-id** command to modify the LDP Router ID.

Configuring a Transport Address

By default, the transport address used by an LDP session is the same as the LDP Router ID.

Run the **transport-address** command to specify the primary address of an interface or an IP address as the transport address used by the LDP session on the interface.

You can configure the transport address in either of the following modes:

- Run the **mpls ldp transport-address** command in interface configuration mode.
- Run the **transport-address** command in LDP configuration mode to globally configure the transport address used by all LDP sessions.

Configuring the Hello Packet Interval

- The default Hello packet interval used by the basic LDP discovery mechanism is 5s.
- You can change the Hello packet interval in the range from 1s to 65,535s in interface configuration mode.

Configuring the Hello Packet Keep-Alive Period

- By default, the default Hello packet keep-alive period used by the basic LDP discovery mechanism is 15s.
- You can change the Hello packet keep-alive period in the range from 1s to 65,535s in interface configuration mode. If it is set to 65,535s, the keep-alive period is infinitely long.

Configuring the Keepalive Packet Keep-Alive Period

- By default, the keep-alive period of keepalive packets sent through an interface is 45s.
- You can change the keep-alive period in the range from 15s to 65,535s. The keepalive packet interval is one-third of its keep-alive period.

Configuring the Maximum PDU Size

- LDP entities exchange messages as protocol data units (PDUs). By default, the maximum PDU size is 4,096. You can change it in the range from 256 to 4,096.

Configuring LDP MD5 Authentication

- You can configure MD5 authentication of TCP connections to improve LDP session security.
- By default, LDP MD5 authentication is not configured.

- Run the **neighbor** *ip-address* **password** [**0 | 7**] *pwd-string* command to configure MD5 authentication of the TCP connection with the specified address.

1.3.4 LDP Session Protection

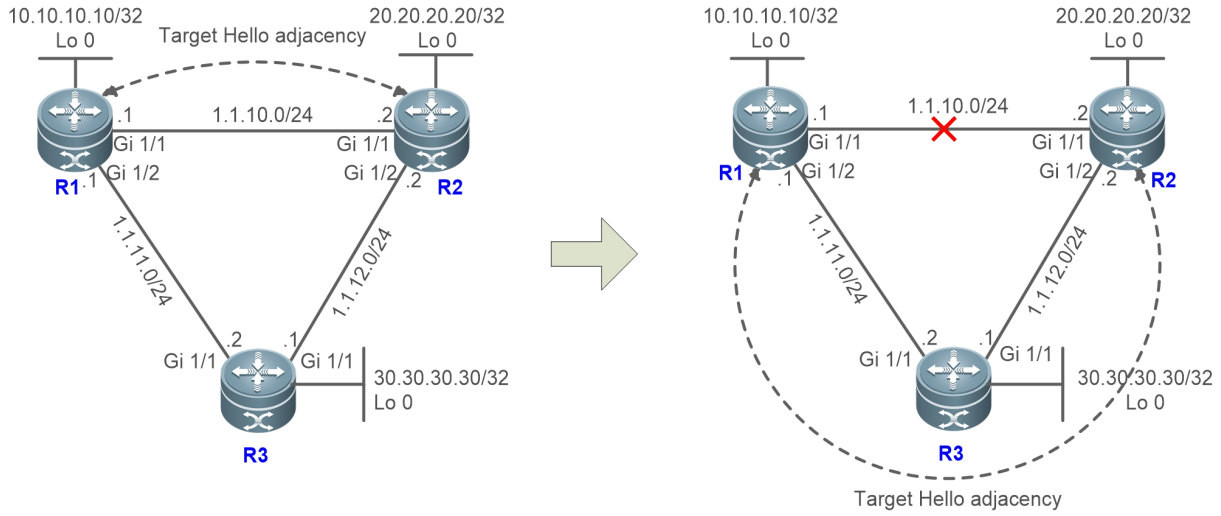
Faulty links will seriously affect network convergence and aggregation because the Hello adjacency between associated interfaces for LDP sessions will be disconnected when links fail and thereby LDP sessions will be closed. When links recover from failures, LSRs need to establish LDP sessions and exchange label binding messages again, which reduces the efficiency of LDP convergence.

LDP session protection can protect LDP sessions between two directly connected LSRs and LDP over TE sessions, and accelerate LDP convergence in case of link failures, especially when failures last for only a short time.

Working Principle

After LDP sessions between two LSRs are protected, a target Hello adjacency will be established between the LSRs. When the link between them fails, if alternative paths exist, the target Hello adjacency will be maintained. When the link fails, the link Hello adjacency will be disconnected, but the target Hello adjacency will maintain LDP sessions. When the link recovers from failures, LSRs do not need to establish LDP sessions again, which improves the efficiency of LDP convergence.

Figure 1- 5 LDP session protection



As shown in the figure, R1, R2, and R3 are interconnected with each other at the link layer with OSPF enabled. LDP is enabled between R1 and R2. A link Hello adjacency will be established between R1 and R2 so that LDP sessions will be established. To protect LDP sessions between R1 and R2, you can enable LDP session protection on R1 and R2. In this way, a target Hello adjacency will be established between R1 and R2.

When the link between R1 and R2 fails, the link Hello adjacency will be disconnected. However, the target Hello adjacency will be maintained because of another path R1-R3-R2. Therefore, LDP sessions will not be closed.

When the link recovers from failures, R1 and R2 do not need to establish LDP sessions or exchange label binding messages again.

Related Configuration

Configuring LDP Session Protection

- By default, LDP instance session protection is disabled.
- Run the **session protection** command to enable or disable LDP session protection.
- You can enable LDP session protection on the two LSRs engaged in an LDP session. If only one LSR supports LDP session protection, you can configure this function on this LSR and run the **discovery targeted-hello accept** command on the other LSR to receive and reply to target Hello packets (or run the **neighbor ip-address** command to specify the remote peer).

1.3.5 LDP LSP

The LDP LSPs described in this section are the LSPs that LDP establishes on IP unicast routes.

Working Principle

Process of Establishing an LSP

Establishing an LSP means binding FECs with labels and notifying the operation to the adjacent LSR. This process is implemented through LDP based on RFC3036.

After LDP sessions are established, two LSRs can exchange label messages to bind labels and establish LSPs for FECs.

Labels can be distributed in Downstream on Demand (DOD) mode and Downstream Unsolicited (DU) mode. In DOD mode, an LSR replies to a label binding message only when receiving a label request from the adjacent LSR. In DU mode, an LSR distributes labels to the adjacent LSR without receiving requests.

Labels can be controlled in independent mode and ordered mode. In independent mode, each LSR can notify its adjacent LSR of label-FEC binding. In independent DOD mode, an LSR can immediately reply to label mapping requests from an upstream LSR rather than wait for label mapping messages from the next hop. In independent DU mode, an LSR can notify label mapping whenever it is ready to switch labels for an FEC.

In ordered mode, an LSR distributes upstream the label bound with an FEC only when it receives the label mapping message from the next hop or when it is the egress for the FEC; otherwise, the LSR must wait for the label mapping message from a downstream LSR. In ordered DU mode, an LSR notifies label mapping to an upstream LSR only when it is the egress for an FEC or when it receives the label from a downstream LSR. If downstream LDP session labels are distributed in DOD mode, the LSR will send downstream the label request from an upstream LSR.

If a label mapping message is not from the next hop of an FEC or if the message does not match any FEC, the message can be processed based on the label retention mode.

In liberal label retention mode, FEC label binding learned from the neighbor is retained; in conservative label retention mode, FEC label binding learned from the neighbor is not retained.

In conservative label retention mode, there are only a few labels to be used and maintained; however, when routes change, label values need to be obtained again, which increases the response time. In liberal label retention mode, LSRs respond to route changes quickly; however, unnecessary labels need to be distributed and maintained.

LSP Loop Control

A loop detection mechanism is required. Two methods are available to prevent a loop: maximum hop count and path vector.

Maximum hop count: The number of LSRs is included in a label binding message. It is increased by 1 when an LSR passes. When the number exceeds the specified maximum, it indicates that a loop occurs, and LSP establishment will be stopped.

Path vector: The LSR ID is recorded in the vector list of a label binding message when an LSR passes. When receiving a label binding message, an LSR checks whether its ID is recorded in the vector list. If the ID is not recorded, the LSR will add it to the list. If the ID is recorded, it indicates that a loop occurs, and LSP establishment will be stopped.

Policy-Based Label Distribution Control

By default, LDP distributes labels to all valid Interior Gateway Protocol (IGP) routes, except BGP routes. In special cases, you may want to distribute labels only to specified routes or LDP peers to reduce labels and LSPs used, in order to alleviate device and network burden. To do this, you can configure label distribution policies to achieve either or both of the following effects:

- Distribute Label Mapping messages with corresponding FECs only to specified IP routes.
- Distribute Label Mapping messages only to specified LDP peers.

Policy-Based Label Reception Control

By default, LDP receives all Label Mapping messages over LDP sessions in DU mode and DOD mode. If an LDP session is in DOD mode, LDP sends the peer Label Request messages with FECs indicating the peer as the next hop. You can configure reception policies to achieve the following effects:

- Only receive the Label Mapping messages with FECs mapped to the specified IP route from the specified peer.
- Only send the Label Request messages with the specified FEC to the specified LDP peer.
- Do not receive Label Mapping messages from the specified peer.
- Do not send Label Request messages to the specified peer.

Related Configuration

Configuring the Label Distribution Control Mode

- By default, label distribution is controlled in independent mode.
- Run the **lsp-control-mode** command to configure the LDP label distribution control mode.

Configuring the Label Distribution Mode

- By default, labels are distributed in DU mode.
- Run the **mpls ldp distribution-mode** command in interface configuration mode to configure the LDP label distribution mode on an interface.

Configuring the Label Retention Mode

- By default, labels are retained in liberal mode.
- Run the **label-retention-mode** command to configure the label retention mode.

Configuring Loop Detection

- By default, loop detection is disabled.

- Run the **loop-detection** command to configure loop detection.

Configuring the Maximum Number of Hops

- Run the **mpls ldp max-hop-count** command in interface configuration mode to configure the maximum number of hops in loop detection. The value ranges from 1 to 255. The default value is 254.
- After loop detection is enabled, loop occurs if the number of hops traversed by an LDP message exceeds the threshold.

Configuring the Maximum Number of LSRs in the Path Vector List

- Run the **mpls ldp max-path-vector** command to configure the maximum number of LSRs in the path vector list used for loop detection. After loop detection is enabled, loop occurs if an LSR detects that its LSR ID is contained in the path vector list or the number of LSR IDs in the path vector list exceeds the threshold. The default value is 254.
- Run the **mpls ldp max-path-vector** command in interface configuration mode to configure the maximum number of LSRs in the path vector list.

Configuring the Label Release Message Propagation Mode

- When an FEC is invalid, LDP will send a label release message from upstream to downstream to revoke the label bound to the FEC. The label release message propagation mode determines whether each receiving LDP entity on an LSP propagates the label release message to the downstream.
- By default, the receiving LDP entity does not propagate the label release message to the downstream.
- Run the **propagate-release** command to configure the label release message propagation mode.

Configuring Label Distribution Policies

- Run the **advertise-labels for host-routes** command to distribute labels only to the host routes with a 32-bit mask in VRF tables. By default, labels are distributed to host routes with various mask lengths.
- Run the **advertise-labels for bgp-routes** [**acl** *acl-name*] command to distribute labels to BGP routes. By default, LDP does not distribute labels to BGP routes.
- Run the **advertise-labels for default-route** command to distribute labels to default routes. By default, LDP distributes implicit null labels with the value of 3 to default routes.
- Run the a**dvertise-labels for acl** *prefix-access-list* [**to** *peer-access-list*] command to distribute labels only to FECs and LDP peers that meet access control list (ACL) rules.

Configuring Label Reception Policies

- Run the **neighbor** *ip-address* **labels accept** *acl-name* command to configure label reception policies. By default, LDP receives label binding messages over all sessions.

Configuring Explicit Null Label Distribution

- By default, LDP distributes implicit null labels to FECs with local devices as the egress (for example, direct routes).
- Run the **explicit-null** command to distribute explicit null labels to all direct routes or direct routes that meet specified ACL rules.

1.3.6 Inter-Domain LDP LSP

When an autonomous system (AS) is divided into multiple IGP areas, to downsize the routing table and improve the efficiency, you need to configure host route aggregation on area border routers (ABRs). You can establish an LSP across multiple LGP areas for the route through the inter-area LDP LSP technology.

Such LSP is established by extending the longest route match.

Working Principle

Inter-area LSP

An inter-area LSP crosses multiple IGP areas of an AS.

ABR

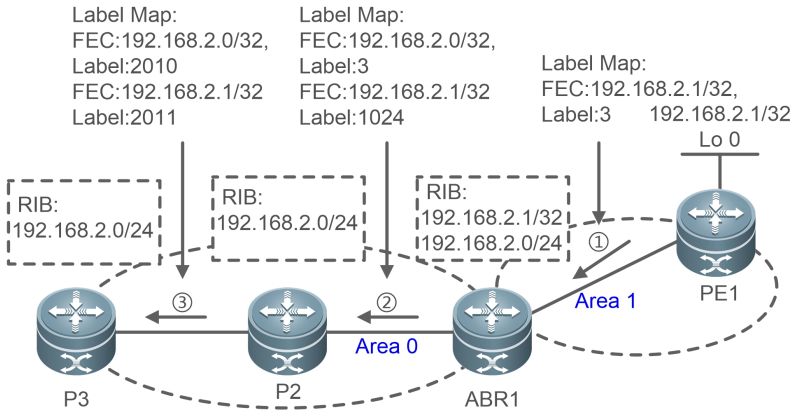
For OSPF, an ABR is a node connecting backbone areas to other areas; for IS-IS, an ABR is a node between Level 1 and Level 2.

When an AS is divided into multiple IGP areas and host routes are aggregated on ABRs, you can enable the inter-area LDP LSP function to establish an LSP across multiple IGP areas for the host routes.

When the inter-area LDP LSP function is disabled, LDP forwards labels only to accurately matched routes and distributes label mapping messages to all LDP neighbors. When the function is enabled, if there is a route with the longest mask that is matched to the FEC in a label mapping message and the neighbor sending the label mapping message corresponds to the next hop of the route, LDP will add forwarding entries and distributes label mapping messages to LDP neighbors (based on FECs in received label mapping messages rather than FECs matched to the route with the longest mask in the routing table).

The following figure shows the process of establishing an inter-area LDP LSP.

Figure 1-6



Based on the figure, the following describes the process of establishing an inter-area LDP LSP.

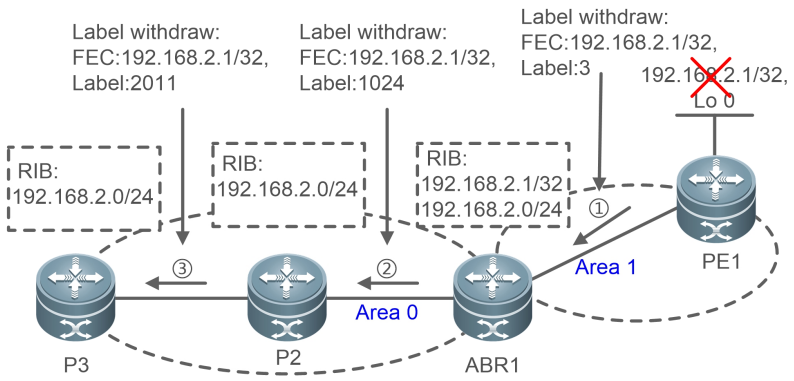
- On ABR1, aggregated routes are notified to backbone areas. Both aggregated route (192.168.2.0/24) and host route (192.168.2.1/32) are available in the routing table of ABR1.
- PE1 sends ABR1 a label mapping message of the FEC 192.168.2.1/32.

- ABR1 sends P2 label mapping messages of FECs 192.168.2.0/24 and 192.168.2.1/32, respectively.
- Only the aggregated route 192.168.2.0/24 is available in the routing table of P2. P2 sends P3 the label mapping message of the FEC 192.168.2.0/24. As there is a route with the longest mask (192.168.2.0/24) that is matched to the FEC 192.168.2.1/32 and ABR1 corresponds to the next hop of the route, P2 also sends P3 the label mapping message of the FEC 192.168.2.1/32.
- P3 repeats the actions of P2.
- An inter-area LSP across P3, P2, ABR1, and PE1 (192.168.2.1/32) is established between Area 0 and Area 1.

This process is similar to the process in ordered DU mode (as only aggregate routes are available upstream, the inter-area LDP LSP technology is inapplicable to the DOD mode). The LSP established in this way is the shortest IGP path.

If the MPLS forwarding entries added based on the longest match principle need to be deleted because of label cancellation messages from a downstream device or deletion of the matched route with the longest mask, the label cancellation messages need to be sent upstream (whether label distribution is controlled in ordered or independent mode). The following figure shows the process of canceling an inter-area LSP.

Figure 1-7



If the matched route with the longest mask changes (a better route is available or the route is deleted), the routing platform will notify LDP to update or delete the MPLS forwarding entries added based on the longest match principle to improve the accuracy of LSPs.

Related Configuration

Configuring the Inter-Domain LDP LSP Feature

- By default, the inter-area LDP LSP function is disabled.
- Run the **inter-area-lsp** command to enable or disable the inter-domain LDP LSP feature.
- The inter-domain LDP LSP feature only takes effect for aggregated host routes. If a longest match ACL rule is configured for aggregated non-host routes, this rule will be invalid.

1.4 Configuration

Configuration	Description and Command	
Configuring MPLS Forwarding	(Mandatory) It is used to enable MPLS.	
	mpls ip	Enables MPLS forwarding globally.
	label-switching	Enables MPLS forwarding on an interface.
	(Optional) It is used to configure MPLS forwarding parameters.	
	mpls mtu	Configures the MPLS MTU of an interface.
	mpls ip fragment	Configures MPLS packet fragmentation.
	mpls ip icmp-error pop	Configures the processing mode for ICMP error packets that are generated during MPLS packet forwarding.
	mpls ip ttl propagate public	Configures MPLS TTL replication.
Configuring a Static MPLS LSP	(Mandatory) It is used to enable MPLS.	
	mpls static ftn out-label nexthop	Configures static FTN.
	mpls static ilm in-label forward-action swap-label nexthop fec	Configures static ILM.
Configuring an LDP Session	(Mandatory) It is used to establish an LDP session.	
	mpls router ldp	Enables LDP.
	mpls ip	Enables LDP on an interface.
	neighbor	Creates a remote LDP peer.
	discovery targeted-hello accept	Configures the reception of target LDP Hello packets.
	(Optional) It is used to configure LDP session parameters.	
	ldp router-id interface loopback	Configures an LDP Router ID.
	mpls ldp transport-address	Configures the transport address used by the local LDP session on an interface.
	transport-address	Configures the transport address used by the local LDP session in global configuration mode.
	mpls ldp hello-interval	Configures the Hello packet interval on every interface.
	discovery targeted-hello	Configures the Hello packet interval or keep-alive period used by the extended LDP discovery mechanism.
	mpls ldp hello-holdtime	Configures the Hello packet keep-alive period on every interface.
	mpls ldp keepalive-holdtime	Configures the keepalive packet keep-alive period on an interface.

Configuration	Description and Command	
	targeted-session holdtime	Configures the keepalive packet keep-alive period used by the remote LDP session.
	mpls ldp max-pdu	Configures the maximum length of every LDP message.
	neighbor *ip-address* **password**	Enables LDP MD5 authentication.
	route wait label-mapping *seconds*	Configures the waiting time for label mapping during route updating.
Configuring LDP Session Protection	(Mandatory) It is used to enable LDP session protection.	
	session protection	Configures LDP session protection.
Configuring an LDP LSP	(Optional) It is used to establish an LDP LSP.	
	loop-detection	Enables loop detection.
	mpls ldp max-hop-count	Configures the maximum number of hops in loop detection on an interface.
	mpls ldp max-path-vector	Configures the maximum number of LSRs in the path vector list used for loop detection on an interface.
	propagate-release	Enables label release message propagation.
	lsp-control-mode	Configures the LDP control mode globally.
	mpls ldp distribution-mode	Configures the LDP label distribution mode on every interface.
	label-retention-mode	Configures the label retention mode.
	advertise-labels	Configures label distribution policies for FECs mapped to IP routes.
	neighbor *ip-address* **labels accept** *acl-name*	Configures LSRs to filter Label Mapping messages sent by specified LDP peers based on specified ACL rules.
	explicit-null	Configures distribution of explicit null labels to all direct routes or direct routes that meet specified ACL rules, or configures distribution of explicit null labels to neighbors that meet specified ACL rules and implicit null labels to other neighbors.
Configuring the Inter-Domain LDP LSP Feature	(Mandatory) It is used to enable the inter-domain LDP LSP feature.	
	inter-area-lsp	Configures the inter-domain LDP LSP feature.

1.4.1 Configuring MPLS Forwarding

Configuration Effect

- Enable MPLS forwarding on all LSRs and corresponding interfaces traversed by an LSP. The device forwards packet labels first rather than IP packets. When label forwarding fails, the device forwards IP packets.

Notes

- Enable MPLS forwarding in global configuration mode and interface configuration mode.

Configuration

Enabling MPLS Forwarding Globally

- (Mandatory) Perform this configuration in global configuration mode.
- (Mandatory) Perform this configuration in global configuration mode on every LSR traversed by an LSP, unless otherwise specified.

Enabling MPLS Forwarding on an Interface

- (Mandatory) Perform this configuration in interface configuration mode.
- (Mandatory) Perform this configuration in interface configuration mode on every interface traversed by an LSP, unless otherwise specified.

Configuring the MPLS MTU of an Interface

- (Optional) Perform this configuration in interface configuration mode.
- This configuration is required only in special cases.

Configuring MPLS Packet Fragmentation

- (Optional) Perform this configuration in global configuration mode.
- Perform this configuration on every LSR traversed by an LSP, unless otherwise specified. If MPLS packet fragmentation is not enabled, LSRs cannot forward large packets that exceed the MTU size. By default, MPLS packet fragmentation is enabled.

Configuring the ICMP Error Packet Processing Mode

- (Optional) Perform this configuration in global configuration mode.
- This configuration is required only in special cases.

Configuring MPLS TTL Replication

- (Optional) Perform this configuration in global configuration mode.
- Disable MPLS TTL replication on every LSR traversed by an LSP if you want to hide the MPLS network topology.

Verification

Run the **show mpls summary** command to display the list of interfaces enabled with MPLS forwarding.

Run the **show run** command to check whether the configuration takes effect.

Related Commands

Enabling MPLS Forwarding Globally

Command	**mpls ip**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling MPLS Forwarding on an Interface

Command	**label-switching**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	MPLS interface commands can be run only on Layer-3 interfaces. Therefore, non-Layer-3 interfaces need to be converted into Layer-3 interfaces.

Configuring the MPLS MTU of an Interface

Command	**mpls mtu** *bytes*
Parameter Description	bytes: Indicates the length of an MPLS packet, which ranges from 64 to 1,500 bytes by default.
Command Mode	Interface configuration mode
Usage Guide	1. The MPLS MTU must not exceed the size of transmitted packets. You can run the **mpls mtu** command to adjust the MTU in interface configuration mode. 2. After encapsulating MPLS packets, devices that support MPLS packet fragmentation will fragment the packets whose length exceeds the MPLS MTU based on the MPLS MTU. Devices that do not support MPLS packet fragmentation will discard the packets whose length exceeds the MPLS MTU.

Fragmenting MPLS Packets

Command	**mpls ip fragment**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	1. This command is valid for IP packets only. 2. On devices that support MPLS packet fragmentation, you can use the **no** form of this command to disable MPLS packet fragmentation. On devices that do not support MPLS packet fragmentation, this command is invalid.

Processing ICMP Error Packet

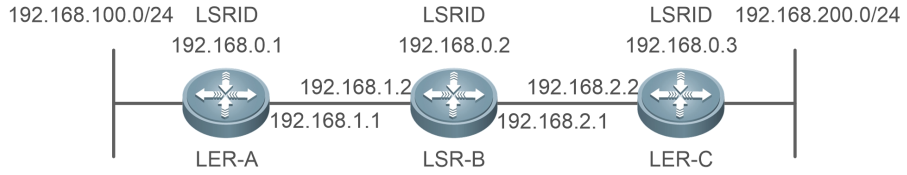
Command	**mpls ip icmp-error pop** *labels*
Parameter Description	*labels:* Indicates the number of labels of packets to be processed.
Command Mode	Global configuration mode
Usage Guide	You can run this command to have packets processed based on the number of labels. If the number of labels exceeds the configured value, the device will forward the packets through the LSP of the label stack. If the number of labels is below or equal to the configured value, the device will forward the ICMP error packets based on the IP routing table of the FEC corresponding to stack-top labels.

Configuring MPLS TTL Replication

Command	**mpls ip ttl propagate public**
Parameter Description	**public:** Enables or disables TTL replication for sent packets. **vpn:** Enables or disables TTL replication for forwarded packets.
Command Mode	Global configuration mode
Usage Guide	When TTL replication is enabled, if the inner header TTL is smaller than the outer header TTL, the inner header TTL will not be copied or changed.

Configuration Example

Configuring MPLS Forwarding

Scenario **Figure 1-8**	
Configuration Steps	● Configure MPLS forwarding on the ingress LSR. ● Configure MPLS forwarding on transit LSRs. ● Configure MPLS forwarding on the penultimate-hop LSR.
LER-A	FS(config)# mpls ip FS(config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# label-switching FS(config-if-GigabitEthernet 2/2)# end
LSR-B	FS(config)# mpls ip FS(config)# interface gigabitEthernet 2/1

	FS(config-if-GigabitEthernet 2/1)# label-switching FS(config-if-GigabitEthernet 2/1)# exit FS(config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# label-switching FS(config-if-GigabitEthernet 2/2)# exit FS(config)# end
LER-C	FS(config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# label-switching FS(config-if-GigabitEthernet 2/1)# end
Verification	Check whether MPLS forwarding is enabled in interface configuration mode on all routers.
LER-A	A# show mpls summary Minimum label configured: 16 Maximum label configured: 1048575 Per label-space information: Label-space 0 is using minimum label: 16 and maximum label: 1048575 Label-Switching Interface: Interface Label space Gi2/2 0 Total number of mpls interface is 1
LSR-B	B# show mpls summary Minimum label configured: 16 Maximum label configured: 1048575 Per label-space information: Label-space 0 is using minimum label: 16 and maximum label: 1048575 Label-Switching Interface: Interface Label space Gi2/1 0 Gi2/2 0

	Total number of mpls interface is 2
LER-C	C# show mpls summary Minimum label configured: 16 Maximum label configured: 1048575 Per label-space information: Label-space 0 is using minimum label: 16 and maximum label: 1048575 Label-Switching Interface: Interface Label space Gi2/1 0 Total number of mpls interface is 1

Common Errors

- MPLS forwarding is not enabled globally.
- LDP is not enabled in interface configuration mode.

1.4.2 Configuring a Static MPLS LSP

Configuration Effect

Establish a static LSP

Notes

- Enable MPLS forwarding in global configuration mode and interface configuration mode.
- When you configure static TFN with the label value of 3, configure IP routes with the same next hop for forwarding.

Configuration

Configuring MPLS Forwarding on All LSRs

- Perform this configuration according to section 1.4.1 "Configuring MPLS Forwarding."

Configuring Static FTN on the Ingress LSR

- (Mandatory) Perform this configuration in global configuration mode.

Configuring Static ILM on Transit LSRs

- (Mandatory) Perform this configuration in global configuration mode.

Verification

Run the **show mpls forwarding-table** command to display MPLS entries.

Related Commands

Configuring Static FTN on the Ingress LSR

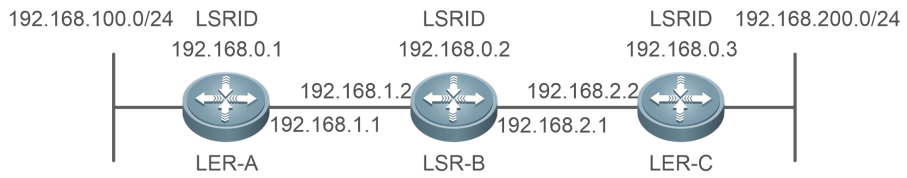
Command	**mpls static ftn** *ip-address/mask* **out-label** *label* **nexthop** *interface nexthop-ip*
Parameter Description	*ip-address/mask:* Indicates the address of an FEC, namely the destination address. **out-label** *label:* Indicates the outgoing label of the FEC. The value ranges from 16 to 1048575, or is 3. **nexthop** *interface-name nexthop-ip:* Indicates the next hop of the FEC, including the next-hop outgoing interface and IP address.
Command Mode	Global configuration mode
Usage Guide	You can run this command to add an FTN entry to the global FTN table. Upon receipt of an IP packet, the device on which MPLS is enabled finds the next hop based on the destination address of the IP packet by following the longest match principle, and then forwards the label of the IP packet. This command is valid for FTN entries whose destination address and mask are 0 only when the default route is available in the IP routing forwarding table.

Configuring Static ILM on Transit LSRs

Command	**mpls static ilm in-label** *in_label* **forward-action swap-label** *swap_label* **nexthop** *interface nexthop-ip* **fec** *ip-address/mask*
Parameter Description	*in_label:* Indicates the incoming label of an ILM entry. The value ranges from 16 to 1025. **forward-action:** Specifies the forwarding action of the ILM entry. *label:* Specifies the value of the outgoing label when swap-label action is performed. The value ranges from 16 to 1048575, or is 0 or 3. *vrf-name:* Specifies the VPN for the ILM entry (namely VRF) when pop-l3vpn-nexthop action is performed. *interface-name:* Specifies the egress when pop-l2vc-destport action is performed. **nexthop** *interface-name nexthop-ip:* Specifies the next hop, including the next-hop egress and IP address. **fec:** Indicates the FEC for which the ILM entry is created. *ip-address/mask:* Specifies a destination network based on the FEC format of global or Layer-3 VPN application. *vc_id:* Specifies a virtual circuit (VC) instance based on the FEC format of Layer-2 VPN application. *The value ranges from 1 to 2147483647.* *vc-peer-addr:* Indicates the address of a VC peer.
Command Mode	Global configuration mode
Usage Guide	You can run this command to add an ILM entry to the ILM table. Upon receipt of an IP packet, the device on which MPLS is enabled finds the next hop based on the IP packet label value by following the longest match principle, and then forwards the label.

Configuration Example

Configuring a Static MPLS LSP Network

Scenario **Figure 1-9**	
Configuration Steps	● Configure MPLS forwarding and static FTN on the ingress LSR. ● Configure MPLS forwarding and static ILM on transit LSRs. ● Configure MPLS forwarding and static ILM on the penultimate-hop LSR.
LER-A	FS(config)# mpls ip FS(config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# label-switching FS(config-if-GigabitEthernet 2/2)# exit FS(config)# mpls static ftn 192.168.200.0/24 out-label 16 nexthop gigabitEthernet 2/2 192.168.1.2 FS(config)# end
LSR-B	FS(config)# mpls ip FS(config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# label-switching FS(config-if-GigabitEthernet 2/1)# exit FS(config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# label-switching FS(config-if-GigabitEthernet 2/2)# exit FS(config)# mpls static ilm in-label 16 forward-action swap-label 3 nexthop gigabitEthernet 2/2 192.168.2.2 fec 192.168.200.0/24 FS(config)# mpls static ilm in-label 17 forward-action swap-label 3 nexthop gigabitEthernet 2/1 192.168.1.1 fec 192.168.100.0/24 FS(config)# end
LER-C	FS(config)# mpls ip FS(config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# label-switching FS(config-if-GigabitEthernet 2/1)# exit FS(config)# mpls static ftn 192.168.100.0/24 out-label 17 nexthop gigabitEthernet 2/1 192.168.2.1 FS(config)# end

Verification	Check whether MPLS entries have been generated on all routers.
LER-A	A# show mpls forwarding-table Label Operation Code: PH--PUSH label PP--POP label SW--SWAP label SP--SWAP topmost label and push new label DP--DROP packet PC--POP label and continue lookup by IP or Label PI--POP label and do ip lookup forward PN--POP label and forward to nexthop PM--POP label and do MAC lookup forward PV--POP label and output to VC attach interface IP--IP lookup forward s--stale Local Outgoing OP FEC Outgoing Nexthop label label interface -- 16 PH 192.168.200.0/24 Gi2/2 192.168.1.2
LSR-B	B# show mpls forwarding-table Label Operation Code: PH--PUSH label PP--POP label SW--SWAP label SP--SWAP topmost label and push new label DP--DROP packet PC--POP label and continue lookup by IP or Label PI--POP label and do ip lookup forward PN--POP label and forward to nexthop

	PM--POP label and do MAC lookup forward PV--POP label and output to VC attach interface IP--IP lookup forward s--stale Local Outgoing OP FEC Outgoing Nexthop label label interface 16 imp-null PP 192.168.200.0/24 Gi2/2 192.168.2.2 17 imp-null PP 192.168.100.0/24 Gi2/1 192.168.1.1
LER-C	C# show mpls forwarding-table Label Operation Code: PH--PUSH label PP--POP label SW--SWAP label SP--SWAP topmost label and push new label DP--DROP packet PC--POP label and continue lookup by IP or Label PI--POP label and do ip lookup forward PN--POP label and forward to nexthop PM--POP label and do MAC lookup forward PV--POP label and output to VC attach interface IP--IP lookup forward s--stale Local Outgoing OP FEC Outgoing Nexthop label label interface -- 17 PH 192.168.100.0/24 Gi2/1 192.168.2.1

Common Errors

- MPLS forwarding is not enabled globally.
- MPLS forwarding is not enabled in interface configuration mode.

1.4.3 Configuring an LDP Session

Configuration Effect

- Establish an LDP session between two LSRs to exchange label binding messages.

Notes

- A reachable IP route to the transport address used by the peer to establish an LDP session must exist.

Configuration

Configuring IP Routes

- (Mandatory) Configure static routes in global configuration mode or configure routes through OSPF or other IGP protocols.
- For details, see the *Configuring IP Routes*.

Configuring LDP

- (Mandatory) Perform this configuration on the two LSRs engaged in an LDP session.

Configuring an LDP Router ID

- (Optional) Perform this configuration in config-mpls-router configuration mode.
- It is recommended that an LDP Router ID be configured on each of the two LSRs engaged in an LDP session to ensure stable and reliable LDP.

You must configure the LDP Router ID in actual application; otherwise, if the Router ID is changed during protocol restart, some functions (such as LDP MD5 authentication using the Router ID) may fail.

Configuring the LDP Capability on an Interface or Configuring the Remote Peer

- Enable the LDP capability in interface configuration mode if a local LDP session needs to be established.
- Configure the remote peer in config-mpls-router configuration mode if a remote LDP session needs to be established.

Configuring a Transport Address

- (Optional) If no transport address is configured, the LDP Router ID is used as the transport address.
- Configure a global transport address in config-mpls-router configuration mode if a local LDP session needs to be established. If a neighbor requires a specified transport address, configure this address on the interface connected to the neighbor.
- This configuration is not required if a remote LDP session needs to be established.

Configuring the Hello Packet Interval

- This configuration is required only in special cases.
- Perform this configuration in interface configuration mode if a local LDP session needs to be established.
- Perform this configuration in config-mpls-router configuration mode if a remote LDP session needs to be established.

Configuring the Hello Packet Keep-Alive Period

- This configuration is required only in special cases.
- Perform this configuration in interface configuration mode if a local LDP session needs to be established.

- Perform this configuration in config-mpls-router configuration mode if a remote LDP session needs to be established.

Configuring the Keepalive Packet Keep-Alive Period

- This configuration is required only in special cases.
- Perform this configuration in interface configuration mode if a local LDP session needs to be established.
- Perform this configuration in config-mpls-router configuration mode if a remote LDP session needs to be established.

Configuring the Maximum PDU Size

- This configuration is required only in special cases.
- Perform this configuration in interface configuration mode.

Configuring LDP MD5 Authentication

- You can configure LDP MD5 authentication on the two LSRs engaged in an LDP session to improve session security.
- (Optional) Perform this configuration in config-mpls-router configuration mode.

Configuring the Waiting Time for Label Mapping During Route Updating

- This configuration is required only in special cases.
- (Optional) It is performed in global configuration mode.

Verification

- Run the **show mpls ldp discovery** command to display the LDP peer discovery state.
- Run the **show mpls ldp neighbor** command to display the LDP session state.
- Run the **show mpls ldp bindings** command to display the LDP label binding state.
- Run the **show run** command to check whether the configuration takes effect.

Related Commands

Enabling LDP Globally

Command	**mpls router ldp** [*vrf-name*]
Parameter Description	*vrf-name:* Enables or disables LDP of a specified VRF instance.
Command Mode	Global configuration mode
Usage Guide	The number of LDP instances is limited by the number of VRF instances. Each VRF instance can enable one LDP instance. By default, LDP instances of global VRF instances are enabled (or disabled).

Configuring an LDP Router ID

Command	**ldp router-id** { *ip-address* **|** **interface** *interface-name* [**force**] }
Parameter	*ip-address:* Specifies a static IP address as the LDP router ID. The ID takes effect immediately.

Description	*interface-name* [**force**]: Specifies the primary address of a specified interface as the LDP router ID. If **force** is specified, the ID must be forced to take effect immediately.
Configuration Mode	config-mpls-router mode
Usage Guide	Configures an LDP router ID.

Enabling LDP on an Interface

Command	**mpls ip**
Parameter Description	**N/A**
Configuration Mode	Interface configuration mode
Usage Guide	Enables LDP on an interface.

Configuring a Remote LDP Peer

Command	**neighbor** *ip-address*
Parameter Description	*ip-address:* Indicates the ID of the peer LSR.
Configuration Mode	config-mpls-router mode
Usage Guide	Creates a remote LDP peer.

Receiving Target Hello Packets Through LDP

Command	**discovery targeted-hello accept** [**from** *acl-name*]
Parameter Description	*acl-name:* Indicates the name of ACL rules. Only the target Hello packets from neighbors specified by ACL rules can be received.
Configuration Mode	config-mpls-router mode
Usage Guide	When remote peers of two devices are established, you can configure a neighbor at one end and enable this function at the other end. To delete a remote peer, you only need to delete the neighbor configuration.

Configuring a Transport Address

Command	**mpls ldp transport-address** { **interface** | *ip-address* }
Parameter Description	**interface:** An LDP session uses the primary address of an interface. *ip-address:* An LDP session uses the specified IP address.
Configuration Mode	Interface configuration mode
Usage Guide	This command is valid for local LDP sessions established by the basic LDP discovery mechanism and invalid for remote LDP sessions established by the extended LDP discovery mechanism. It is valid for new local LDP sessions and invalid for existing local LDP sessions.

Command	**transport-address** { **interface** | *ip-address* | *interface-name* }
Parameter Description	**interface:** A local LDP session uses the primary IP address of an interface as a transport address. *ip-address:* All local LDP sessions use a specified IP address as the transport address. *interface-name:* All local LDP sessions use the primary IP address of a specified interface as the transport address.
Configuration Mode	config-mpls-router mode
Usage Guide	This command is valid for local LDP sessions established by the basic LDP discovery mechanism and invalid for remote LDP sessions established by the extended LDP discovery mechanism. Remote LDP sessions always use the LDP LSR ID as the transport address. If transport addresses are configured in both global and interface configuration modes, the transport address configured in interface configuration mode is preferred.

Configuring the Hello Packet Interval

Command	**mpls ldp hello-interval** *seconds*
Parameter Description	*seconds:* Indicates the Hello packet interval, which ranges from 1 to 65,535 seconds.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the LDP link Hello packet interval.

Command	**discovery targeted-hello** *interval seconds*
Parameter Description	*seconds:* Indicates the Hello packet interval, which ranges from 1 to 65,535 seconds.
Configuration Mode	config-mpls-router mode
Usage Guide	Configures the LDP target Hello packet interval.

Configuring the Hello Packet Keep-Alive Period

Command	**mpls ldp hello-holdtime** *seconds*
Parameter Description	*seconds:* Indicates the Hello packet keep-alive period, which ranges from 1 to 65,535 seconds. 65,535 seconds indicate that a timeout will never occur.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the link Hello packet keep-alive period.

Command	**discovery targeted-hello** *holdtime seconds*
Parameter Description	Indicates the target Hello packet keep-alive period, which ranges from 1 to 65,535 seconds. 65,535 seconds indicate infinite hold time.
Configuration	config-mpls-router mode

Mode	
Usage Guide	Configures the target Hello packet keep-alive period.

Configuring the Keepalive Packet Keep-Alive Period

Command	**mpls ldp keepalive-holdtime** *seconds*
Parameter Description	*seconds:* Indicates the keep-alive packet keep-alive period, which ranges from 15 to 65,535 seconds.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the keep-alive packet keep-alive period for local LDP sessions in interface configuration mode.

Command	**targeted-session holdtime** *seconds*
Parameter Description	*seconds:* Indicates the keep-alive packet keep-alive period, which ranges from 15 to 65,535 seconds.
Configuration Mode	config-mpls-router mode
Usage Guide	The default keep-alive packet keep-alive period is used for restoration of local LDP sessions.

Configuring the Maximum PDU Size

Command	**mpls ldp max-pdu** *max-pdu*
Parameter Description	*max-pdu:* Indicates the maximum PDU length for message exchange, which ranges from 256 to 4,096 bytes.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the maximum PDU length.

Configuring LDP MD5 Authentication

Command	**neighbor** *ip-address* **password** [**0 | 7**] *pwd-string*
Parameter Description	*ip-address:* Indicates the transport address of the peer LSR. [**0 | 7**]: (Optional) **0** indicates a plaintext key; **7** indicates a cyphertext key. By default, a plaintext key is entered. *pwd-string:* Indicates a password string, which is case-sensitive. A plaintext password string contains 1 to 25 characters, while a cyphertext password string contains 1 to 52 characters.
Configuration Mode	config-mpls-router mode
Usage Guide	Configures MD5 authentication of TCP connections to the peer LSR.

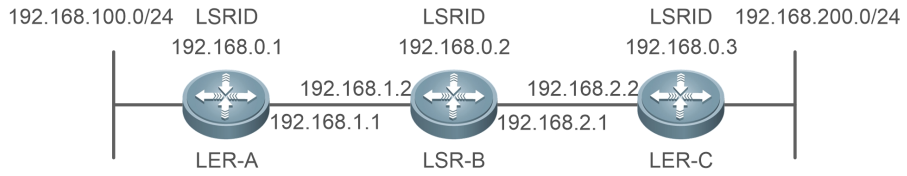
Configuring the Waiting Time for Label Mapping

Command	**route wait label-mapping** *seconds*
Parameter Description	*seconds*: Configures the waiting time for label mapping during route updating.

Configuration Mode	Global configuration mode
Usage Guide	During IP route updating, an LSP corresponding to the new next hop perhaps is not yet established. However, LDP will delete the LSP corresponding to the old next hop so that the traffic will be interrupted until a new LSP is established. To solve this problem, you can run this command to have LDP delay deleting the old LSP until a new LSP is established. During this period, once a new LSP is established, the traffic will be switched to the new LSP. If a new LSP is not established within this period, the LDP will delete the old LSP.

Configuration Example

Establishing a Local LDP Session

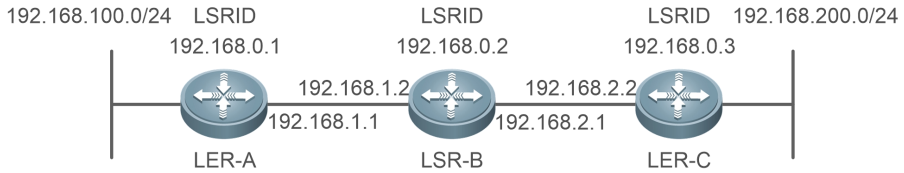
Scenario **Figure 1-10**	
Configuration Steps	● Configure an IPv4 unicast routing protocol (such as OSPF) on each router and check that a unicast route is reachable to its loopback interface. ● Enable LDP globally. ● Configure an LDP Router ID. ● Enable LDP on an interface.
LER-A	FS(config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit FS(config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# mpls ip FS(config-if-GigabitEthernet 2/2)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.100.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.1 0.0.0.0 area 0 FS (config-router)# network 192.168.1.0 0.0.0.255 area 0 FS(config-router)# end
LSR-B	FS (config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit

	FS (config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# mpls ip FS(config-if-GigabitEthernet 2/1)# exit FS (config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# mpls ip FS(config-if-GigabitEthernet 2/2)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.1.0 0.0.0.255 area 0 FS (config-router)# network 192.168.2.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.2 0.0.0.0 area 0 FS (config-router)# end
LER-C	FS (config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit FS (config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# mpls ip FS(config-if-GigabitEthernet 2/1)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.200.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.3 0.0.0.0 area 0 FS (config-router)# network 192.168.2.0 0.0.0.255 area 0 FS (config-router)# end
Verification	
LER-A	A# show mpls ldp neighbor Default VRF: Peer LDP Ident: 192.168.0.2:0; Local LDP Ident: 192.168.0.1:0 TCP connection: 192.168.0.2.54184 - 192.168.0.1.646 State: OPERATIONAL; Msgs sent/recv: 18/19; UNSOLICITED Up time: 00:02:49 Graceful Restart enabled; Peer reconnect time (msecs): 300000

	LDP discovery sources: Link Peer on GigabitEthernet 2/2, Src IP addr: 192.168.1.2 Addresses bound to peer LDP Ident: 192.168.2.1 192.168.1.2 192.168.0.2
LSR-B	B# show mpls ldp neighbor Default VRF: Peer LDP Ident: 192.168.0.3:0; Local LDP Ident: 192.168.0.2:0 TCP connection: 192.168.0.3.1101 - 192.168.0.2.646 State: OPERATIONAL; Msgs sent/recv: 20/21; UNSOLICITED Up time: 00:03:24 Graceful Restart enabled; Peer reconnect time (msecs): 300000 LDP discovery sources: Link Peer on GigabitEthernet 2/2, Src IP addr: 192.168.2.2 Addresses bound to peer LDP Ident: 192.168.2.2 192.168.0.3 Peer LDP Ident: 192.168.0.1:0; Local LDP Ident: 192.168.0.2:0 TCP connection: 192.168.0.1.646 - 192.168.0.2.54184 State: OPERATIONAL; Msgs sent/recv: 7/8; UNSOLICITED Up time: 00:00:04 Graceful Restart enabled; Peer reconnect time (msecs): 300000 LDP discovery sources: Link Peer on GigabitEthernet 2/1, Src IP addr: 192.168.1.1 Addresses bound to peer LDP Ident: 192.168.1.1 192.168.0.1
LER-C	C# show mpls ldp neighbor Default VRF: Peer LDP Ident: 192.168.0.2:0; Local LDP Ident: 192.168.0.3:0 TCP connection: 192.168.0.2.646 - 192.168.0.3.1101 State: OPERATIONAL; Msgs sent/recv: 35/36; UNSOLICITED Up time: 00:07:09 Graceful Restart enabled; Peer reconnect time (msecs): 300000

	LDP discovery sources: Link Peer on GigabitEthernet 2/1, Src IP addr: 192.168.2.1 Addresses bound to peer LDP Ident: 192.168.2.1 192.168.1.2 192.168.0.2

Establishing a Remote LDP Session

Scenario **Figure 1-11**	
Configuration Steps	● Configure an IPv4 unicast routing protocol (such as OSPF) on each router and check that a unicast route is reachable to its loopback interface. ● Enable LDP globally. ● Configure an LDP Router ID. ● Configure the remote peer.
LER-A	FS(config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS(config-mpls-router)# neighbor 192.168.0.3 FS(config-mpls-router)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.100.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.1 0.0.0.0 area 0 FS (config-router)# network 192.168.1.0 0.0.0.255 area 0 FS(config-router)# end
LSR-B	FS (config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.1.0 0.0.0.255 area 0 FS (config-router)# network 192.168.2.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.2 0.0.0.0 area 0 FS (config-router)# end

LER-C	FS (config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS(config-mpls-router)# neighbor 192.168.0.1 FS (config-mpls-router)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.200.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.3 0.0.0.0 area 0 FS (config-router)# network 192.168.2.0 0.0.0.255 area 0 FS (config-router)# end
Verification	Check whether an LDP session is established between LER A and LER C.
LER-A	A# show mpls ldp neighbor Default VRF: Peer LDP Ident: 192.168.0.3:0; Local LDP Ident: 192.168.0.1:0 TCP connection: 192.168.0.3.1102 - 192.168.0.1.646 State: OPERATIONAL; Msgs sent/recv: 7/8; UNSOLICITED Up time: 00:00:55 Graceful Restart enabled; Peer reconnect time (msecs): 300000 LDP discovery sources: Targeted Hello 192.168.0.1 -> 192.168.0.3, active, passive; Addresses bound to peer LDP Ident: 192.168.2.2 192.168.0.3
LER-C	C# show mpls ldp neighbor Default VRF: Peer LDP Ident: 192.168.0.1:0; Local LDP Ident: 192.168.0.3:0 TCP connection: 192.168.0.1.646 - 192.168.0.3.1102 State: OPERATIONAL; Msgs sent/recv: 7/8; UNSOLICITED Up time: 00:00:09 Graceful Restart enabled; Peer reconnect time (msecs): 300000 LDP discovery sources: Targeted Hello 192.168.0.3 -> 192.168.0.1, active, passive;

	Addresses bound to peer LDP Ident: 192.168.1.1 192.168.0.1

Common Errors

- IPv4 unicast routing is incorrectly configured.
- LDP is not enabled globally.
- LDP is not enabled interface configuration mode (or a remote peer is not configured).

1.4.4 Configuring LDP Session Protection

Configuration Effect

Protect LDP sessions between two directly connected LSRs.

Notes

You can enable LDP session protection on the two LSRs engaged in an LDP session. If only one LSR supports LDP session protection, configure LDP session protection on this LSR and run the **discovery targeted-hello accept** command on the other LSR to enable it to receive and reply to target Hello packets (or run the **neighbor ip-address** command to specify the remote peer).

Configuration

Configuring an LDP Session

- (Mandatory) Perform this configuration according to section 1.4.3 "Configuring an LDP Session."

Configuring LDP Session Protection

- (Mandatory) Perform this configuration in config-mpls-router configuration mode.
- Enable LDP session protection on the two LSRs engaged in an LDP session (see "Notes"), unless otherwise specified.

Verification

- Run the **show mpls ldp parameters** command to check whether LDP session protection is enabled.
- Run the **show mpls ldp neighbor** command to check whether LDP session protection takes effect.

Related Commands

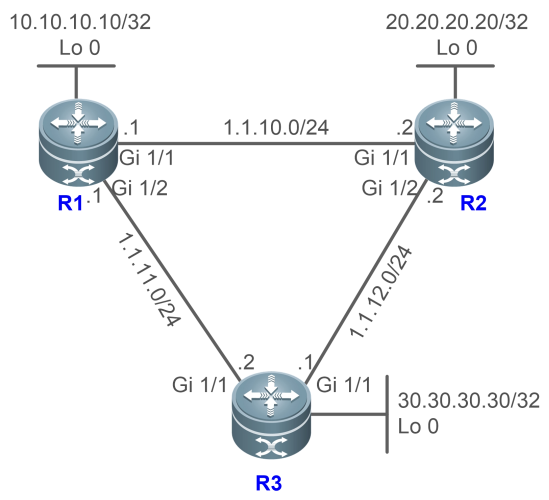
Configuring LDP Session Protection

Command	**session protection** [**for acl** *acl_name*] [**duration** { **infinite** | *seconds* }]
Parameter Description	**for acl** *acl_name:* Configures ACL rules. Only the LDP sessions of LDP neighbors whose LSR IDs meet ACL rules are protected. **duration:** Configures the duration of session protection. When links fail and cause a Hello adjacency without associated interfaces, session protection will be enabled, and a target Hello adjacency will maintain LDP sessions. **infinite:** The target Hello adjacency will maintain LDP sessions infinitely.

	seconds: Indicates the duration of session protection, which ranges from 30 to 2,147,483 seconds.
Configuration Mode	config-mpls-router mode
Usage Guide	Enables LDP protection.

Configuration Example

Enabling LDP Session Protection Between LSRs

Scenario Figure 1-12	
Configuration Steps	● On all nodes, configure the interface IP address, configure OSPF, and enable LDP. ● Enable LDP session protection.
R1	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 10.10.10.10 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.10.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2

	FS(config-if-GigabitEthernet 1/2)# ip address 1.1.11.1 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.10.0 0.0.0.255 area 0 FS(config-router)# network 1.1.11.0 0.0.0.255 area 0 FS(config-router)# network 10.10.10.10 0.0.0.0 area 0 FS(config-router)# exit FS# configure terminal FS(config)# mpls router ldp FS(config-router)# session protection FS(config-router)# session protection duration 300 FS(config-router)# exit
R2	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 20.20.20.20 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.10.2 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.12.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit

	FS(config)# router ospf 1 FS(config-router)# network 1.1.10.0 0.0.0.255 area 0 FS(config-router)# network 1.1.12.0 0.0.0.255 area 0 FS(config-router)# network 20.20.20.20 0.0.0.0 area 0 FS(config-router)# exit FS# configure terminal FS(config)# mpls router ldp FS(config-router)# session protection FS(config-router)# session protection duration 300 FS(config-router)# exit
R3	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 30.30.30.30 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.11.2 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.12.1 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.11.0 0.0.0.255 area 0 FS(config-router)# network 1.1.12.0 0.0.0.255 area 0 FS(config-router)# network 30.30.30.30 0.0.0.0 area 0

	FS(config-router)# exit FS# configure terminal FS(config)# mpls router ldp FS(config-router)# session protection FS(config-router)# session protection duration 300 FS(config-router)# exit
Verification	● Run the **show mpls ldp parameters** command to check whether the inter-area LDP LSP function is enabled. ● Run the **show mpls ldp neighbor detail** command to check whether an LDP entry is added to the route based on the longest match principle.
R1	FS# show mpls ldp parameters Default VRF: Protocol version: 1 Ldp Router ID: 8.8.8.8 Control Mode: INDEPENDENT Propogate Release: FALSE Label Merge: TRUE Label Retention Mode: LIBERAL Loop Detection Mode: off Inter Area Lsp: TRUE Session Protection: TRUE Targeted Session Keepalive HoldTime/Interval: 180/60 sec Targeted Hello HoldTime/Interval: 90/10 sec FS# show mpls ldp neighbor detail Default VRF: Peer LDP Ident: 20.20.20.20:0; Local LDP Ident: 10.10.10.10:0 TCP connection: 20.20.20.20.1030 - 10.10.10.10.646 State: OPERATIONAL; Msgs sent/recv: 6/7; UNSOLICITED Up time: 00:00:15 Graceful Restart disabled LDP discovery sources: Targeted Hello 10.10.10.10 -> 20.20.20.20, active, passive;

	Hold time: 45 sec, hello interval: 5 sec Addresses bound to peer LDP Ident: 1.1.10.2 1.1.20.2 20.20.20.20 Our is PASSIVE KA hold time: 45 sec; Proposed local/peer: 45/45 sec Peer distribute label mode: UNSOLICITED Peer loop detection: Disabled Peer Path Vector Limit: 0; Max PDU Length: 4096 LDP Session Protection enabled, state: Protecting duration: 300 seconds holdup time remaining: 60 seconds

1.4.5 Configuring an LDP LSP

Configuration Effect

Establish an LSP on IP unicast routes through LDP.

Notes

- LDP needs to use unicast routes existing in the network; therefore, IP unicast routes must be configured.

Configuration

Configuring MPLS Forwarding

- (Mandatory) Configure MPLS forwarding according to the section "Configuring Basic PIM-SM Functions 1".

Configuring LDP Sessions

- (Mandatory) Configure LDP sessions according to the section "Configuring LDP Sessions 1".

Configuring Loop Detection

- (Optional) Perform this configuration in config-mpls-router configuration mode.
- If IP routes are statically configured instead of being generated by IGP, you are advised to enable loop detection on every LSR traversed by the LSP to avoid loop caused by misconfiguration.

Configuring the Maximum Number of Hops

- Perform this configuration in interface configuration mode.
- This configuration is required only in special cases.

Configuring the Maximum Number of LSRs in the Path Vector List

- Perform this configuration in interface configuration mode.
- This configuration is required only in special cases.

Configuring the Label Release Message Propagation Mode

- Perform this configuration in config-mpls-router configuration mode.
- This configuration is required only in special cases.

Configuring the LDP Label Distribution Control Mode

- Perform this configuration in config-mpls-router configuration mode.
- This configuration is required only in special cases.

Configuring the LDP Label Distribution Mode

- Perform this configuration in interface configuration mode.
- This configuration is required only in special cases.

Configuring the LDP Label Retention Mode

- Perform this configuration in config-mpls-router configuration mode.
- This configuration is required only in special cases.

Configuring Label Distribution Policies

- (Optional) Perform this configuration in config-mpls-router configuration mode.

By default, LDP distributes labels only to IGP routes but not to BGP routes. To distribute labels to BGP routes, run the **advertise-labels for bgp-routes** command.

By default, LDP distributes implicit null labels with the value of 3 to default routes; therefore, an LSP is not established on default routes.

Configuring Label Reception Policies

- Perform this configuration in config-mpls-router configuration mode.
- This configuration is required only in special cases.

Configuring Explicit Null Label Distribution Policies

- Perform this configuration in config-mpls-router configuration mode.
- This configuration is required only in special cases.

Explicit null label distribution policies can be configured only for global LDP instances but not for VRF instances.

Verification

- Run the **show mpls ldp discovery** command to display the LDP peer discovery state.
- Run the **show mpls ldp neighbor** command to display the LDP session state.

- Run the **show mpls ldp bindings** command to display the LDP label binding state.
- Run the **show mpls forwarding-table** command to display the LSP entry state.

Related Commands

Configuring Loop Detection

Command	**loop-detection**
Parameter Description	N/A
Configuration Mode	config-mpls-router mode
Usage Guide	Enables loop detection.

Configuring the Maximum Hop Count

Command	**mpls ldp max-hop-count** *number*
Parameter Description	*number:* Indicates the maximum hop count allowed for loop detection. The value ranges from 1 to 255.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the maximum hop count for loop detection.

Configuring the Maximum Number of LSRs in the Path Vector List

Command	**mpls ldp max-path-vector** *number*
Parameter Description	*number:* Indicates the maximum number of of LSRs in the path vector list. The value ranges from 0 to 255.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the maximum number of LSRs in the path vector list.

Propagating Label Release Messages

Command	**propagate-release**
Parameter Description	N/A
Configuration Mode	config-mpls-router mode
Usage Guide	Propagates label release messages downstream.

Configuring the LDP Label Distribution Control Mode

Command	**lsp-control-mode** { **independent | orderd** }
Parameter Description	**independent:** Indicates the independent control mode. **ordered:** Indicates the ordered control mode.

Configuration Mode	The independent control mode is used.
Usage Guide	Configures the label distribution control mode.

Configuring the LDP Label Distribution Mode

Command	**mpls ldp distribution-mode** { **du** | **dod** }
Parameter Description	**dod:** Indicates the DOD distribution mode. **du:** Indicates the DU distribution mode.
Configuration Mode	Interface configuration mode
Usage Guide	Configures the label distribution mode.

Configuring the LDP Label Retention Mode

Command	**label-retention-mode** { **liberal** | **conservation** }
Parameter Description	**liberal:** Indicates the liberal label retention mode. **conservative:** Indicates the conservative label retention mode.
Configuration Mode	config-mpls-router mode
Usage Guide	Configures the label retention mode.

Configuring Label Distribution Policies

Command	**advertise-labels** [**for host-routes** | **for bgp-routes** [**acl** *acl_name*] | **for default-route** | **for acl** *prefix-access-list* [**to** *peer-access-list*]
Parameter Description	**for host-routes:** Distributes labels to host routes (with a 32-bit mask) only. **for bgp-routes** [**acl** *acl_ame*]: Distributes labels to BGP routes. You can use the **acl** keyword to have labels distributed to eligible BGP routes only. **for default-route:** Distributes labels whose value is not 3 to default routes. **for acl** *prefix-access-list:* Specifies the route prefix. **to** *peer-access-list:* Specifies the neighbors to which label binding messages should be sent.
Configuration Mode	config-mpls-router mode
Usage Guide	Configures label distribution policies for FECs mapped to IP routes. You can use the **no** form of this command to recover the default.

Configuring Label Reception Policies

Command	**neighbor** *ip-address* **labels accept** *acl-name*
Parameter Description	*ip-address:* Indicates the ID of the peer LSR. *acl-name:* Specifies the name of ACL rules.
Configuration Mode	config-mpls-router mode

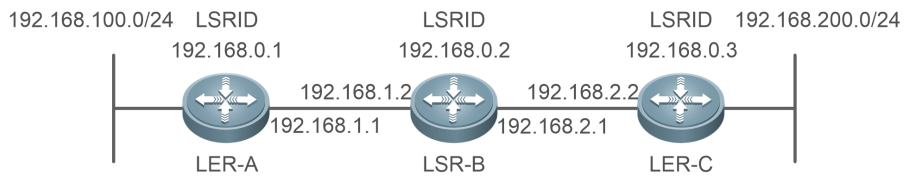
Usage Guide	Configures label reception policies.

Configuring Explicit Null Label Distribution Policies

Command	**explicit-null** [**for** *prefix_acl*] [**to** *peer_acl*]
Parameter Description	**for** *prefix-acl:* Specifies the address prefix of the direct route for replacing implicit null labels with explicit null labels. **to** *peer-acl:* Specifies the LDP peer for replacing implicit null labels with explicit null labels.
Configuration Mode	config-mpls-router mode
Usage Guide	Distributes explicit null labels to all direct routes or to routes meeting ACL rules.

Configuration Example

Establishing an LDP LSP

Scenario **Figure 1-13**	
Configuration Steps	● Enable MPLS forwarding globally. ● Enable LDP globally. ● Enable LDP on an interface. ● Configure an LDP Router ID. ● Enable label forwarding on an interface. ● Configure an IPv4 unicast routing protocol (such as OSPF) on each router and check that a unicast route is reachable to its loopback interface.
LER-A	FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit FS(config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# mpls ip FS(config-if-GigabitEthernet 2/2)# label-switching FS(config-if-GigabitEthernet 2/2)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.100.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.1 0.0.0.0 area 0

	FS (config-router)# network 192.168.1.0 0.0.0.255 area 0 FS(config-router)# end
LSR-B	FS (config)# mpls ip FS (config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit FS (config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# mpls ip FS(config-if-GigabitEthernet 2/1)# label-switching FS(config-if-GigabitEthernet 2/1)# exit FS (config)# interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)# mpls ip FS(config-if-GigabitEthernet 2/2)# label-switching FS(config-if-GigabitEthernet 2/2)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.1.0 0.0.0.255 area 0 FS (config-router)# network 192.168.2.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.2 0.0.0.0 area 0 FS (config-router)# end
LER-C	FS (config)# mpls ip FS (config)# mpls router ldp FS(config-mpls-router)# ldp router-id interface loopback 0 force FS (config-mpls-router)# exit FS (config)# interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)# mpls ip FS(config-if-GigabitEthernet 2/1)# label-switching FS(config-if-GigabitEthernet 2/1)# exit FS (config)# router ospf 10 FS (config-router)# network 192.168.200.0 0.0.0.255 area 0 FS (config-router)# network 192.168.0.3 0.0.0.0 area 0 FS (config-router)# network 192.168.2.0 0.0.0.255 area 0

	FS (config-router)# end
Verification	Check whether LSP entries are correctly created on all routers.
LER-A	A# show mpls forwarding-table Label Operation Code: PH--PUSH label PP--POP label SW--SWAP label SP--SWAP topmost label and push new label DP--DROP packet PC--POP label and continue lookup by IP or Label PI--POP label and do ip lookup forward PN--POP label and forward to nexthop PM--POP label and do MAC lookup forward PV--POP label and output to VC attach interface IP--IP lookup forward Local Outgoing OP FEC Outgoing Nexthop label label interface -- imp-null PH 192.168.0.2/32 Gi2/2 192.168.1.2 -- 1024 PH 192.168.0.3/32 Gi2/2 192.168.1.2 -- imp-null PH 192.168.2.0/24 Gi2/2 192.168.1.2 1024 imp-null PP 192.168.0.2/32 Gi2/2 192.168.1.2 1025 1024 SW 192.168.0.3/32 Gi2/2 192.168.1.2 1026 imp-null PP 192.168.2.0/24 Gi2/2 192.168.1.2
LSR-B	B# show mpls forwarding-table Label Operation Code: PH--PUSH label PP--POP label SW--SWAP label SP--SWAP topmost label and push new label DP--DROP packet

	PC--POP label and continue lookup by IP or Label PI--POP label and do ip lookup forward PN--POP label and forward to nexthop PM--POP label and do MAC lookup forward PV--POP label and output to VC attach interface IP--IP lookup forward s--stale Local Outgoing OP FEC Outgoing Nexthop label label interface -- imp-null PH 192.168.0.1/32 Gi2/1 192.168.1.1 -- imp-null PH 192.168.0.3/32 Gi2/2 192.168.2.2 1024 imp-null PP 192.168.0.3/32 Gi2/2 192.168.2.2 1025 imp-null PP 192.168.0.1/32 Gi2/1 192.168.1.1
LER-C	C# show mpls forwarding-table Label Operation Code: PH--PUSH label PP--POP label SW--SWAP label SP--SWAP topmost label and push new label DP--DROP packet PC--POP label and continue lookup by IP or Label PI--POP label and do ip lookup forward PN--POP label and forward to nexthop PM--POP label and do MAC lookup forward PV--POP label and output to VC attach interface IP--IP lookup forward Local Outgoing OP FEC Outgoing Nexthop label label interface -- 1025 PH 192.168.0.1/32 Gi2/1 192.168.2.1 -- imp-null PH 192.168.0.2/32 Gi2/1 192.168.2.1 -- imp-null PH 192.168.1.0/24 Gi2/1 192.168.2.1

	1024	imp-null	PP 192.168.0.2/32	Gi2/1	192.168.2.1
	1025	imp-null	PP 192.168.1.0/24	Gi2/1	192.168.2.1
	1026	1025	SW 192.168.0.1/32	Gi2/1	192.168.2.1

Common Errors

- IPv4 unicast routing is incorrectly configured.
- MPLS forwarding is not enabled globally.
- LDP is not enabled globally.
- LDP is not enabled on an interface.
- Label forwarding is not enabled on an interface.

1.4.6 Configuring the Inter-Domain LDP LSP Feature

Configuration Effect

When an AS is divided into multiple IGP areas and host routes are aggregated on ABRs, establish an LSP across multiple IGP areas for the host routes.

Notes

- The inter-domain LDP LSP feature takes effect only when the free label retention mode is configured.
- The inter-domain LDP LSP feature is not applicable in DOD label distribution mode.

Configuration Steps

Configuring IGP

- (Mandatory) Perform this configuration according to the *Configuring OSPF* or *Configuring IS-IS*.

Enabling Area Border Routers (ABRs) to Aggregate Host Routes

- (Mandatory) Perform this configuration according to the *Configuring OSPF* or *Configuring IS-IS*.

Configuring MPLS Forwarding

- (Mandatory) Configure MPLS forwarding according to the section "Configuring Basic PIM-SM Functions 1".

Configuring LDP Sessions

- (Mandatory) Configure LDP sessions according to the section "Configuring LDP Sessions 1".

Configuring the Inter-Domain LDP LSP Feature

- (Mandatory) Perform this configuration in config-mpls-router configuration mode.
- After route aggregation, enable the inter-domain LDP LSP feature on every LSR traversed by the corresponding LSP.

Verification

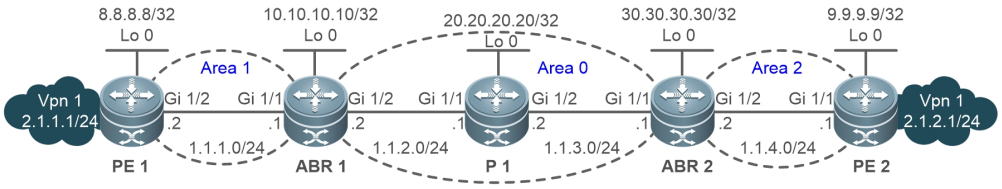
- Display configurations of LDP instances to check whether the inter-area LDP LSP function is enabled.
- Check whether an MPLS entry is added to the route compliant with the longest match rule on every router.

Related Commands

Command	**inter-area-lsp** [**for acl** *acl_name*]
Parameter Description	**for acl** *acl_name*: Configures ACL rules and establishes inter-area LSPs for the aggregated host routes (with a 32-bit mask) that meet ACL rules only.
Configuration Mode	config-mpls-router mode
Usage Guide	Enables the inter-area LDP LSP function on all devices between PEs.

Configuration Example

Establishing an LSP (Across Different Areas) Between PE1 and PE2, and Enabling the Inter-area LDP LSP Function (OSPF Network)

Scenario **Figure 1-14**	
Configuration Steps	- On all nodes, configure the interface IP address, configure OSPF, and enable LDP. - On PEs, configure VPN routing instances and MP-IBGP neighbors, and propagate VPN routing information. - Announce aggregated routes on ABRs. - Enable the inter-area LDP LSP function on all nodes.
PE1	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 8.8.8.8 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.1.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip

	FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.1.0 0.0.0.255 area 1 FS(config-router)# network 8.8.8.8 0.0.0.0 area 1 FS(config-router)# exit FS(config)# ip vrf vpn1 FS(config-vrf)# rd 65001:20 FS(config-vrf)# route-target both 65001:20 FS(config-vrf)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip vrf forwarding vpn1 FS(config-if-GigabitEthernet 1/1)# ip address 2.1.1.1 255.255.255.0 FS(config)# router bgp 65001 FS(config-router)# neighbor 9.9.9.9 remote-as 65001 FS(config-router)# neighbor 9.9.9.9 update-source Loopback 0 FS(config-router)# address-family vpnv4 FS(config-router-af)# neighbor 9.9.9.9 activate FS(config-router-af)# neighbor 9.9.9.9 send-community both FS(config-router-af)# exit FS(config-router)# address-family ipv4 vrf vpn1 FS(config-router-af)# redistribute connected FS(config-router-af)# exit FS(config-router)# exit FS(config)# mpls router ldp FS(config-router)# inter-area-lsp FS(config-router)# exit
ABR1	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 10.10.10.10 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip

	FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.1.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.2.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.1.0 0.0.0.255 area 1 FS(config-router)# network 1.1.2.0 0.0.0.255 area 0 FS(config-router)# network 10.10.10.10 0.0.0.0 area 0 FS(config-router)# exit FS(config)# router ospf 1 FS(config-router)# area 1 range 8.0.0.0 255.0.0.0 FS(config-router)# exit
P1	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 20.20.20.20 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.2.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip

	FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.3.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.2.0 0.0.0.255 area 0 FS(config-router)# network 1.1.3.0 0.0.0.255 area 0 FS(config-router)# network 20.20.20.20 0.0.0.0 area 0 FS(config-router)# exit
PE2	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 9.9.9.9 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.4.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.4.0 0.0.0.255 area 2 FS(config-router)# network 9.9.9.9 0.0.0.0 area 2 FS(config-router)# exit FS(config)# ip vrf vpn1 FS(config-vrf)# rd 65001:20 FS(config-vrf)# route-target both 65001:20

	FS(config-vrf)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip vrf forwarding vpn1 FS(config-if-GigabitEthernet 1/2)# ip address 2.1.2.1 255.255.255.0 FS(config)# router bgp 65001 FS(config-router)# neighbor 8.8.8.8 remote-as 65001 FS(config-router)# neighbor 8.8.8.8 update-source Loopback 0 FS(config-router)# address-family vpnv4 FS(config-router-af)# neighbor 8.8.8.8 activate FS(config-router-af)# neighbor 8.8.8.8 send-community both FS(config-router-af)# exit FS(config-router)# address-family ipv4 vrf vpn1 FS(config-router-af)# redistribute connected FS(config-router-af)# exit FS(config-router)# exit FS(config)# mpls router ldp FS(config-router)# inter-area-lsp FS(config-router)# exit
ABR2	FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 30.30.30.30 255.255.255.255 FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.3.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2

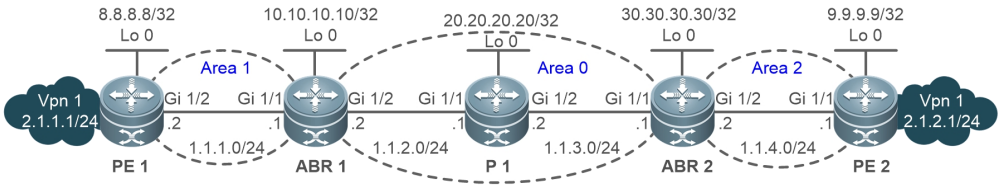
	FS(config-if-GigabitEthernet 1/2)# ip address 1.1.4.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router ospf 1 FS(config-router)# network 1.1.3.0 0.0.0.255 area 0 FS(config-router)# network 1.1.4.0 0.0.0.255 area 2 FS(config-router)# network 30.30.30.30 0.0.0.0 area 2 FS(config-router)# exit FS(config)# router ospf 1 FS(config-router)# area 2 range 9.0.0.0 255.0.0.0 FS(config-router)# exit
Verification	● Run the **show mpls ldp parameters** command to check whether the inter-area LDP LSP function is enabled. ● Run the **show mpls ldp bindings** command to check whether an LDP entry is added to the route based on the longest match principle.
PE1	FS# show mpls ldp parameters Default VRF: Protocol version: 1 Ldp Router ID: 8.8.8.8 Control Mode: INDEPENDENT Propogate Release: FALSE Label Merge: TRUE Label Retention Mode: LIBERAL Loop Detection Mode: off Inter Area Lsp: TRUE Session Protection: FALSE Targeted Session Keepalive HoldTime/Interval: 180/60 sec Targeted Hello HoldTime/Interval: 90/10 sec FS# show mpls ldp bindings Default VRF:

```
lib entry: 10.10.10.10/32
    local binding:   to lsr:    10.10.10.10:0, label: 1027
    local binding:   to lsr:    30.30.30.30:0, label: 1027
    remote binding: from lsr: 10.10.10.10:0, label: imp-null
    remote binding: from lsr: 30.30.30.30:0, label: 1030 (not in FIB)
lib entry: 9.9.9.9/32, (IA)
    local binding:   to lsr:    10.10.10.10:0, label: 1028
    local binding:   to lsr:    30.30.30.30:0, label: 1028
    remote binding: from lsr: 10.20.10.10:0, label: 1027 (not in FIB)
    remote binding: from lsr: 30.30.30.30:0, label: 1031
```

Establishing an LSP (Across Different Areas) Between PE1 and PE2, and Enabling the Inter-area LDP LSP Function (IS-IS Network)

Scenario

Figure 1-15



Configuration Steps

- On all nodes, configure the interface IP address and IS-IS.
- On PEs, configure VPN routing instances and MP-IBGP neighbors, and propagate VPN routing information.
- Announce aggregated routes on ABRs.
- Enable the inter-area LDP LSP function on all nodes.

P1

```
FS(config)# router isis
FS(config-router)# net 50.0001.0000.0000.0001.00
FS(config-router)# is-type level-2-only
FS(config-router)# exit
FS(config)# interface Loopback 0
FS(config-if-Loopback 0)# ip address 20.20.20.20 255.255.255.255
FS(config-if-Loopback 0)# ip router isis
FS(config-if-Loopback 0)# exit
FS(config)# mpls ip
FS(config)# mpls router ldp
```

	FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.2.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# ip router isis FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.3.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# ip router isis FS(config-if-GigabitEthernet 1/2)# exit
ABR1	FS(config)# router isis FS(config-router)# net 49.0001.0000.0000.0002.00 FS(config-router)# is-type level-1-2 FS(config-router)# exit FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 10.10.10.10 255.255.255.255 FS(config-if-Loopback 0)# ip router isis FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip address 1.1.1.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# ip router isis

	FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.2.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# ip router isis FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router isis FS(config-router)# summary-address 8.0.0.0 255.0.0.0 FS(config-router)# exit
PE1	FS(config)# router isis FS(config-router)# net 49.0001.0000.0000.0003.00 FS(config-router)# is-type level-1 FS(config-router)# exit FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 8.8.8.8 255.255.255.255 FS(config-if-Loopback 0)# ip router isis FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.1.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# ip router isis FS(config-if-GigabitEthernet 1/2)# exit FS(config)# ip vrf vpn1 FS(config-vrf)# rd 65001:20 FS(config-vrf)# route-target both 65001:20

	FS(config-vrf)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip vrf forwarding vpn1 FS(config-if-GigabitEthernet 1/1)# ip address 2.1.1.1 255.255.255.0 FS(config)# router bgp 65001 FS(config-router)# neighbor 9.9.9.9 remote-as 65001 FS(config-router)# neighbor 9.9.9.9 update-source Loopback 0 FS(config-router)# address-family vpnv4 FS(config-router-af)# neighbor 9.9.9.9 activate FS(config-router-af)# neighbor 9.9.9.9 send-community both FS(config-router-af)# exit FS(config-router)# address-family ipv4 vrf vpn1 FS(config-router-af)# redistribute connected FS(config-router-af)# exit FS(config-router)# exit FS(config)# mpls router ldp FS(config-router)# inter-area-lsp FS(config-router)# exit
ABR2	FS(config)# router isis FS(config-router)# net 49.0002.0000.0000.0002.00 FS(config-router)# is-type level-1-2 FS(config-router)# exit FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 30.30.30.30 255.255.255.255 FS(config-if-Loopback 0)# ip router isis FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1

	FS(config-if-GigabitEthernet 1/1)# ip address 1.1.3.1 255.255.255.0 FS(config-if-GigabitEthernet 1/1)# mpls ip FS(config-if-GigabitEthernet 1/1)# label-switching FS(config-if-GigabitEthernet 1/1)# ip router isis FS(config-if-GigabitEthernet 1/1)# exit FS(config)# interface gigabitEthernet 1/2 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.4.2 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# ip router isis FS(config-if-GigabitEthernet 1/2)# exit FS(config)# router isis FS(config-router)# summary-address 9.0.0.0 255.0.0.0 FS(config-router)# exit
PE2	FS(config)# router isis FS(config-router)# net 49.0002.0000.0000.0003.00 FS(config-router)# is-type level-1 FS(config-router)# exit FS(config)# interface Loopback 0 FS(config-if-Loopback 0)# ip address 9.9.9.9 255.255.255.255 FS(config-if-Loopback 0)# ip router isis FS(config-if-Loopback 0)# exit FS(config)# mpls ip FS(config)# mpls router ldp FS(config-mpls-router)# ldp rouer-id interface Loopback 0 force FS(config-mpls-router)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/2)# ip address 1.1.4.1 255.255.255.0 FS(config-if-GigabitEthernet 1/2)# mpls ip FS(config-if-GigabitEthernet 1/2)# label-switching FS(config-if-GigabitEthernet 1/2)# ip router isis

	FS(config-if-GigabitEthernet 1/2)# exit FS(config)# ip vrf vpn1 FS(config-vrf)# rd 65001:20 FS(config-vrf)# route-target both 65001:20 FS(config-vrf)# exit FS(config)# interface gigabitEthernet 1/1 FS(config-if-GigabitEthernet 1/1)# ip vrf forwarding vpn1 FS(config-if-GigabitEthernet 1/1)# ip address 2.1.2.1 255.255.255.0 FS(config)# router bgp 65001 FS(config-router)# neighbor 8.8.8.8 remote-as 65001 FS(config-router)# neighbor 8.8.8.8 update-source Loopback 0 FS(config-router)# address-family vpnv4 FS(config-router-af)# neighbor 8.8.8.8 activate FS(config-router-af)# neighbor 8.8.8.8 send-community both FS(config-router-af)# exit FS(config-router)# address-family ipv4 vrf vpn1 FS(config-router-af)# redistribute connected FS(config-router-af)# exit FS(config-router)# exit FS(config)# mpls router ldp FS(config-router)# inter-area-lsp FS(config-router)# exit
Verification	● Run the **show mpls ldp parameters** command to check whether the inter-area LDP LSP function is enabled. ● Run the **show mpls ldp bindings** command to check whether an LDP entry is added to the route based on the longest match principle.
P1	FS# show mpls ldp parameters Default VRF: Protocol version: 1 Ldp Router ID: 8.8.8.8 Control Mode: INDEPENDENT Propogate Release: FALSE

```
    Label Merge: TRUE
    Label Retention Mode: LIBERAL
Loop Detection Mode: off
Inter Area Lsp: TRUE
    Session Protection: FALSE
    Targeted Session Keepalive HoldTime/Interval: 180/60 sec
Targeted Hello HoldTime/Interval: 90/10 sec
FS# show mpls ldp bindings
Default VRF:
  lib entry: 10.10.10.10/32
      local binding:   to lsr:    10.10.10.10:0, label: 1027
      local binding:   to lsr:    30.30.30.30:0, label: 1027
      remote binding: from lsr: 10.10.10.10:0, label: imp-null
      remote binding: from lsr: 30.30.30.30:0, label: 1030 (not in FIB)
  lib entry: 9.9.9.9/32, (IA)
      local binding:   to lsr:    10.10.10.10:0, label: 1028
      local binding:   to lsr:    30.30.30.30:0, label: 1028
      remote binding: from lsr: 10.20.10.10:0, label: 1027 (not in FIB)
      remote binding: from lsr: 30.30.30.30:0, label: 1031
```

1.5 Monitoring and Maintenance

Clearing

Running the **clear** commands may lose vital information and thereby interrupt services.

Description	Command
Resets existing LDP sessions.	**clear mpls ldp neighbor** [**all | vrf** *vrf-name*] [* | *ip-address*]

Displaying

Description	Command
Displays basic MPLS information.	**show mpls summary**

Description	Command
Displays the MPLS forwarding table.	**show mpls forwarding-table** [**detail**] **show mpls forwarding-table** { *ip-address/mask* | *ipv6-address/mask* } [**detail**] **show mpls forwarding-table label** *label* [**detail**] **show mpls forwarding-table interface** *interface-name* [**detail**] **show mpls forwarding-table next-hop** { *ip-address* | *ipv6-address* } [**detail**] **show mpls forwarding-table** { **ftn** | **ilm** } [**ip** | **ipv6** | **vc**] [**detail**] **show mpls forwarding-table** { **vrf** *vrf-name* | **global** } [**ftn** | **ilm**] [**ip** | **ipv6**] [**detail**] **show mpls forwarding-table summary**
Displays usage of the MPLS label pool.	**show mpls label-pool**
Detects LSP connectivity.	**ping mpls ipv4** *ip-address/mask* [**repeat** *repeat*] [**ttl** *time-to-live*] [**timeout** *timeout*] [**size** *size*] [**interval** *mseconds*] [**source** *ip-address*] [**destination** *ip-address*] [**force-explicit-null**] [**pad** *pattern*] [**reply mode** { **ipv4** | **router-alert** }] [**dsmap**] [**flags fec**] [**verbose**]
Detects LSR nodes that LSPs pass through.	**traceroute mpls ipv4** *ip-address/mask* [**timeout** *timeout*] [**ttl** *ttl*] [**source** *ip-address*] [**destination** *ip-address*] [**force-explicit-null**] [**reply mode** { **ipv4 | router-alert** }] [**flags fec**] [**verbose**]
Displays FEC-label binding.	**show mpls ldp bindings** [**all | vrf** *vrf-name*] [*ip-address/mask* | **label label**] [**remote | local**]
Displays discovered LDP neighbors.	**show mpls ldp discovery** [**all | vrf** *vrf-name*] [**detail**]
Displays LDP interfaces.	**show mpls ldp interface** [**all | vr**f *vrf-name* | *interface-name*]
Displays LDP neighbors.	**show mpls ldp neighbor** [**all | vrf** *vrf-name*] [*ip-address*] [**detail**]
Displays LDP attributes.	**show mpls ldp parameters** [**all** | **vrf** *vrf-name*]
Displays MPLS express forwarding.	**show mpls ref ftn-ipv4** [**global** | **vrf** *vrf_name*] [*ip_address / mask* [**detail**]] **show mpls ref ftn-ipv6** [**global** | **vrf** *vrf_name*] [*ipv6_address / mask* [**detail**]] **show mpls ref ilm** [**in-label** *label* [**detail**] | **summary**] **show mpls ref nhlfe** [[*nhlfe_id* [**detail**]] | **summary**] **show mpls ref summary**
Displays the routing information base (RIB).	**show mpls rib** [**all** | **vrf** *vrf-name*] [**ipv6**]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs LDP session messages.	**debug mpls ldp message**
Debugs LDP label binding.	**debug mpls ldp binding**
Debugs internal processing of MPLS entries.	**debug mpls**
Debugs MPLS messages.	**debug mpls msg** [**send** | **recv**]
Debugs the MPLS express forwarding control plane.	**debug mpls ref control**
Debugs MPLS express forwarding.	**debug mpls ref packet**
Debugs events during execution of the **ping mpls** or **traceroute mpls** command.	**debug mpls lspv event**
Debugs packets during execution of the **ping mpls** or **traceroute mpls** command.	**debug mpls lspv packet**

Debugs packet encapsulation TLV during execution of the **ping mpls** or **traceroute mpls** command.	**debug mpls lspv tlv**
Debugs errors during execution of the **ping mpls** or **traceroute mpls** command.	**debug mpls lspv error**

2 Configuring BGP/MPLS IP VPN

2.1 Overview

Conventional virtual private networks (VPNs) transmit private network data in public networks over tunneling protocols such as Generic Routing Encapsulation (GRE), Layer 2 Tunneling Protocol (L2TP), and Point-to-Point Tunneling Protocol (PPTP). BGP/MPLS IP VPN is a VPN implementation mode between L2TP and Layer 3 Tunneling Protocol (L3TP). Label Switched Paths (LSPs) are established in public networks over the Label Distribution Protocol (LDP) for Multiprotocol Label Switching (MPLS). MPLS VPN connects the geographically different branches of a private network over LSPs to form a unified network. MPLS VPN supports the interworking between different VPNs.

Carriers implement MPLS in IP backbone networks to provide the VPN service to users' private networks. MPLS allows users to build VPNs using routers instead of dedicated VPN devices. MPLS also allows carriers to scale VPNs easily.

The router devices and router icons in this document represent general routers and Layer-3 switches that run routing protocols. Specific device forms are indicated by "router series products" or "switch series products".

2.2 Applications

Application	Description
Single-AS VPN Service Model	Carriers provide the VPN service to users' IP networks. Multiple sites of a user's VPN are connected to the same IP/MPLS autonomous system (AS).
Inter-AS VPN Service Model	Carriers provide the VPN service to users' IP networks. Multiple sites of a user's VPN are connected to different IP/MPLS ASs.

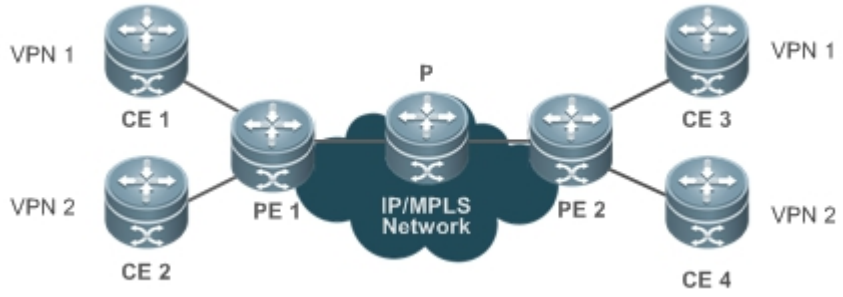
2.2.1 Single-AS VPN Service Model

Scenario

Carriers provide the MPLS VPN service to users' IP networks. Multiple sites of a user's VPN are connected to the same IP/MPLS AS and communicate with each other by exchanging VPN routes across the AS.

In Figure 2- 1, two sites of VPN1 are connected to the same IP/MPLS AS.

Figure 2- 1



Deployment

- Provider edge (PE) 1 and PE2 communicate with the provider (P) device over the Open Shortest Path First (OSPF) and Label Distribution Protocol (LDP), in order to implement the MPLS public network function.
- The customer edge (CE) devices communicate with the PE devices over OSPF, and the PEs communicate with each other over the Border Gateway Protocol (BGP), in order to implement OSPF access to MPLS L3VPN.
- The CEs communicate with the PEs over the Routing Information Protocol (RIP), and the PEs communicate with each other over BGP, in order to implement RIP access to MPLS L3VPN.
- The CEs communicate with the PEs over the Intermediate System to Intermediate System (IS-IS), and the PEs communicate with each other over BGP, in order to implement IS-IS access to MPLS L3VPN.
- The CEs communicate with the PEs over BGP, and the PEs communicate with each other over BGP, in order to implement BGP access to MPLS L3VPN.
- The CEs and the PEs exchange routes based on static configurations, and the PEs communicate with each other over BGP, in order to implement static access to MPLS L3VPN.

2.2.2 Inter-AS VPN Service Model

Scenario

Carriers provide the MPLS VPN service to users' IP networks. Multiple sites of a user's VPN are connected to different IP/MPLS ASs and communicate with each other by exchanging VPN routes across different ASs.

In Figure 2- 2, two sites of VPN1 are connected to two IP/MPLS ASs.

Figure 2- 2

Deployment

- The MPLS public network function is implemented between AS1 and AS2.
- Autonomous system boundary routers (ASBRs) 1 and ASBR2 are CEs that implement MPLS L3VPN over inter-AS Option A.
- The PEs communicate with each other over BGP to implement MPLS L3VPN over inter-AS Option B.
- The PEs communicate with each other over BGP to implement MPLS L3VPN over inter-AS Option C.

2.3 Features

Overview

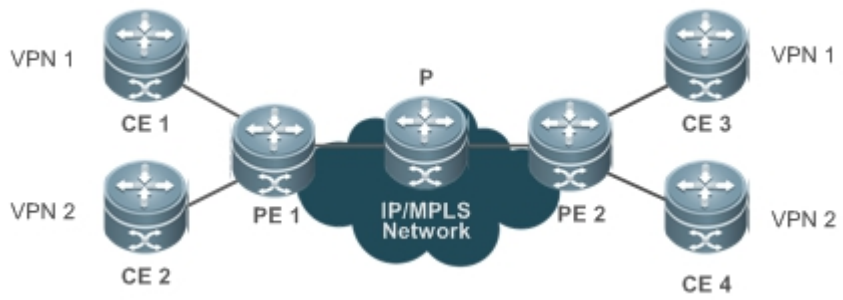
Feature	Description
BGP/MPLS IP VPN Basic Functions	Uses legacy IPv4 backbone networks to provide the VPN service to users' IPv4 networks.
Inter-AS VPN Service Model	Provides the VPN service to users' IPv4 networks across multiple IPv4 backbone networks.
OSPF VPN Extended Attributes	Implements OSPF between PE and CE to simplify CE configuration and management.

2.3.1 BGP/MPLS IP VPN Basic Functions

Network Structure

CE, PE, and P play important roles in the BGP/MPLS VPN network structure, as shown in Figure 2- .

Figure 2- 4



- CEs are located on the edge of user networks and belong logically to users' VPNs. A port on each CE is connected directly to a PE. A CE can be a router, switch, or host, and it may or may not support MPLS.
- PEs are located on the edge of the service provider (SP)'s backbone network and belong logically to the SP. A PE can be a router, ATM switch, or frame relay (FR) switch. PEs are connected directly to CEs, and one PE can be connected to different CEs. PEs receive VPN information from CEs, send VPN information to other PEs, receive VPN information from other PEs, and distribute the received VPN information to corresponding CEs. PEs must support MPLS.
- The P router is the core device of the SP's backbone network. The P router implements routing and fast forwarding. It is not connected to CEs. The P router must support MPLS. It knows the route for any destination in the backbone network, but does not know the routes destined for the VPNs.

Application Features

- VPN tunnels are established on the SP's PEs, rather than on users' CEs. VPN routes are exchanged between PEs. Users do not need to maintain VPN information.
- Existing routing protocols are used to dynamically establish VPN tunnels and advertise routes, making it easy to scale VPNs.
- The use of overlapping addresses is supported, allowing different VPN users to use the same address space.
- VPN service content is exchanged by label switching rather than route distribution in the SP's network.

- VPNs deliver security comparable to that of leased lines.

Implementation Mechanism

- LSP tunnels are established over LDP in the backbone network. (This process is completed when the SP's network topology is stable.)
- Packets are forwarded according to their labels and the local mapping table.
- MP-BGP and BGP extended attributes are used to transmit VPN routes and carry VPN attributes, labels, and other content.
- VPN routes are managed by creating multiple routing tables, maintaining VPN routes, and other methods.

VRF

A VPN routing and forwarding (VRF) table is used to solve local route conflicts.

Each PE-CE connection must be associated with a VRF table. A PE may have multiple VRF tables used to exchange routes with CEs. Each VRF table can be considered as a virtual router, which is connected to a CE and receives routes from the CE or advertises VPN routes to the CE. VRF tables allow PEs to support different VPNs that use the same address space without causing local route conflicts.

A VRF table consists of:

- An independent routing table
- A set of interfaces that belong to the VRF table
- A group of routing protocols only used by the VRF table

RD

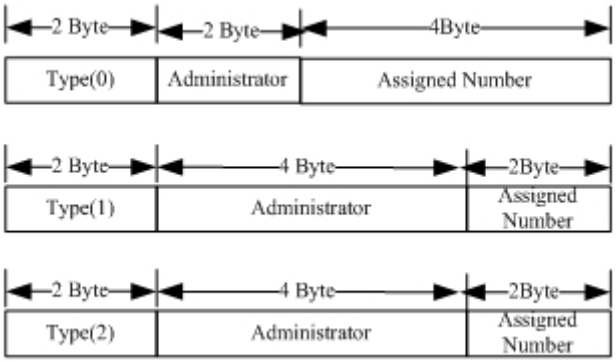
A route distinguisher (RD) is an important attribute of VPN routes and used to solve route conflicts during the transmission process.

When routes are advertised over BGP in the backbone network to different VPNs with overlapping addresses, BGP selects the best route among those addresses for advertising; as a result, some VPNs cannot receive corresponding routes. To address this problem, different RDs are added to the overlapping addresses, allowing BGP, when making routing decisions, to distinguish those addresses according to the different RDs carried in VPN routes. In this way, all VPNs receive respective routes. RDs are only used to distinguish identical network addresses, and they are not required when different VPNs do not have overlapping addresses.

A unique RD is assigned to one VPN, so that different VPNs have different RDs to ensure normal transmission of routes over the backbone network. RDs are defined in the format of XX: XX, for example, RD 1: 100, in which 1 is an AS number in the backbone network and 100 is a user-assigned number. A VPN route carries only one RD.

An RD consists of the **Type** field, an Administrator sub-area, and an Assigned Number. Depending on the value of the **Type** field, the RD is encoded in one of the following three formats:

Figure 2- 5 RD Structure



- When **Type** is **0**, the Administrator sub-area uses a two-byte public AS number, and the Assigned Number has four bytes and is managed by the SP.
- When **Type** is **1**, the Administrator sub-area uses a four-byte global IPv4 address, and the Assigned Number has two bytes and is managed by the SP.
- When **Type** is **2**, the Administrator sub-area uses a four-byte AS number, and the Assigned Number has two bytes and is managed by the SP.

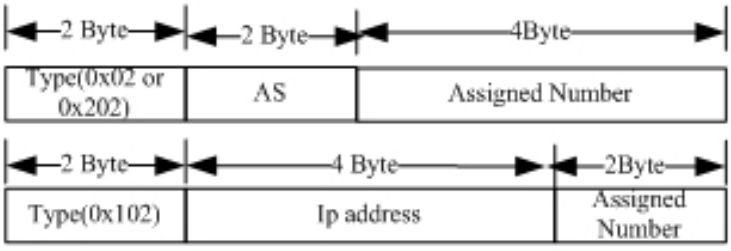
Route-Target

The route target (RT) is an important attribute of VPN routes and used to indicate a route selection method.

RT attributes are classified into the export RT attribute and the import RT attribute. When a PE receives a VPN route from a CE, the PE adds the export RT attribute to the VPN route and advertises the route to the peer PE. When receiving a route from the peer PE, the PE determines whether to import the route to a VRF table according to the import RT attribute. The import principle is as follows: The PE imports the incoming VPN route to a VRF table only when the VPN route has an RT attribute the same as the import RT attribute of the VRF table. In this way, the distribution of VPN routes is controlled flexibly. A VPN route can carry multiple RT values.

The BGP extended community attribute defines the RT encoding structure, as shown in Figure 2- .

Figure 2- 5 RT Structure



The RT is defined in a way similar to the RD. The RT with **Type** set to 0x02 or 0x202 must have a public AS number. The RT with **Type** set to 0x102 must have a global IPv4 address.

MP-BGP

VPN routes are transmitted over BGP in the backbone network, and carry the export RT attribute according to the BGP extended community attribute.

The traditional BGP4 protocol only transmits IPv4 routes. BGP must be extended for transmitting RD-included VPN routes. BGP is highly scalable and supports the definition of new attributes. Multi-Protocol BGP (MP-BGP) introduces new attributes based on BGP to support multiple protocols. MP-BGP can carry VPN routing information in the format of RD+IP address prefix. When receiving IPv4 routes from VPN users, MP-BGP adds RDs to the IPv4 routes to convert the routes into VPN-IPv4 routes, and then transmits the VPN-IPv4 routes between PEs in the backbone network.

2.3.2 Inter-AS VPN Service Model

Inter-AS VPN allows different sites of a user's VPN located in different ASs to communicate with each other by exchanging VPN routes across those ASs.

RFC4364 proposes the following three inter-AS VPN solutions:

- Option A: VRF-to-VRF mode
- Option B: single-hop MP-EBGP mode
- Option C: multi-hop MP-EBGP mode

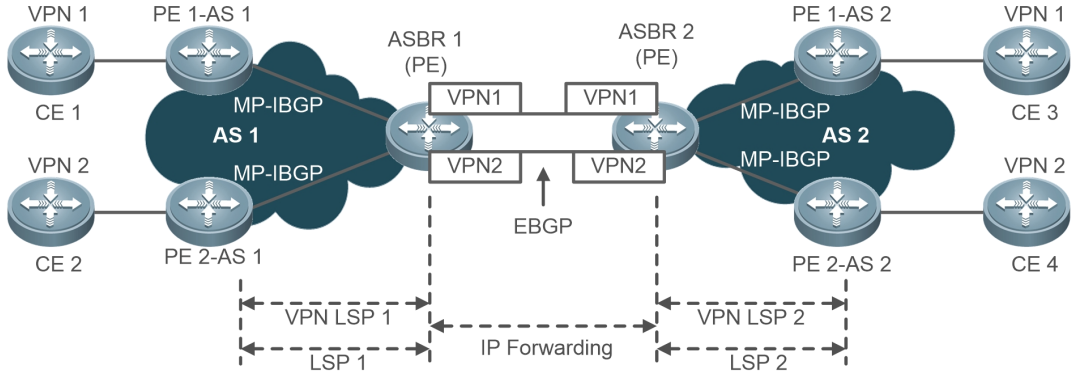
2.3.2.1 Option A: VRF-to-VRF Mode

The Option A (VRF-to-VRF) solution is easy to implement. The ASBRs in ASs create a VRF table for every VPN that requires inter-AS communication, and bind an interface to each VRF table. The VRF tables between ASBRs use their bound interfaces to exchange VPN routes.

The purposes of creating VRF tables and binding interfaces are:

- To receive VPN routes in the local AS
- To set up EBGP connections between VRF tables in different ASs to exchange IPv4 routes.

Figure 2- 7 VRF-to-VRF Inter-AS VPN



In Figure 2-7, the VRF tables between the two ASBRs establish an EBGP session to exchange IPv4 routes, and the ASBRs and the PEs establish an MP-IBGP session to exchange VPN routes. The VRF tables engaged in the EBGP session consider the peer as a CE. VRF-to-VRF

configuration is similar to intra-AS configuration, requiring to establish an MP-IBGP session between ASBR and PE to exchange VPN routes and establish an EBGP session between VRF tables on different ASBRs in BGP VRF address family configuration mode to exchange IPv4 routes.

Features and Limitations

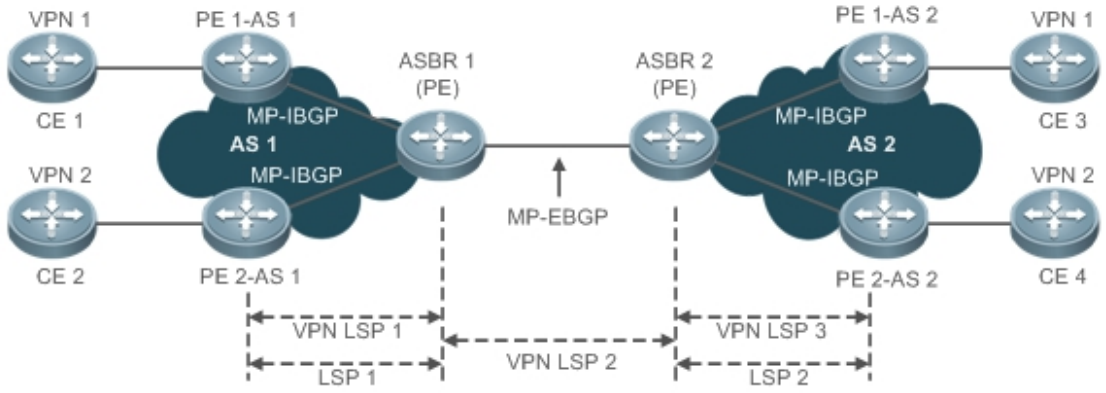
The Option A (VRF-to-VRF) solution is easy to implement and deploy over MP-IBGP. However, ASBRs must configure and bind an interface (a logical sub-interface) to every VPN that requires inter-AS communication. The number of bound interfaces must be at least equal to the number of the VPNs. In addition, VPNs must be configured one by one on the ASBRs. This causes scalability issues. The creation of sub-interfaces for individual VPNs also poses additional requirements on ASBRs. Therefore, the Option A solution is applicable to networks with few VPNs that require inter-AS communication.

Option A configuration is similar to BGP/MPLS VPN configuration. The configuration process is omitted in this document.

2.3.2.2 Option B: Single-hop MP-EBGP Mode

In the Option A solution, ASBRs must configure a VRF table for each VPN and bind an interface to the VRF table. This is because VPN routes can only be carried over MP-IBGP, but cannot be exchanged directly between EBGP peers. VRF configuration on ASBRs is not required if VPN routes can be exchanged directly between EBGP peers. The Option B solution extends MP-IBGP for direct distribution of VPN routes between ASBRs. The solution is also called single-hop MP-EBGP. Figure 2- shows the topology.

Figure 2- 8 Option B Inter-AS VPN



Features and Limitations

The Option B solution removes the need to allocate sub-interfaces to sites of each VPN on ASBRs and the need to establish inter-AS LSPs. Instead, VPN routes are transmitted directly between single-hop MP-EBGP neighbors. Because VPN routes are stored and propagated by ASBRs between ASs, the ASBRs are under great burden when handling a large amount of VPN routes. The ASBRs also need to forward public IP packets. This poses high requirements on ASBRs. Because ASBRs disable the RT filter function for handling incoming VPN routes, the VPN routes on a PE may be propagated to the ASBR in a different AS, which may cause VPN route leakage. To address this problem, SPs that exchange VPN routes must implement a confidence-based routing protocol, enabling ASBRs to trust each other and implement route filtering policies. The Option B solution is applicable to networks with many VPNs that require inter-AS communication.

Implementation

The Option B solution is implemented by the following methods:

- The ASBR retains the next hop in incoming VPN routes.
- The ASBR changes the next hop in incoming VPN routes.

The two methods are described in the following.

Method 1: The ASBR Retains the Next Hop in Incoming VPN Routes

When an ASBR receives a VPN route from the peer ASBR in a different AS, the ASBR sends the VPN route with the next hop unchanged to its MP-IBGP neighbors in the local AS. The PEs and the ASBR in the same AS exchange VPN routes over an MP-IBGP session, and two ASBRs exchange VPN routes directly over an MP-EBGP session. Because the ASBR does not change the next hop in the VPN route received from the MP-EBGP peer, the PEs (MP-IBGP neighbors) in the local AS must have the route destined for the next hop (an ASBR in the other AS). To do this, the ASBR redistributes this route to IGP in the local AS, in order to make the destination ASBR reachable. Then an LSP is established over LDP.

The configuration procedure is as follows:

11. Configure PEs and CEs to exchange routes.
12. Configure IGP and MPLS signaling for ASs.
13. Disable the default RT filter function on ASBRs.
14. Configure the PEs and the ASBR in the same AS to exchange VPN routes.
15. Configure ASBRs to establish an MP-EBGP session.
16. (Optional) Configure a route map for filtering VPN routes.
17. Configure IGP redistribution of the route destined for the ASBR in the other AS.

Method 2: The ASBR Changes the Next Hop in Incoming VPN Routes

When an ASBR receives a VPN route from the peer ASBR in a different AS, the ASBR changes the next hop in the incoming VPN route before sending the route to PEs in the local AS. The PEs and the ASBR in the same AS exchange VPN routes over an MP-IBGP session. The two ASBRs in different ASs exchange VPN routes over an MP-EBGP session. When receiving a VPN route from one ASBR, the other ASBR changes the route's next hop to itself before sending the route to MP-IBGP peers in the local AS.

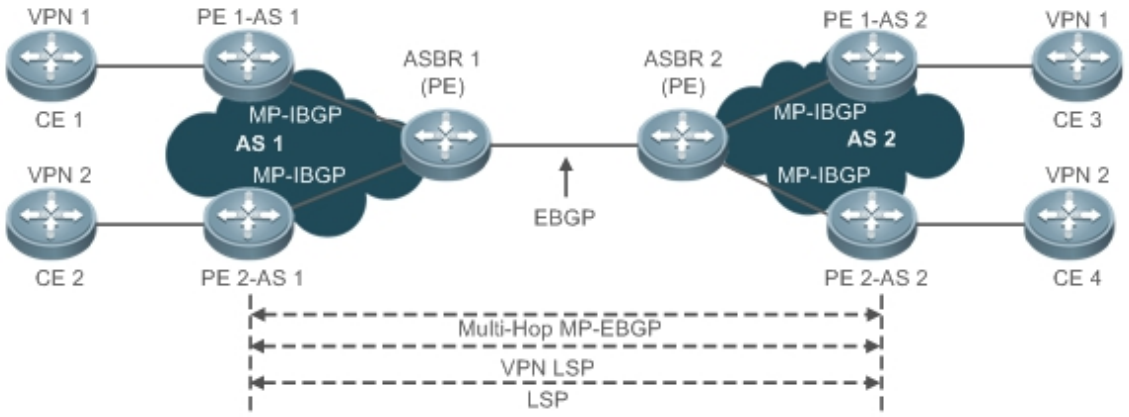
The configuration procedure is as follows:

17. Configure PEs and CEs to exchange routes.
18. Configure IGP and MPLS signaling for ASs.
19. Disable the default RT filter function on ASBRs.
20. Configure the PEs and the ASBR in the same AS to establish an MP-IBGP session and enable the ASBR to change the next hop in incoming VPN routes to itself.
21. Configure ASBRs to establish an MP-EBGP session.
22. (Optional) Configure a route map for filtering VPN routes.

2.3.2.3 Option C: Multi-hop MP-EBGP Mode

The Option A and Option B solutions implement inter-AS VPN networking. A common feature of the two solutions is that ASBRs maintain and advertise VPN routes. ASBRs may become a bottleneck of network scalability when the ASBR in each AS needs to advertise a large amount of inter-AS VPN routes. The Option C (multi-hop MP-EBGP) solution is developed to address the scalability issue. The solution allows PEs in different ASs to exchange VPN routes directly over a multi-hop MP-EBGP session, removing the need for ASBRs to maintain and distribute VPN routes.

Figure 2- 9 Multi-hop MP-EBGP



Features and Limitations

In multi-hop MP-EBGP mode, only PEs maintain VPN routes, whereas ASBRs do not store VPN user information. The configuration procedure is relatively complex. The solution is applicable to inter-AS VPN service provision in large scale.

Implementation

The Option C solution is implemented by the following methods:

- Enable IPv4 label distribution only between EBGP neighbors.
- Enable IPv4 label distribution between EBGP neighbors and between IBGP neighbors.

For Option C scalability, a route reflector (RR) is deployed in each AS, and the RRs between two ASs establish a multi-hop MP-EBGP session to exchange VPN routes. From the deployment perspective, Option C is also called the inter-RR multi-hop MP-EBGP solution.

The configuration of each Option C implementation method is described in the following.

Method 1: Enable IPv4 Label Switching Only Between EBGP Neighbors

In this implementation method, BGP routes must be redistributed using the IGP protocol (for example, OSPF or RIP) on ASBRs, so that each device in one AS has the route destined for each PE in the other AS. LSPs are established by distributing labels over LDP to the routes destined for PEs between ASs. IPv4 label switching is enabled between the directly connected inter-AS ASBRs. BGP assumes the role of MPLS signaling to distribute labels to the routes destined for PEs between ASs for LSP establishment.

The configuration procedure is as follows:

18. Configure PEs and CEs in each AS to exchange routes.
19. Configure IGP and MPLS signaling for ASs.
20. Configure ASBRs to establish an EBGP session in order to distribute labels to IPv4 routes.
21. Configure ASBRs to redistribute to IGP the routes of PEs that are learned over EBGP.
22. Configure a multi-hop MP-EBGP session.

Method 2: Enable IPv4 Label Switching Between EBGP Neighbors and Between IBGP Neighbors

As Method 1 enables IPv4 label switching only between EBGP neighbors, IGP and LDP in one AS need to maintain the routes destined for PEs in the other AS. That is, the inter-AS routes of PEs must be propagated to all devices between ASs. In actual application, for AS security consideration, the routes of PEs in one AS should not be propagated to each device in the other AS. Rather, the routes must be maintained only by BGP and transparent to IGP and LDP in the same AS. To meet this requirement, you can enable IPv4 label switching between EBGP neighbors and between IBGP neighbors.

Different from Method 1, Method 2 does not require IGP on ASBRs to redistribute BGP routes and does not require LDP to distribute labels to BGP routes. LDP is only responsible for LSP establishment in the local AS. However, IPv4 label switching must be enabled between IBGP neighbors and between EBGP neighbors for inter-AS LSP establishment, and PEs must support Stacking of three consecutive labels.

The configuration procedure is as follows:

1. Configure PEs and CEs in each AS to exchange routes.
2. Configure IGP and MPLS signaling for ASs.
3. Configure PEs and ASBRs to establish an IBGP session in order to distribute labels to IPv4 routes.
4. Configure ASBRs to establish an EBGP session in order to distribute labels to IPv4 routes.
5. Configure a multi-hop MP-EBGP session.

2.3.4 OSPF VPN Extended Attributes

OSPF is typically used as an IGP internal routing protocol for users' VPNs in many application cases. No other routing protocols are required if OSPF is implemented between PE and CE, thus simplifying CE configuration and management.

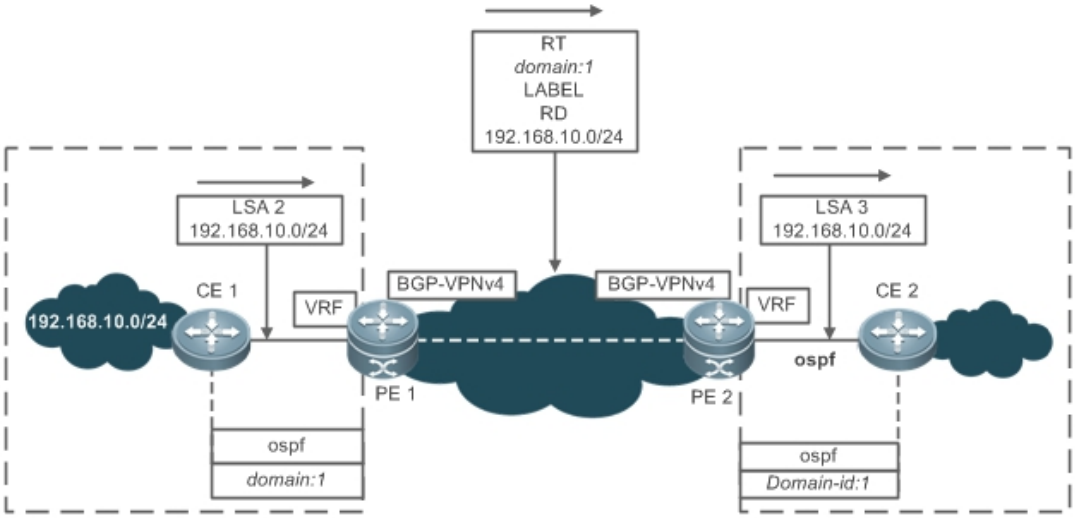
The following describes the OSPF extended attributes between PE and CE.

Domain ID

A domain ID indicates the OSPF domain to which a route belongs. When an OSPF route for a VPN site that a CE learns is advertised as a type 1, 2, or 3 link-state advertisement (LSA) to a PE and redistributed to BGP to form a VPN route, the domain ID is also redistributed with the route to BGP and advertised as an extended community attribute of the VPN route. When the peer PE receives the VPN route and redistributes it to a VRF OSPF instance, the domain ID is also redistributed with the route to the VRF OSPF instance. If the VRF OSPF instance confirms that the domain ID of the route is the same as its domain ID, it advertises the route as an internal route to a CE. If the VRF OSPF instance confirms that the domain ID of the route is different from its domain ID, it advertises the route as an external route to the CE.

In Figure 2- 3, CE1 advertises routes in the local OSPF domain as type 2 LSAs to PE1, which then forms VPN routes and advertises the routes to PE2. PE2 redistributes incoming routes to a VRF OSPF instance. Because the domain IDs of the VPN routes are the same as the domain ID of the VRF OSPF instance, the routes are advertised as internal routes to CE2.

Figure 2- 3

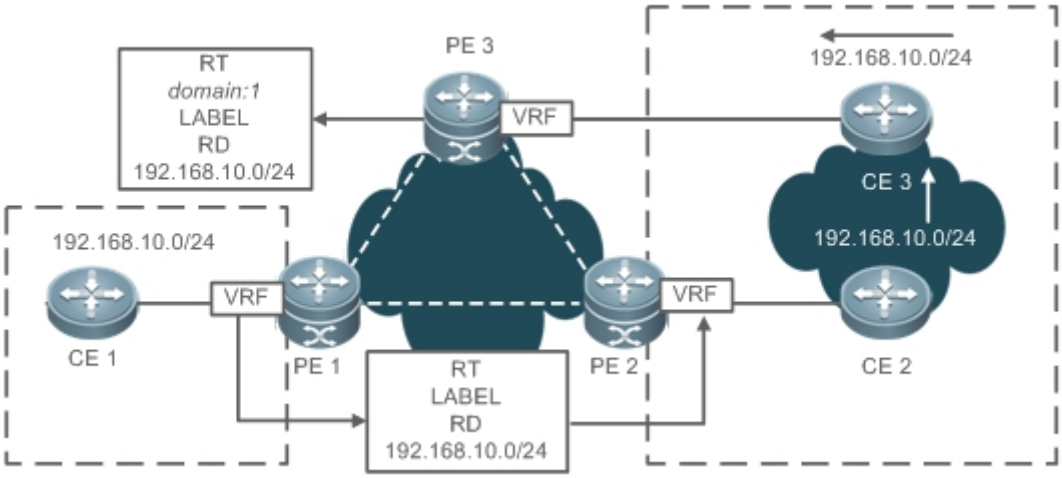


DN Bit

The down-bit (DN bit) is used for loop detection between PE and CE running OSPF. Loops may occur when a PE advertises a learned VPN route to its connected VPN site and the VPN site transmits the route over OSPF to another connected PE, which then propagates the route.

In Figure 2- 4, PE1 advertises the route 192.168.10.0/24 to PE2 and PE3. CE2 advertises the route to CE3 over OSPF, which then advertises the route to PE3. In this way, the route is redistributed to BGP on PE3, which preferentially selects OSPF for route redistribution and converts the route into a VPNv4 route before advertising the route. In this case, a loop may occur.

Figure 2- 4



To avoid loops, PEs insert the DN bit as an optional field into type 3, 5, or 7 LSAs before advertising the LSAs to CEs. When other PEs receive LSAs containing the DN bit, OSPF on the PEs exclude the LSAs from OSPF route calculation.

VPN Route Tag

The VPN route tag provides another method of detecting loops between PE and CE running OSPF. When a PE and a CE communicate over OSPF, the corresponding VRF OSPF instance on the PE has a VPN route tag by default. When the VRF OSPF instance imports a VPN route and advertises it as a type 5 or 7 LSA to the CE, the LSA carries the VPN route tag. In an environment where one VPN site is connected to multiple PEs, when a PE receives a type 5 or 7 LSA that carries the same VPN route tag as that of its VRF OSPF instance, the LSA is excluded from OSPF route calculation.

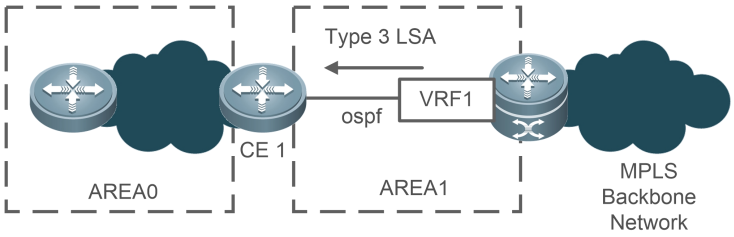
Area Deployment

In normal cases, the link between a PE and a CE can be any OSPF area. If the PE-CE link is a non-0 area, the OSPF area where the CE is located considers the PE as an area border router (ABR). This may cause problems due to the following characteristics of OSPF with the ABR feature:

- The ABR calculates only type 3 LSAs of the backbone area.
- The ABR forwards only type 3 LSAs from the backbone area to other areas.

In Figure 2- , if the PE-CE link is a non-0 area, the PE redistributes the VPNv4 route advertised over MP-BGP to OSPF and restores the route to a type 3 LSA before adverting it to CE1. Because CE1 does not calculate LSAs of non-backbone areas, the type 3 LSA is not advertised to the router in Area0. As a result, VPN sites may fail to learn routes between each other. Therefore, non-0 areas between PE and CE must be deployed correctly.

Figure 2- 16

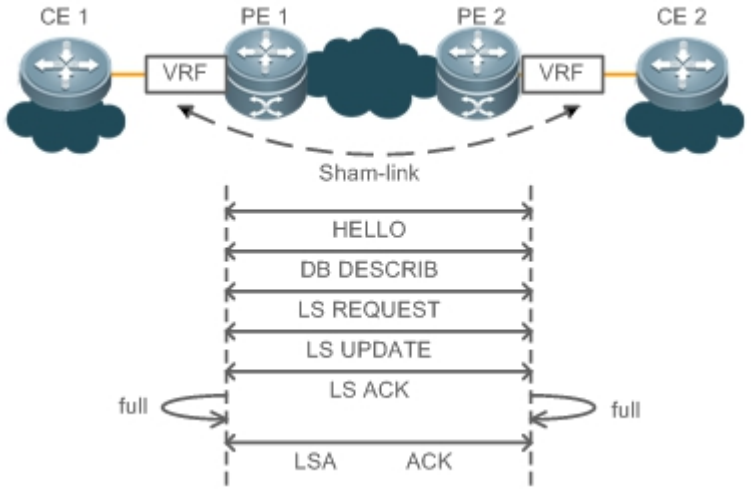


In L3VPN application, if PEs and CEs exchange VPN routes over OSPF, you are advised not to deploy the backbone area at a VPN site. In the actual condition, if the backbone area has PEs as well as routers of a VPN site, at least one router series product of the VPN site must be connected to a PE, and the CE-PE link must belong to Area0. In this way, inter-area routes and external routes can be transmitted between the PEs and the VPN site.

Sham Link

A sham link is not an actual link, but a virtual link between the VRF tables on two PEs. Like OSPF links, sham links have OSPF interfaces and can be used to send OSPF packets, establish neighbor relationships, and send LSAs. When LSAs are flooded out over sham links, all OSPF route types remain unchanged, as shown in Figure 2- 17.

Figure 2- 17



The sham link established between VRF OSPF instances on different PEs have the following functions:

- The sham link transmits VPN routes to the peer PE over MP-IBGP, and original OSPF routes are imported on a best-effort basis on the peer PE. The sham link does not support real communication of OSPF topology information. An OSPF link is established over a sham link to connect OSPF instances of various sites and maintain complete topology information.
- Different sites of the same VPN exchange routes over the MPLS backbone network. These sites are connected by a backdoor link to support inter-site communication when the MPLS backbone network is not available. Two sites of a VPN that are located in the same OSPF area and connected by a backdoor link exchange routes over the MPLS backbone network as well as the backdoor link. Because the intra-AS routes exchanged over the backdoor link take precedence over the inter-AS routes exchanged over the MPLS network, routes of the two sites are forwarded preferentially over the backdoor link. This does not comply with the intention of VPN users connecting to the backdoor link. The sham link is introduced to ensure proper use of the backdoor link.

2.4 Configuration

Interfaces used by BGP/MPLS IP VPN must be Layer-3 routing interfaces; otherwise, configuration fails.

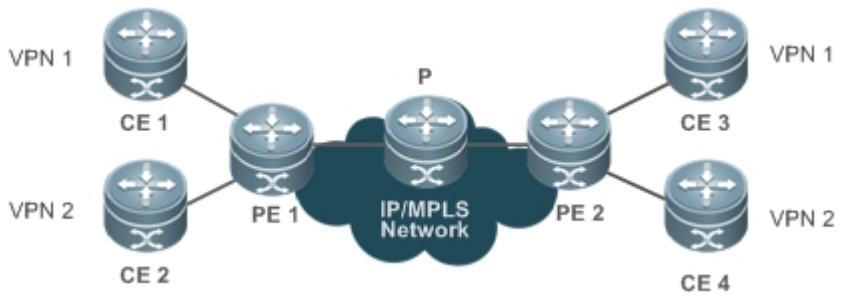
Configuration	Description
Configuring BGP MPLS IP VPN Basic Functions	(Mandatory) It allows carrier networks to provide the single-AS VPN service.
Configuring the Inter-AS VPN Service Model: Option A	(Optional) It is applicable to carrier networks with light traffic of the inter-AS VPN service.
Configuring an Inter-AS VPN Service Model: Option B	(Optional) It is applicable to carrier networks with heavy traffic of the inter-AS VPN service.
Configuring an Inter-AS VPN Service Model: Option C	(Optional) It is applicable to carrier networks requiring large-scale VPN service provision.
Configuring OSPF VPN Extended Attributes	(Optional) It is applicable to VPN sites that run OSPF.

2.4.1 Configuring BGP MPLS IP VPN Basic Functions

Configuration Effect

Build the BGP/MPLS VPN service in a single AS.

Figure 2- 17



Notes

- The LDP router ID must consist of 32 bits.
- The BGP router ID must consist of 32 bits.
- The **mpls ip** and **label-switch** commands must be enabled on PEs' interfaces connected to a public network.

Configuration Steps

The configuration of BGP/MPLS VPN basic functions includes:

- (Mandatory) Configure an MPLS network.
- (Mandatory) Configure a VPN routing instance.
- (Mandatory) Configure PEs to exchange VPN routes.
- (Mandatory) Configure PEs and CEs to exchange routes.
- (Optional) Configure a method of VPN label distribution.
- (Optional) Configure VPN route import/export policies.
- (Optional) Configure static L3VPN FEC to NHLFE (FTN) and incoming label map (ILM).

Configuring an MPLS Network

Before you implement MPLS in the backbone network, configure MPLS-specific LDP on the P and PE devices for public tunnel establishment. The configuration includes configuring LDP on corresponding MPLS devices and enabling MPLS packet forwarding on every interface.

Perform configuration on the P and PE devices according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **mpls ip**	Enables MPLS globally.

FS(config)# **mpls router ldp**	Configures LDP and enters LDP configuration mode.
FS(config-mpls-router)# **ldp router-id interface loopback** *id* [**force**]	Configures the LDP router ID, which typically uses the IP address of a loopback interface.
FS(config-mpls-router)# **exit**	Exits LDP configuration mode.
Ruiije(config)# **interface type** *ID*	Enters interface configuration mode.
FS(config-if-*type ID*)# **ip address** *ip-address mask*	Configures the interface IP address.
FS(config-if-*type ID*)# **label-switching**	Enables MPLS packet forwarding on the interface connected to the public network.
FS(config-if-*type ID*)# **mpls ip**	Enables LDP on the interface.
FS(config-if-*type ID*)# **show running-config**	Displays existing configurations.

Configuring a VPN Routing Instance

VPN routing instances are also called VRF tables, which are configured on the PEs. The CEs and the P do not have VRF tables.

To configure a VRF table, you need to define the VRF table, configure RD and RT values for the VRF table, and associate the VRF table with an interface.

Perform configuration on the PEs according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **ip vrf** *vrf-name*	Creates a VRF table and enters VRF configuration mode.
FS(config-vrf)# **rd** *rd-value*	Configures an RD value.
Ruiije(config-vrf)# **route-target** { **both** | **export** | **import** } *rt-value*	Configures RT values.
FS(config-vrf)# **exit**	Exits VRF configuration mode.
FS(config-if-*type ID*)# **interface type** *ID*	Enters interface configuration mode.
FS(config-if-*type ID*)# **ip vrf forwarding** *vrf-name*	Associates the VRF table with an interface.
FS(config-if-*type ID*)# **ip address** *address mask*	Configures the interface IP address.
FS(config-if-*type ID*)# **show running-config**	Displays existing configurations.

After an RD value is defined or BGP VRF is enabled for a VRF table on a PE, the RD value cannot be modified or deleted. To modify the RD value, delete the VRF table, create a new one, and then set its RD value.

The RD values of two different VRF tables on a PE must not be the same.

After the **ip vrf forwarding vrf-name** command is executed, the IP address of the specified interface is deleted. You need to configure an IP address again in interface configuration mode.

Configuring PEs to Exchange VPN Routes

To configure PEs to exchange VPN routes (not IPv4 routes) over BGP, enter VPN address family configuration mode to enable the capability of transmitting VPN routes to the peer PE.

Perform configuration on the PEs according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-num*	Creates BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *ip-address* **remote-as** *asn-number*	Configures a BGP session.
Ruiije(config-router)# **neighbor** *ip-address* **update-source** *interface-name*	Specifies the address of a loopback interface as the source address for the MP-IBGP session.
Ruiije(config-router)# **address-family vpnv4**	Enters VPN address family configuration mode.
FS(config-router-af)# **neighbor** *ip-address* **activate**	Activates the function of exchanging VPN routes over BGP.
FS(config-router-af)# **show running-config**	Displays existing configurations.

Configuring PEs and CEs to Exchange Routes

Four methods are available for configuring the VRF tables on PEs and the CEs to exchange routes.

Method 1: Implement BGP Between PE and CE

After a BGP session is established between PEs, establish a BGP session between PE and CE to connect all CEs over BGP. Redistribute VPN routes between CEs and BGP to build a fully meshed VPN.

Enter VRF address family configuration mode on the PEs and configure a BGP session between PE and CE according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *pe-asn-num*	Creates BGP and enters BGP configuration mode.
FS(config-router)# **address-family ipv4 vrf** *vrf-name*	Performs configuration and enters BGP VRF address family configuration mode.
Ruiije(config-router-af)# **neighbor** *ip-address* **remote-as** *ce-asn-num*	Establishes an EBGP session between PE and CE.

If no RD value is defined for the specified VRF table, the system prompts that the RD value is not configured when you run the **address-family ipv4 vrf vrf-name** command, and you are not allowed to enter address family configuration mode under the VRF table.

Configure PE-CE BGP sessions on the CEs and introduce VPN routes to BGP according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *ce-asn-num*	Creates BGP and enters BGP configuration mode.
Ruiije(config-router)# **neighbo**r *ip-address* **remote-as** *pe-asn*	Establishes an EBGP session between PE and CE.
FS(config-router)**# redistribute ospf** *ospf-id*	Redistributes OSPF routes over BGP.
Ruiije(config-router)# **exit**	Exits BGP configuration mode.
FS(config)# **router ospf** *ospf-id*	Enters OSPF configuration mode.
Ruiije(config-router)# **redistribute bgp subnets**	Redistributes BGP routes over OSPF.
Ruiije(config-router)# **exit**	Exits OSPF configuration mode.

Users' VPNs may run any dynamic protocols or use static routes. OSPF is used as an example.

Method 2: Implement OSPF Between PE and CE

If users' VPNs run OSPF, connect PEs and CEs over OSPF, and no additional configuration is required on the CEs.

On the PEs, configure an OSPF instance for each VRF table connected to a CE, and use the OSPF instance to exchange routes with the CE. By redistributing BGP routes, OSPF transmits VPN routes received from the peer PE to the corresponding CE. By redistributing OSPF routes, BGP transmits to the peer PE the VPN routes that CEs send to the connected PE.

OSPF must be enabled on the CEs' interfaces connected to the PEs, and no passive port (or other restrictions on route exchange) must be configured.

Perform configuration on the PEs according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router ospf** *ospf-id vrf-name*	Creates an OSPF instance and enters OSPF configuration mode.
FS(config-router)# **network** *prefix mask* **area** *area-id*	Configures OSPF links.
Ruiije(config-router)# **redistribute bgp subnets**	Redistributes BGP routes over OSPF.
Ruiije(config-router)# **exit**	Exits OSPF configuration mode.
FS(config)# **router bgp** *asn*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **address-family ipv4 vrf** *vrf-name*	Enters BGP VRF configuration mode.
FS(config-router-af)# **redistribute ospf** *ospf-id*	Redistributes OSPF routes.
FS(config-router-af)# **show running-config**	Displays existing configurations.

Method 3: Implement RIP Between PE and CE

If users' VPNs run RIP, connect PEs and CEs over RIP, and no additional configuration is required on the CEs.

The VRF tables on PEs exchange routes with the CEs over RIP. By redistributing BGP routes, RIP transmits VPN routes received from the peer PE to the corresponding CE. By redistributing RIP routes, BGP transmits to the peer PE the VPN routes that CEs send to the connected PE.

RIP must be enabled on the CEs' interfaces connected to the PEs, and no restrictions on route exchange must be configured.

Perform configuration on the PEs according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router rip**	Creates an RIP instance and enters RIP configuration mode.
Ruiije(config-router)# **address-family ipv4 vrf** *vrf-name*	Enters RIP VRF address family configuration mode.
FS(config-router-af)# **version 2**	Configures the RIP version number.
Ruiije(config-router-af)# **network** *network-number* [*wildcard*]	Configures RIP for CE-PE communication.
Ruiije(config-router-af)# **redistribute bgp**	Redistributes BGP routes over RIP.
Ruiije(config-router-af)# **exit**	Exits address family configuration mode.
FS(config)# **router bgp** *asn*	Configures BGP and enters BGP configuration mode.

FS(config-router)# **address-family ipv4 vrf** *vrf-name*	Enters BGP VRF configuration mode.
FS(config-router-af)# **redistribute rip**	Redistributes RIP routes.
FS(config-router-af)# **show running-config**	Displays existing configurations.

Method 4: Configure Static Routes Between PE and CE

Static routes are typically used in simple network environments.

Point static routes from PE to CE, and redistribute the static routes to BGP VRF tables. Point static routes from CE to PE, and introduce the static routes to users' VPNs.

Perform configuration on the PEs according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **ip route vrf** *vrf-name* **prefix** *mask interface-name nexthop*	Configures static routes.
FS(config)# r**outer bgp** *asn*	Enters BGP configuration mode.
FS(config-router)# **address-family ipv4 vrf** *vrf-name*	Enters BGP VRF address family configuration mode.
FS(config-router-af)# **redistribute static**	Redistributes static routes.
FS(config-router-af)# **show running-config**	Displays existing configurations.

The configuration on the CEs is omitted.

(Optional) Configuring a Method of VPN Label Distribution

The RFC4364 describes two label distribution methods for L3VPN: per-route label distribution and per-VRF label distribution. Per-route label distribution enables fast forwarding, whereby packets are forwarded to the next hop just by querying an ILM table. However, this method requires the maintenance of large-capacity ILM tables. Per-VRF label distribution allows all routes of a VRF table to share a label distributed to the VRF table, thus greatly reducing the ILM table capacity. However, this method is less efficient at forwarding packets than per-route label distribution, because the former requires two table queries. The first query gets the VRF table to which the packet belongs from the ILM table, and the second query gets the packet forwarding route from the VRF table based on the destination IP address.

By default, L3VPN adopts per-VRF label distribution. You can run the **alloc-label** command in VRF configuration mode to change the default label distribution method, and you can also apply different label distribution methods to different VRF tables.

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **ip vrf** *vrf-name*	Creates a VRF table and enters VRF configuration mode.
FS(config-vrf)# **alloc-label per-vrf**	Distributes a label to all routes of the VRF table. All VPN routes distributed over MP-BGP use the same label.
FS(config-vrf)# **alloc-label per-route**	Distributes a label to each route of the VRF table. Each VPN route distributed over MP-BGP uses a different label.
FS(config-vrf)# **show running-config**	Displays existing configurations.

When you change the label distribution method, MP-BGP cancels all VPN routes that have been advertised and re-advertises the routes.

(Optional) Configuring VPN Route Import/Export Policies

You need to define the import RT attribute in VRF configuration mode to specify which routes are imported to VRF tables, and define the export RT attribute to specify which RT attributes are carried in distributed routes. The import and export RT attributes take effect for all routes. In special application scenarios, you can control the import and export of VPN routes accurately based on policies.

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **ip vrf** *vrf-name*	Creates a VRF table and enters VRF configuration mode.
FS(config-vrf)# **import map** *routemap-name*	Configures a policy for importing remote VPN routes to local VPN routes based on a route map.
FS(config-vrf)# **export map** *routemap-name*	Sets the extended community attribute of the VPN routes distributed from the local end to the remote end based on a route map.
FS(config-vrf)# **show running-config**	Displays existing configurations.

The rules defined by the **Import map** command take effect following the import-specific extended community attribute defined in VRF configuration mode. That is, the VPN routes received from the remote end must have the extended community attribute defined by the **route-target import** command in VRF configuration mode before the routes are filtered based on the rules defined by the **Import map** command.

(Optional) Configuring Static L3VPN FTN and ILM

In normal cases, labels are distributed to VPN routes over MP-BGP, and public LSP tunnels are established over LDP in public networks. You can configure static LSPs for VPN label distribution and private LSP tunnel establishment.

Configuring L3VPN FTN on PEs

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **mpls static l3vpn-ftn** *vrf-name ip-address/mask* **out-label** *out-label* **remote-pe** *ip-address*	Configures a static private FTN, with the forwarding equivalence class (FEC) egress on the other PE. You need to specify a private label and the egress PE. The address of the egress PE is used for public LSP tunnel matching.
FS(config)# **mpls static l3vpn-ftn** *vrf-name fec-prefix/fec-mask* **local-forward nexthop** *interface-name nexthop-ip*	Configures a static private FTN, with the FEC egress on the local PE. You need to specify the egress for the local PE and the next hop address. (The other VRF table is typically used as the outbound interface and the next hop.) This command is used when multiple VRF tables on the local PE belong to the same VPN.

FS# **show running-config**	Displays existing configurations.

Configuring L3VPN ILM on PEs

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **mpls static ilm in-label** *in-label* **forward-action pop-l3vpn-nexthop** *vrf-name* **nexthop** *interface-name nexthop-ip-address* **fec** *ip-address/mask*	Configures L3VPN ILM on the PE. You need to specify the incoming label, outbound interface, and next hop address.
FS# **show running-config**	Displays existing configurations.

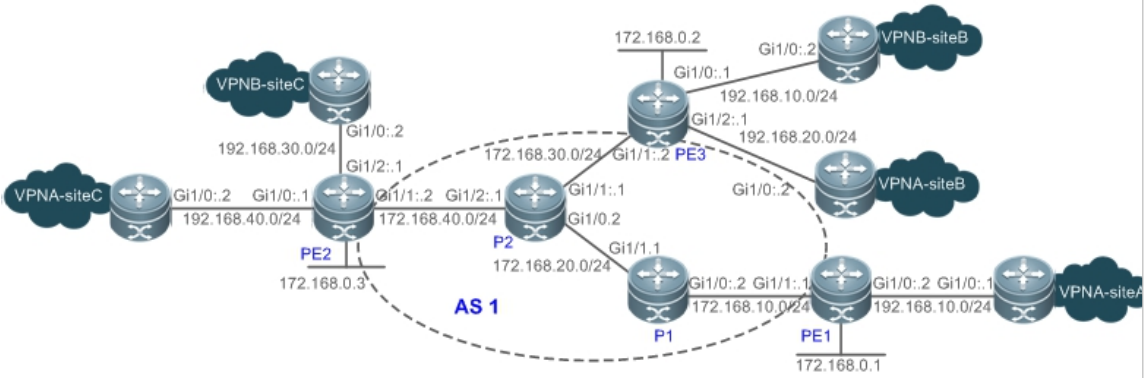
The static private FTN and ILM configurations take effect after the corresponding public LSP tunnel is established. For how to establish a public LSP tunnel, see "Configuring an MPLS Network" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions." You can establish a public LSP tunnel over LDP or based on static configuration.

Verification

- Run the **show mpls forwarding-table** command to display the corresponding MPLS entries.
- Check that VPN sites can communicate with each other.

Configuration Example

Intranet Configuration

Scenario **Figure 2- 18**	 Requirements: Two VPNs (VPNA and VPNB) exist. VPNA maintains sites at SiteA, SiteB, and SiteC, whereas VPNB maintains sites at SiteB and SiteC. Enable communication between users at different sites of VPNA and between users at different sites of VPNB, and disable communication between VPNA and VPNB.
Configuration Steps	• On PE1, configure a loopback interface, create a VRF table, define RD and RT values, and associate VPNA with the corresponding interface. Configure BGP, and establish an MP-IBGP session with PE2 and PE3 respectively. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. • On PE2 and PE3, configure a loopback interface, create two VRF tables named VPNA and VPNB, define RD and RT values, and associate the VRF tables with the corresponding interfaces. Configure BGP, and establish an MP-IBGP

	session with the neighboring PE. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● Configure VPNA-SiteA to establish an EBGP session with PE1. Configure VPNA-SiteB to establish an EBGP session with PE3. Configure VPNA-SiteC to establish an EBGP session with PE2. Configure VPNB-SiteB to establish an EBGP session with PE3. Configure VPNB-SiteC to establish an EBGP session with PE2. ● On P1 and P2, configure a loopback interface, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol.
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-if-Loopback 0)# ip address 172.168.0.1 255.255.255.255 PE1# configure terminal PE1(config)# ip vrf VPNA PE1(config-vrf)# rd 1:100 PE1(config-vrf)# route-target both 1:100 PE1(config-vrf)# exit PE1(config)# interface GigabitEthernet 1/0 PE1(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPNA PE1(config-if-GigabitEthernet 1/0)# ip address 192.168.10.2 255.255.255.0 PE1(config-if-GigabitEthernet 1/0)# exit PE1(config)# router bgp 1 PE1(config-router)# neighbor 172.168.0.2 remote-as 1 PE1(config-router)# neighbor 172.168.0.2 update-source loopback 0 PE1(config-router)# neighbor 172.168.0.3 remote-as 1 PE1(config-router)# neighbor 172.168.0.3 update-source loopback 0 PE1(config-router)# address-family vpnv4 PE1(config-router-af)# neighbor 172.168.0.2 activate PE1(config-router-af)# neighbor 172.168.0.3 activate PE1(config-router-af)# exit PE1(config)# router bgp 1 PE1(config)# address-family ipv4 vrf VPNA PE1(config-router-af)# neighbor 192.168.10.1 remote-as 65002 PE1(config-router-af)# exit PE1(config)# mpls ip

	PE1(config)# mpls router ldp PE1(config-mpls-router)# ldp router-id interface loopback 0 force PE1(config-mpls-router)# exit PE1(config)# interface GigabitEthernet 1/1 PE1(config-if-GigabitEthernet 1/1)# ip address 172.168.10.1 255.255.255.0 PE1(config-if-GigabitEthernet 1/1)# label-switching PE1(config-if-GigabitEthernet 1/1)# mpls ip PE1(config-if-GigabitEthernet 1/1)# exit PE1(config)# router ospf 10 PE1(config-router)# network 172.168.10.0 0.0.0.255 area 0 PE1(config-router)# network 172.168.0.1 0.0.0.0 area 0 PE1(config-router)# end
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-if-Loopback 0)# ip address 172.168.0.3 255.255.255.255 PE2# configure terminal PE2(config)# ip vrf VPNA PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target both 1:100 PE2(config-vrf)# exit PE2(config)# ip vrf VPNB PE2(config-vrf)# rd 1:200 PE2(config-vrf)# route-target both 1:200 PE2(config-vrf)# exit PE2# configure terminal PE2(config)# interface GigabitEthernet 1/0 PE2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPNA PE2(config-if-GigabitEthernet 1/0)# ip address 192.168.40.1 255.255.255.0 PE2(config-if-GigabitEthernet 1/0)# exit PE2(config)# interface GigabitEthernet 1/2 PE2(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPNB

```
PE2(config-if-GigabitEthernet 1/2)# ip address 192.168.30.1 255.255.255.0
PE2(config-if-GigabitEthernet 1/2)# exit
PE2# configure terminal
PE2(config)# router bgp 1
PE2(config-router)# neighbor 172.168.0.1 remote-as 1
PE2(config-router)# neighbor 172.168.0.1 update-source loopback 0
PE2(config-router)# neighbor 172.168.0.2 remote-as 1
PE2(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE2(config-router)# address-family vpnv4
PE2(config-router-af)# neighbor 172.168.0.1 activate
PE2(config-router-af)# neighbor 172.168.0.2 activate
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# router bgp 1
PE2(config)# address-family ipv4 vrf VPNA
PE2(config-router-af)# neighbor 192.168.40.2 remote-as 65003
PE2(config-router-af)# exit
PE2(config)# address-family ipv4 vrf VPNB
PE2(config-router-af)# neighbor 192.168.30.2 remote-as 65004
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# mpls ip
PE2(config)# mpls router ldp
PE2(config-mpls-router)# ldp router-id interface loopback 0 force
PE2(config-mpls-router)# exit
PE2(config)# interface GigabitEthernet 1/1
PE2(config-if-GigabitEthernet 1/1)# ip address 172.168.40.2 255.255.255.0
PE2(config-if-GigabitEthernet 1/1)# label-switching
PE2(config-if-GigabitEthernet 1/1)# mpls ip
PE2(config-if-GigabitEthernet 1/1)# end
PE2# configure terminal
PE2(config)# router ospf 10
```

	PE2(config-router)# network 172.168.40.0 255.255.255.0 area 0 PE2(config-router)# network 172.168.0.3 255.255.255.255 area 0 PE2(config-router)# end
PE3	PE3# configure terminal PE3(config)# interface loopback 0 PE3(config-if-Loopback 0)# ip address 172.168.0.2 255.255.255.255 PE3# configure terminal PE3(config)# ip vrf VPNA PE3(config-vrf)# rd 1:100 PE3(config-vrf)# route-target both 1:100 PE3(config-vrf)# exit PE3(config)# ip vrf VPNB PE3(config-vrf)# rd 1:200 PE3(config-vrf)# route-target both 1:200 PE3(config-vrf)# exit PE3# configure terminal PE3(config)# interface GigabitEthernet 1/0 PE3(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPNB PE3(config-if-GigabitEthernet 1/0)# ip address 192.168.10.1 255.255.255.0 PE3(config-if-GigabitEthernet 1/0)# exit PE3(config)# interface GigabitEthernet 1/2 PE3(config-if-GigabitEthernet 1/2)#ip vrf forwarding VPNA PE3(config-if-GigabitEthernet 1/2)#ip address 192.168.20.1 255.255.255.0 PE3(config-if-GigabitEthernet 1/2)# exit PE3# configure terminal PE3(config)# router bgp 1 PE3(config-router)# neighbor 172.168.0.1 remote-as 1 PE3(config-router)# neighbor 172.168.0.1 update-source loopback 0 PE3(config-router)# neighbor 172.168.0.3 remote-as 1 PE3(config-router)# neighbor 172.168.0.3 update-source loopback 0 PE3(config-router)# address-family vpnv4

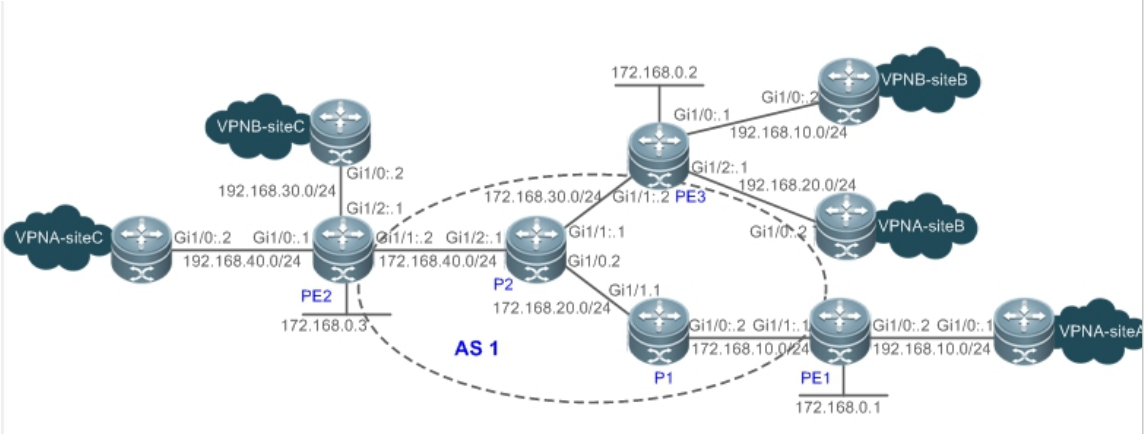
	PE3(config-router-af)# neighbor 172.168.0.1 activate PE3(config-router-af)# neighbor 172.168.0.3 activate PE3(config-router-af)# end PE3# configure terminal PE3(config)# router bgp 1 PE3(config)# address-family ipv4 vrf VPNA PE3(config-router-af)# neighbor 192.168.20.2 remote-as 65003 PE3(config-router-af)# exit PE3(config)# address-family ipv4 vrf VPNB PE3(config-router-af)# neighbor 192.168.10.2 remote-as 65004 PE3(config-router-af)# end PE3# configure terminal PE3(config)# mpls ip PE3(config)# mpls router ldp PE3(config-mpls-router)# ldp router-id interface loopback 0 force PE3(config-mpls-router)# exit PE3(config)# interface GigabitEthernet 1/1 PE3(config-if-GigabitEthernet 1/1)# ip address 172.168.30.2 255.255.255.0 PE3(config-if-GigabitEthernet 1/1)# label-switching PE3(config-if-GigabitEthernet 1/1)# mpls ip PE3(config-if-GigabitEthernet 1/1)# end PE3# configure terminal PE3(config)# router ospf 10 PE3(config-router)# network 172.168.30.0 255.255.255.0 area 0 PE3(config-router)# network 172.168.0.2 255.255.255.255 area 0 PE3(config-router)# end
VPNA-SiteA	VPNA-SITEA# configure terminal VPNA-SITEA(config)# interface GigabitEthernet 1/0 VPNA-SITEA(config-if-GigabitEthernet 1/0) ip address 192.168.10.1 255.255.255.0 VPNA-SITEA(config-if-GigabitEthernet 1/0) exit VPNA-SITEA(config)# router bgp 65002

	VPNA-SITEA(config-router)# neighbor 192.168.10.2 remote-as 1 VPNA-SITEA(config-router-af)# end
VPNA-SiteB	VPNA-SITEB# configure terminal VPNA-SITEB(config)# interface GigabitEthernet 1/0 VPNA-SITEB(config-if-GigabitEthernet 1/0) ip address 192.168.20.2 255.255.255.0 VPNA-SITEB(config-if-GigabitEthernet 1/0) exit VPNA-SITEB(config)# router bgp 65003 VPNA-SITEB(config-router)# neighbor 192.168.20.1 remote-as 1 VPNA-SITEB(config-router-af)# end
VPNA-SiteC	VPNA-SITEC# configure terminal VPNA-SITEC(config)# interface GigabitEthernet 1/0 VPNA-SITEC(config-if-GigabitEthernet 1/0) ip address 192.168.40.2 255.255.255.0 VPNA-SITEC(config-if-GigabitEthernet 1/0) exit VPNA-SITEC(config)# router bgp 65003 VPNA-SITEC(config-router)# neighbor 192.168.40.1 remote-as 1 VPNA-SITEC(config-router-af)# end
VPNB-SiteB	VPNB-SITEB# configure terminal VPNB-SITEB(config)# interface GigabitEthernet 1/0 VPNB-SITEB(config-if-GigabitEthernet 1/0) ip address 192.168.10.2 255.255.255.0 VPNB-SITEB(config-if-GigabitEthernet 1/0) exit VPNB-SITEB(config)# router bgp 65004 VPNB-SITEB(config-router)# neighbor 192.168.10.1 remote-as 1 VPNB-SITEB(config-router-af)# end
VPNB-SiteC	VPNB-SITEC# configure terminal VPNB-SITEC(config)# interface GigabitEthernet 1/0 VPNB-SITEC(config-if-GigabitEthernet 1/0) ip address 192.168.30.2 255.255.255.0 VPNB-SITEC(config-if-GigabitEthernet 1/0) exit VPNB-SITEC(config)# router bgp 65004 VPNB-SITEC(config-router)# neighbor 192.168.30.1 remote-as 1 VPNB-SITEC(config-router-af)# end

P1	P1# configure terminal P1(config)# interface loopback 0 P1(config-if-Loopback 0)# ip address 172.168.0.4 255.255.255.255 P1# configure terminal P1(config)# mpls ip P1(config)# mpls router ldp P1(config-mpls-router)# ldp router-id interface loopback 0 force P1(config-mpls-router)# exit P1(config)# interface GigabitEthernet 1/0 P1(config-if-GigabitEthernet 1/0)# ip address 172.168.10.2 255.255.255.0 P1(config-if-GigabitEthernet 1/0)# label-switching P1(config-if-GigabitEthernet 1/0)# mpls ip P1(config-if-GigabitEthernet 1/0)# exit P1(config)# interface GigabitEthernet 1/1 P1(config-if-GigabitEthernet 1/1)# ip address 172.168.20.1 255.255.255.0 P1(config-if-GigabitEthernet 1/1)# label-switching P1(config-if-GigabitEthernet 1/1)# mpls ip P1(config-if-GigabitEthernet 1/0)# end P1# configure terminal P1(config)# router ospf 10 P1(config-router)# network 172.168.10.0 0.0.0.255 area 0 P1(config-router)# network 172.168.20.0 0.0.0.255 area 0 P1(config-router)# end
P2	P1# configure terminal P1(config)# interface loopback 0 P1(config-if-Loopback 0)# ip address 172.168.0.4 255.255.255.255 P1# configure terminal P1(config)# mpls ip P1(config)# mpls router ldp P1(config-mpls-router)# ldp router-id interface loopback 0 force P1(config-mpls-router)# exit

	P1(config)# interface GigabitEthernet 1/0 P1(config-if-GigabitEthernet 1/0)# ip address 172.168.10.2 255.255.255.0 P1(config-if-GigabitEthernet 1/0)# label-switching P1(config-if-GigabitEthernet 1/0)# mpls ip P1(config-if-GigabitEthernet 1/0)# exit P1(config)# interface GigabitEthernet 1/1 P1(config-if-GigabitEthernet 1/1)# ip address 172.168.20.1 255.255.255.0 P1(config-if-GigabitEthernet 1/1)# label-switching P1(config-if-GigabitEthernet 1/1)# mpls ip P1(config-if-GigabitEthernet 1/0)# end P1# configure terminal P1(config)# router ospf 10 P1(config-router)# network 172.168.10.0 0.0.0.255 area 0 P1(config-router)# network 172.168.20.0 0.0.0.255 area 0 P1(config-router)# end
Verification	● Run the ping command for connection checking.
PE1	PE2, PE3, P1, P2, and VPNA-SiteA are pingable.
PE2	PE1, PE3, P1, P2, VPNA-SiteC, and VPNB-SiteC are pingable.
PE3	PE1, PE2, P1, P2, VPNA-SiteB, and VPNB-SiteB are pingable.
VPNA-SiteA	VPNA-SiteB and VPNA-SiteC are pingable, whereas VPNB-SiteB and VPNB-SiteC are unpingable.
VPNA-SiteB	VPNA-SiteA and VPNA-SiteC are pingable, whereas VPNB-SiteB and VPNB-SiteC are unpingable.
VPNA-SiteC	VPNA-SiteA and VPNA-SiteB are pingable, whereas VPNB-SiteB and VPNB-SiteC are unpingable.
VPNB-SiteB	VPNB-SiteC is pingable, whereas VPNA-SiteA, VPNA-SiteB and VPNA-SiteC are unpingable.
VPNB-SiteC	VPNB-SiteB is pingable, whereas VPNA-SiteA, VPNA-SiteB and VPNA-SiteC are unpingable.
P1	PE1, PE2, PE3, and P2 are pingable, whereas VPN sites are unpingable.
P2	PE1, PE2, PE3, and P1 are pingable, whereas VPN sites are unpingable.

Extranet Configuration

Scenario **Figure 2- 19**	 Requirements: Two VPNs (VPNA and VPNB) exist. Enable communication between users at different sites of VPNA and between users at different sites of VPNB, disable communication between VPNA and VPNB, and allow VPNA and VPNB to access partial shared resources. In Figure 2- , VPNA and VPNB need to access the resources of VPN-SiteA.
Configuration Steps	● On PE1, configure a loopback interface, create a VRF table named VPN_EXTRA, define RD and RT values, and associate the VRF table with the corresponding interface. Configure BGP, and establish an MP-IBGP session with PE2 and PE3 respectively. Configure the capability of exchanging routes with the CE over OSPF. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On PE2 and PE3, configure a loopback interface, create two VRF tables named VPNA and VPNB, define RD and RT values, and associate the VRF tables with the corresponding interfaces. Configure BGP, and establish an MP-IBGP session with PE3 or PE2. Configure the capability of exchanging VPN routes with the CEs over OSPF. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● Configure OSPF on VPNA-SiteA, VPNA-SiteB, VPNA-SiteC, VPNB-SiteB, and VPNB-SiteC for exchanging routes with PEs. ● On P1 and P2, configure a loopback interface, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol.
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-if-Loopback 0)# ip address 172.168.0.1 255.255.255.255 PE1# configure terminal PE1(config)# ip vrf VPN_EXTRA PE1(config-vrf)# rd 1:100 PE1(config-vrf)# route-target both 1:100 PE1(config-vrf)# route-target both 1:200 PE1(config-vrf)# end

```
PE1# configure terminal
PE1(config)# interface GigabitEthernet 1/0
PE1(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPN_EXTRA
PE1(config-if-GigabitEthernet 1/0)# ip address 192.168.100.2 255.255.255.0
PE1(config-if-GigabitEthernet 1/0)# end
PE1# configure terminal
PE1(config)# router bgp 1
PE1(config-router)# neighbor 172.168.0.2 remote-as 1
PE1(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE1(config-router)# neighbor 172.168.0.3 remote-as 1
PE1(config-router)# neighbor 172.168.0.3 update-source loopback 0
PE1(config-router)# address-family vpnv4
PE1(config-router-af)# neighbor 172.168.0.2 activate
PE1(config-router-af)# neighbor 172.168.0.3 activate
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# router ospf 10 vrf VPN_EXTRA
PE1(config-router)# network 192.168.100.0 255.255.255.0 area 0
PE1(config-router)# redistribute bgp subnets
PE1(config-router)# exit
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf VPN_EXTRA
PE1(config-router-af)# redistribute ospf 10
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# mpls ip
PE1(config)# mpls router ldp
PE1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1(config-mpls-router)# exit
PE1(config)# interface GigabitEthernet 1/1
PE1(config-if-GigabitEthernet 1/1)# ip address 172.168.10.1 255.255.255.0
PE1(config-if-GigabitEthernet 1/1)# label-switching
```

	PE1(config-if-GigabitEthernet 1/1)# mpls ip PE1(config-if-GigabitEthernet 1/1)# end PE1# configure terminal PE1(config)# router ospf 1 PE1(config-router)# network 172.168.10.0 0.0.0.255 area 0 PE1(config-router)# network 172.168.0.1 0.0.0.0 area 0 PE1(config-router)# end
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-if-Loopback 0)# ip address 172.168.0.3 255.255.255.255 PE2# configure terminal PE2(config)# ip vrf VPNA PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target both 1:100 PE2(config-vrf)# exit PE2(config)# ip vrf VPNB PE2(config-vrf)# rd 1:200 PE2(config-vrf)# route-target both 1:200 PE2(config-vrf)# exit PE2# configure terminal PE2(config)# interface GigabitEthernet 1/0 PE2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPNA PE2(config-if-GigabitEthernet 1/0)# ip address 192.168.40.1 255.255.255.0 PE2(config-if-GigabitEthernet 1/0)# exit PE2(config)# interface GigabitEthernet 1/2 PE2(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPNB PE2(config-if-GigabitEthernet 1/2)# ip address 192.168.30.1 255.255.255.0 PE2(config-if-GigabitEthernet 1/2)# exit PE2# configure terminal PE2(config)# router bgp 1 PE2(config-router)# neighbor 172.168.0.1 remote-as 1

```
PE2(config-router)# neighbor 172.168.0.1 update-source loopback 0
PE2(config-router)# neighbor 172.168.0.2 remote-as 1
PE2(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE2(config-router)# address-family vpnv4
PE2(config-router-af)# neighbor 172.168.0.1 activate
PE2(config-router-af)# neighbor 172.168.0.2 activate
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# router ospf 10 vrf VPNA
PE2(config-router)# network 192.168.40.0 255.255.255.0 area 0
PE2(config-router)# redistribute bgp subnets
PE2(config-router)# exit
PE2(config)# router ospf 20 vrf VPNB
PE2(config-router)# network 192.168.30.0 255.255.255.0 area 0
PE2(config-router)# redistribute bgp subnets
PE2(config-router)# exit
PE2(config)# router bgp 1
PE2(config-router)# address-family ipv4 vrf VPNA
PE2(config-router-af)# redistribute ospf 10
PE2(config-router-af)# exit
PE2(config-router)# address-family ipv4 vrf VPNB
PE2(config-router-af)# redistribute ospf 20
PE2(config-router-af)# exit
PE2# configure terminal
PE2(config)# mpls ip
PE2(config)# mpls router ldp
PE2(config-mpls-router)# ldp router-id interface loopback 0 force
PE2(config-mpls-router)# exit
PE2(config)# interface GigabitEthernet 1/1
PE2(config-if-GigabitEthernet 1/1)# ip address 172.168.40.2 255.255.255.0
PE2(config-if-GigabitEthernet 1/1)# label-switching
PE2(config-if-GigabitEthernet 1/1)# mpls ip
```

	PE2(config-if-GigabitEthernet 1/1)# end PE2# configure terminal PE2(config)# router ospf 1 PE2(config-router)# network 172.168.40.0 0.0.0.255 area 0 PE2(config-router)# network 172.168.0.3 0.0.0.0 area 0 PE2(config-router)# end
PE3	PE3# configure terminal PE3(config)# interface loopback 0 PE3(config-if-Loopback 0)# ip address 172.168.0.2 255.255.255.255 PE3# configure terminal PE3(config)# ip vrf VPNA PE3(config-vrf)# rd 1:100 PE3(config-vrf)# route-target both 1:100 PE3(config-vrf)# exit PE3(config)# ip vrf VPNB PE3(config-vrf)# rd 1:200 PE3(config-vrf)# route-target both 1:200 PE3(config-vrf)# exit PE3# configure terminal PE3(config)# interface GigabitEthernet 1/0 PE3(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPNB PE3(config-if-GigabitEthernet 1/0)# ip address 192.168.10.1 255.255.255.0 PE3(config-if-GigabitEthernet 1/0)# exit PE3(config)# interface GigabitEthernet 1/2 PE3(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPNA PE3(config-if-GigabitEthernet 1/2)# ip address 192.168.20.1 255.255.255.0 PE3(config-if-GigabitEthernet 1/2)# exit PE3# configure terminal PE3(config)# router bgp 1 PE3(config-router)# neighbor 172.168.0.1 remote-as 1 PE3(config-router)# neighbor 172.168.0.1 update-source loopback 0

```
PE3(config-router)# neighbor 172.168.0.3 remote-as 1
PE3(config-router)# neighbor 172.168.0.3 update-source loopback 0
PE3(config-router)# address-family vpnv4
PE3(config-router-af)# neighbor 172.168.0.1 activate
PE3(config-router-af)# neighbor 172.168.0.3 activate
PE3(config-router-af)# end
PE3# configure terminal
PE3(config)# router ospf 10 vrf VPNA
PE3(config-router)# network 192.168.20.0 255.255.255.0 area 0
PE3(config-router)# redistribute bgp subnets
PE3(config-router)# exit
PE3(config)# router ospf 20 vrf VPNB
PE3(config-router)# network 192.168.10.0 255.255.255.0 area 0
PE3(config-router)# redistribute bgp subnets
PE3(config-router)# exit
PE3(config)# router bgp 1
PE3(config-router)# address-family ipv4 vrf VPNA
PE3(config-router-af)# redistribute ospf 10
PE3(config-router-af)# exit
PE3(config-router)# address-family ipv4 vrf VPNB
PE3(config-router-af)# redistribute ospf 20
PE3(config-router-af)# exit
PE3# configure terminal
PE3(config)# mpls ip
PE3(config)# mpls router ldp
PE3(config-mpls-router)# ldp router-id interface loopback 0 force
PE3(config-mpls-router)# exit
PE3(config)# interface GigabitEthernet 1/1
PE3(config-if-GigabitEthernet 1/1)# ip address 172.168.30.2 255.255.255.0
PE3(config-if-GigabitEthernet 1/1)# label-switching
PE3(config-if-GigabitEthernet 1/1)# mpls ip
PE3(config-if-GigabitEthernet 1/1)# end
```

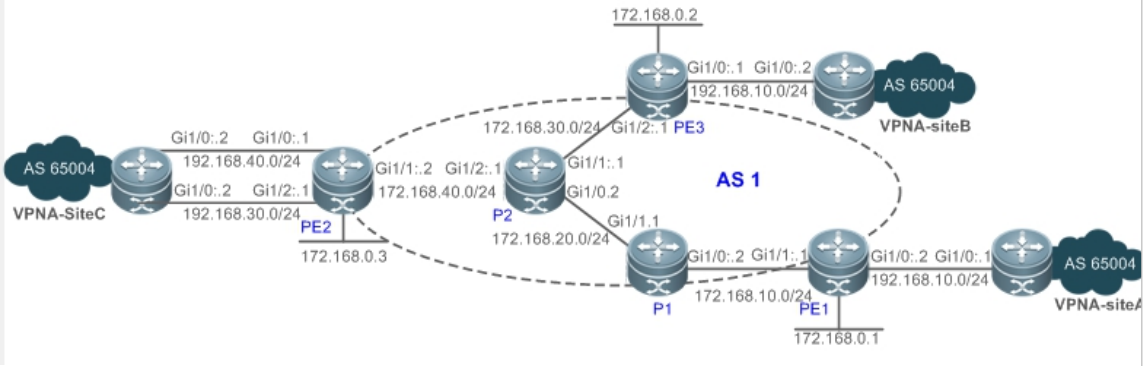
	PE3# configure terminal PE3(config)# router ospf 1 PE3(config-router)# network 172.168.30.0 0.0.0.255 area 0 PE3(config-router)# network 172.168.0.2 0.0.0.0 area 0 PE3(config-router)# end
VPNA-SiteA	VPNA-SITEA# configure terminal VPNA-SITEA(config)# interface GigabitEthernet 1/0 VPNA-SITEA(config-if-GigabitEthernet 1/0)# ip address 192.168.100.1 255.255.255.0 VPNA-SITEA(config-if-GigabitEthernet 1/0)# exit VPNA-SITEA(config)# router ospf 10 VPNA-SITEA(config-router)# network 192.168.100.0 255.255.255.0 area 0 VPNA-SITEA(config-router)# end
VPNA-SiteB	VPNA-SITEB# configure terminal VPNA-SITEB(config)# interface GigabitEthernet 1/0 VPNA-SITEB(config-if-GigabitEthernet 1/0)# ip address 192.168.20.2 255.255.255.0 VPNA-SITEB(config-if-GigabitEthernet 1/0)# exit VPNA-SITEB(config)# router ospf 10 VPNA-SITEB(config-router)# network 192.168.20.0 255.255.255.0 area 0 VPNA-SITEB(config-router)# end
VPNA-SiteC	VPNA-SITEC# configure terminal VPNA-SITEC(config)# interface GigabitEthernet 1/0 VPNA-SITEC(config-if-GigabitEthernet 1/0)# ip address 192.168.40.2 255.255.255.0 VPNA-SITEC(config-if-GigabitEthernet 1/0)# exit VPNA-SITEC(config)# router ospf 10 VPNA-SITEC(config-router)# network 192.168.40.0 255.255.255.0 area 0 VPNA-SITEC(config-router)# end
VPNB-SiteB	VPNB-SITEB# configure terminal VPNB-SITEB(config)# interface GigabitEthernet 1/0 VPNB-SITEB(config-if-GigabitEthernet 1/0)# ip address 192.168.10.2 255.255.255.0 VPNB-SITEB(config-if-GigabitEthernet 1/0)# exit VPNB-SITEB(config)# router ospf 10

	VPNB-SITEB(config-router)# network 192.168.10.0 255.255.255.0 area 0 VPNB-SITEB(config-router)# end
VPNB-SiteC	VPNB-SITEC# configure terminal VPNB-SITEC(config)# interface GigabitEthernet 1/0 VPNB-SITEC(config-if-GigabitEthernet 1/0)# ip address 192.168.30.2 255.255.255.0 VPNB-SITEC(config-if-GigabitEthernet 1/0)# exit VPNB-SITEC(config)# router ospf 10 VPNB-SITEC(config-router)# network 192.168.30.0 255.255.255.0 area 0 VPNB-SITEC(config-router)# end
P1	P1# configure terminal P1(config)# interface loopback 0 P1(config-if-Loopback 0)# ip address 172.168.0.4 255.255.255.255 P1# configure terminal P1(config)# mpls ip P1(config)# mpls router ldp P1(config-mpls-router)# ldp router-id interface loopback 0 force P1(config-mpls-router)# exit P1(config)# interface GigabitEthernet 1/0 P1(config-if-GigabitEthernet 1/0)# ip address 172.168.10.2 255.255.255.0 P1(config-if-GigabitEthernet 1/0)# label-switching P1(config-if-GigabitEthernet 1/0)# mpls ip P1(config-if-GigabitEthernet 1/0)# exit P1(config)# interface GigabitEthernet 1/1 P1(config-if-GigabitEthernet 1/1)# ip address 172.168.20.1 255.255.255.0 P1(config-if-GigabitEthernet 1/1)# label-switching P1(config-if-GigabitEthernet 1/1)# mpls ip P1(config-if-GigabitEthernet 1/0)# end P1# configure terminal P1(config)# router ospf 1 P1(config-router)# network 172.168.10.0 0.0.0.255 area 0 P1(config-router)# network 172.168.20.0 0.0.0.255 area 0

	P1(config-router)# end
P2	P2# configure terminal P2(config)# interface loopback 0 P2(config-if-Loopback 0)# ip address 172.168.0.5 255.255.255.255 P2# configure terminal P2(config)# mpls ip P2(config)# mpls router ldp P2(config-mpls-router)# ldp router-id interface loopback 0 force P2(config-mpls-router)# exit P2(config)# interface GigabitEthernet 1/0 P2(config-if-GigabitEthernet 1/0)# ip address 172.168.20.2 255.255.255.0 P2(config-if-GigabitEthernet 1/0)# label-switching P2(config-if-GigabitEthernet 1/0)# mpls ip P2(config-if-GigabitEthernet 1/0)# exit P2(config)# interface GigabitEthernet 1/1 P2(config-if-GigabitEthernet 1/1)# ip address 172.168.30.1 255.255.255.0 P2(config-if-GigabitEthernet 1/1)# label-switching P2(config-if-GigabitEthernet 1/1)# mpls ip P2(config-if-GigabitEthernet 1/1)# exit P2(config)# interface GigabitEthernet 1/2 P2(config-if-GigabitEthernet 1/2)# ip address 172.168.40.1 255.255.255.0 P2(config-if-GigabitEthernet 1/2)# label-switching P2(config-if-GigabitEthernet 1/2)# mpls ip P2(config-if-GigabitEthernet 1/2)# end P2# configure terminal P2(config)# router ospf 1 P2(config-router)# network 172.168.20.0 0.0.0.255 area 0 P2(config-router)# network 172.168.30.0 0.0.0.255 area 0 P2(config-router)# network 172.168.40.0 0.0.0.255 area 0 P2(config-router)# end
Verification	● Run the ping command for connection checking.

PE1	PE2, PE3, P1, P2, and VPN-SiteA are pingable.
PE2	PE1, PE3, P1, P2, VPNA-SiteC, and VPNB-SiteC are pingable.
PE3	PE1, PE2, P1, P2, VPNA-SiteB, and VPNB-SiteB are pingable.
VPN-SiteA	VPNA-SiteB, VPNA-SiteC, VPNB-SiteB, and VPN-SiteC are pingable.
VPNA-SiteB	VPN-SiteA and VPNA-SiteC are pingable, whereas VPNB-SiteB and VPNB-SiteC are unpingable.
VPNA-SiteC	VPN-SiteA and VPNA-SiteB are pingable, whereas VPNB-SiteB and VPNB-SiteC are unpingable.
VPNB-SiteB	VPN-SiteA and VPNB-SiteC are pingable, whereas VPNA-SiteB and VPNA-SiteC are unpingable.
VPNB-SiteC	VPN-SiteA and VPNB-SiteB are pingable, whereas VPNA-SiteB and VPNA-SiteC are unpingable.
P1	PE1, PE2, PE3, and P2 are pingable.
P2	PE1, PE2, PE3, and P1 are pingable.

Hub-and-Spoke Configuration

Scenario **Figure 2- 20**	 Requirements: Intra-VPN data cannot be exchanged directly, but must be exchanged over a control center. Only the control center has the right to know the resources of the VPN. Users of the VPN can access the resources only with the control center's permission. In Figure 2- , VPNA-SiteA must access the resources of VPNA-SiteB over the control center VPNA-SiteC.
Configuration Steps	● On PE1, configure a loopback interface, create a VRF table named spoke1, define RD and RT values, and associate VPNA with the corresponding interface. Configure BGP, and establish an MP-IBGP session with PE2. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On PE3, configure a loopback interface, create a VRF table named spoke2, define RD and RT values, and associate the VRF table with the corresponding interface. Configure BGP, and establish an MP-IBGP session with PE2. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the

	interface connected to the public network. Configure a backbone routing protocol. ● On PE2, configure a loopback interface, create two VRF tables named from-spoke and from-hub, define RD and RT values, and associate the VRF tables with the corresponding interfaces. Configure BGP, and establish an MP-IBGP session with PE1 and PE3 respectively. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● Configure VPNA-SiteA, VPNA-SiteB, and VPNA-SiteC to establish an EBGP session with PEs.
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-if-Loopback 0)# ip address 172.168.0.1 255.255.255.255 PE1# configure terminal PE1(config)# ip vrf spoke1 PE1(config-vrf)# rd 1:100 PE1(config-vrf)# route-target export 1:200 PE1(config-vrf)# route-target import 1:100 PE1(config-vrf)# end PE1# configure terminal PE1(config)# interface GigabitEthernet 1/0 PE1(config-if-GigabitEthernet 1/0)# ip vrf forwarding spoke1 PE1(config-if-GigabitEthernet 1/0)# ip address 192.168.100.2 255.255.255.0 PE1(config-if-GigabitEthernet 1/0)# end PE1# configure terminal PE1(config)# router bgp 1 PE1(config-router)# neighbor 172.168.0.3 remote-as 1 PE1(config-router)# neighbor 172.168.0.3 update-source loopback 0 PE1(config-router)# address-family vpnv4 PE1(config-router-af)# neighbor 172.168.0.3 activate PE1(config-router-af)# neighbor 172.168.0.3 allowas-in PE1(config-router-af)# end PE1# configure terminal PE1(config)# router bgp 1 PE1(config)# address-family ipv4 vrf spoke1 PE1(config-router-af)# neighbor 192.168.100.1 remote-as 65004

	PE1(config-router-af)# neighbor 192.168.100.1 as-override PE1(config-router-af)# end PE1# configure terminal PE1(config)# mpls ip PE1(config)# mpls router ldp PE1(config-mpls-router)# ldp router-id interface loopback 0 force PE1(config-mpls-router)# exit PE1(config)# interface GigabitEthernet 1/1 PE1(config-if-GigabitEthernet 1/1)# ip address 172.168.10.1 255.255.255.0 PE1(config-if-GigabitEthernet 1/1)# label-switching PE1(config-if-GigabitEthernet 1/1)# mpls ip PE1(config-if-GigabitEthernet 1/1)# end PE1# configure terminal PE1(config)# router ospf 10 PE1(config-router)# network 172.168.10.0 0.0.0.255 area 0 PE1(config-router)# network 172.168.0.1 0.0.0.0 area 0 PE1(config-router)# end
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-if-Loopback 0)# ip address 172.168.0.3 255.255.255.255 PE2# configure terminal PE2(config)# ip vrf from-spoke PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target import 1:300 PE2(config-vrf)# route-target import 1:200 PE2(config-vrf)# exit PE2(config)# ip vrf from-hub PE2(config-vrf)# rd 1:200 PE2(config-vrf)# route-target export 1:100 PE2(config-vrf)# exit PE2# configure terminal

```
PE2(config)# interface GigabitEthernet 1/0
PE2(config-if-GigabitEthernet 1/0)# ip vrf forwarding from-hub
PE2(config-if-GigabitEthernet 1/0)# ip address 192.168.40.1 255.255.255.0
PE2(config-if-GigabitEthernet 1/0)# exit
PE2(config)# interface GigabitEthernet 1/2
PE2(config-if-GigabitEthernet 1/2)# ip vrf forwarding from-spoke
PE2(config-if-GigabitEthernet 1/2)# ip address 192.168.30.1 255.255.255.0
PE2(config-if-GigabitEthernet 1/2)# exit
PE2# configure terminal
PE2(config)# router bgp 1
PE2(config-router)# neighbor 172.168.0.1 remote-as 1
PE2(config-router)# neighbor 172.168.0.1 update-source loopback 0
PE2(config-router)# neighbor 172.168.0.2 remote-as 1
PE2(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE2(config-router)# address-family vpnv4
PE2(config-router-af)# neighbor 172.168.0.1 activate
PE2(config-router-af)# neighbor 172.168.0.2 activate
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# router bgp 1
PE2(config-router)# address-family ipv4 vrf from-spoke
PE2(config-router-af)# neighbor 192.168.30.2 remote-as 65004
PE2(config-router-af)# neighbor 192.168.30.2 as-override
PE2(config-router-af)# exit
PE2(config-router)# address-family ipv4 vrf from-hub
PE2(config-router-af)# neighbor 192.168.40.2 remote-as 65004
PE2(config-router-af)# neighbor 192.168.40.2 allows-in
PE2(config-router-af)# exit
PE2# configure terminal
PE2(config)# mpls ip
PE2(config)# mpls router ldp
PE2(config-mpls-router)# ldp router-id interface loopback 0 force
```

	PE2(config-mpls-router)# exit PE2(config)# interface GigabitEthernet 1/1 PE2(config-if-GigabitEthernet 1/1)# ip address 172.168.40.2 255.255.255.0 PE2(config-if-GigabitEthernet 1/1)# label-switching PE2(config-if-GigabitEthernet 1/1)# mpls ip PE2(config-if-GigabitEthernet 1/1)# end PE2# configure terminal PE2(config)# router ospf 10 PE2(config-router)# network 172.168.40.0 0.0.0.255 area 0 PE2(config-router)# network 172.168.0.3 0.0.0.0 area 0 PE2(config-router)# end
PE3	PE3# configure terminal PE3(config)# interface loopback 0 PE3(config-if-Loopback 0)# ip address 172.168.0.2 255.255.255.255 PE3# configure terminal PE3(config)# ip vrf spoke2 PE3(config-vrf)# rd 1:100 PE3(config-vrf)# route-target export 1:300 PE3(config-vrf)# route-target import 1:100 PE3(config-vrf)# exit PE3# configure terminal PE3(config)# interface GigabitEthernet 1/0 PE3(config-if-GigabitEthernet 1/0)# ip vrf forwarding spoke2 PE3(config-if-GigabitEthernet 1/0)# ip address 192.168.10.1 255.255.255.0 PE3(config-if-GigabitEthernet 1/0)# exit PE3# configure terminal PE3(config)# router bgp 1 PE3(config-router)# neighbor 172.168.0.3 remote-as 1 PE3(config-router)# neighbor 172.168.0.3 update-source loopback 0 PE3(config-router)# address-family vpnv4 PE3(config-router-af)# neighbor 172.168.0.3 activate

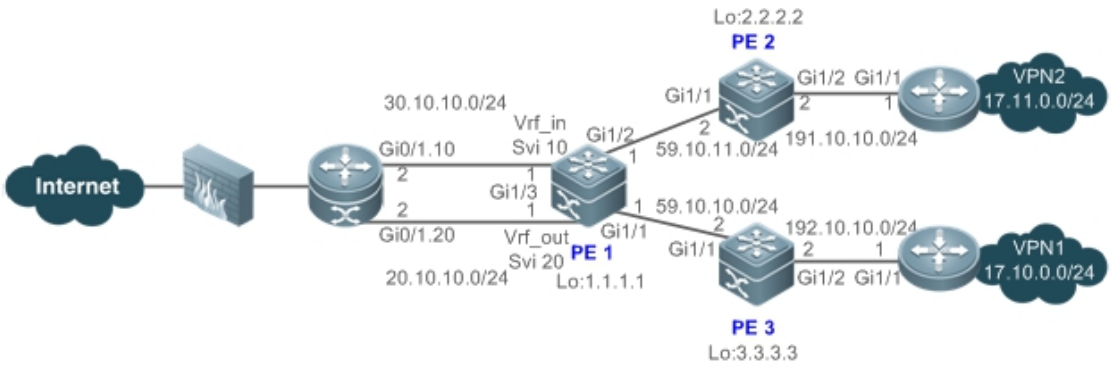
	PE3(config-router-af)# neighbor 172.168.0.3 allowas-in PE3(config-router-af)# end PE3# configure terminal PE3(config)# router bgp 1 PE3(config-router)# address-family ipv4 vrf spoke2 PE3(config-router-af)# neighbor 192.168.10.2 remote-as 65004 PE3(config-router-af)# neighbor 192.168.10.2 as-override PE3(config-router-af)# end PE3# configure terminal PE3(config)# mpls ip PE3(config)# mpls router ldp PE3(config-mpls-router)# ldp router-id interface loopback 0 force PE3(config-mpls-router)# exit PE3(config)# interface GigabitEthernet 1/2 PE3(config-if-GigabitEthernet 1/2)# ip address 172.168.30.2 255.255.255.0 PE3(config-if-GigabitEthernet 1/2)# label-switching PE3(config-if-GigabitEthernet 1/2)# mpls ip PE3(config-if-GigabitEthernet 1/2)# end PE3# configure terminal PE3(config)# router ospf 10 PE3(config-router)# network 172.168.30.0 0.0.0.255 area 0 PE3(config-router)# network 172.168.0.2 0.0.0.0 area 0 PE3(config-router)# end
VPNA-SiteA	VPNA-SITEA# configure terminal VPNA-SITEA(config)# interface GigabitEthernet 1/0 VPNA-SITEA(config-if-GigabitEthernet 1/0)# ip address 192.168.100.1 255.255.255.0 VPNA-SITEA(config-if-GigabitEthernet 1/0)# exit VPNA-SITEA(config)# router bgp 65004 VPNA-SITEA(config-router)# neighbor 192.168.100.2 remote-as 1 VPNA-SITEA(config-router)# exit
VPNA-SiteB	VPNA-SITEB# configure terminal

	VPNA-SITEB(config)# interface GigabitEthernet 1/0 VPNA-SITEB(config-if-GigabitEthernet 1/0)# ip address 192.168.10.2 255.255.255.0 VPNA-SITEB(config-if-GigabitEthernet 1/0)# exit VPNA-SITEB(config)# router bgp 65004 VPNA-SITEB(config-router)# neighbor 192.168.10.1 remote-as 1 VPNA-SITEB(config-router)# exit
VPNA-SiteC	VPNA-SITEC# configure terminal VPNA-SITEC(config)# interface GigabitEthernet 1/0 VPNA-SITEC(config-if-GigabitEthernet 1/0)# ip address 192.168.40.2 255.255.255.0 VPNA-SITEC(config-if-GigabitEthernet 1/0)# exit VPNA-SITEC(config)# interface GigabitEthernet 1/1 VPNA-SITEC(config-if-GigabitEthernet 1/1)# ip address 192.168.30.2 255.255.255.0 VPNA-SITEC(config-if-GigabitEthernet 1/1)# exit VPNA-SITEC(config)# router bgp 65004 VPNA-SITEC(config-router)# neighbor 192.168.30.1 remote-as 1 VPNA-SITEC(config-router)# neighbor 192.168.40.1 remote-as 1 VPNA-SITEC(config-router)# exit
Verification	● Run the ping command for connection checking.
PE1	PE2 and PE3 are pingable.
PE2	PE1 and PE3 are pingable.
PE3	PE1 and PE2 are pingable.
VPNA-SiteA	VPNA-SiteB is pingable. The VPNA-SiteB tracing results show that traffic traverses VPNA-SiteC.
VPNA-SiteB	VPNA-SiteA is pingable. The VPNA-SiteA tracing results show that traffic traverses VPNA-SiteC.
VPNA-SiteC	VPNA-SiteA and VPNA-SiteB are pingable.

Centralized Isolation Control at the Internet Egress

Scenario

Figure 2- 21



Requirements: Disable inter-VPN communication, and configure VPNs to access the Internet over a single device. In Figure 2- , VPN1 and VPN2 must access the Internet over PE1, but the two VPNs are not allowed to communicate with each other.

Two isolation methods are available: centralized isolation control and distributed isolation control.

In centralized isolation control, to allow new VPN sites to access the Internet over a single egress, you only need to add filter rules on the egress CE, without the need to modify the configurations of other VPN sites. This method provides good scalability. The disadvantage is that the isolated traffic is discarded only on the egress CE, which wastes network bandwidth.

Configuration Steps

- On PE2, configure a loopback interface, create a VRF table named VPN2, define RD and RT values, and associate the VRF table with the corresponding interface. Configure the capability of exchanging routes with CE2 over OSPF. Configure an IBGP neighbor relationship with PE1 for IP route distribution. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On CE2, configure the IP address of the interface connected to PE2, and configure the capability of exchanging routes with PE2 over OSPF.
- On PE3, configure a loopback interface and create a VRF table named VPN1. Assume that the export and import RT attributes of VPN1 are both 1: 200. Define RD and RT values, and associate the VRF table with the corresponding interface. Configure the capability of exchanging routes with CE3 over OSPF. Configure an IBGP neighbor relationship with PE1 for IP route distribution. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On PE1, configure a loopback interface and create a trunk interface. Create a VRF table named vrf_out, define RD and RT values, and associate the VRF table with the corresponding interface. Configure a default route destined for the Internet. Create a VRF table named vrf_in, define RD and RT values, and associate the VRF table with the corresponding interface. Configure the capability of exchanging routes with CE1 over EBGP. Configure an IBGP neighbor relationship with PE2 and PE3 respectively. Configure the route exchange function of the vrf_out VRF table. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On CE1, create sub-interfaces, configure an EBGP neighbor relationship with PE1, create an access control list (ACL), and configure ACL application on the Gi0/1.10 sub-interface.

This example uses the configuration of the switch virtual interface (SVI) and the 802.1Q sub-interface to describe how to connect PE1 and CE1. Some devices may not support the two interfaces. PE1 and CE1 can be connected

	by any two links (physical or logical) based on two routing adjacencies between them. You can select a suitable connection method according to the actual condition.
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 PE2# configure terminal PE2(config)# ip vrf VPN1 PE2(config-vrf)# rd 1:200 PE2(config-vrf)# route-target both 1:200 PE2(config-vrf)# end PE2# configure terminal PE2(config)# interface GigabitEthernet 1/2 PE2(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPN1 PE2(config-if-GigabitEthernet 1/2)# ip address 191.10.10.2 255.255.255.0 PE2(config-if-GigabitEthernet 1/2)# end PE2# configure terminal PE2(config)# router ospf 10 vrf VPN1 PE2(config-router)# network 191.10.10.0 0.0.0.255 area 0 PE2(config-router)# default-information originate PE2(config-router)# redistribute bgp subnets PE2(config-router)# exit PE2# configure terminal PE2(config)# router bgp 1 PE2(config-router)# neighbor 1.1.1.1 remote-as 1 PE2(config-router)# neighbor 1.1.1.1 update-source loopback 0 PE2(config-router)# address-family vpnv4 PE2(config-router-af)# neighbor 1.1.1.1 activate PE2(config-router-af)# end PE2(config)# router bgp 1 PE2(config-router)# address-family ipv4 vrf VPN1 PE2(config-router-af)# redistribute ospf 10

	PE2(config-router-af)# end PE2# configure terminal PE2(config)# mpls ip PE2(config)# mpls router ldp PE2(config-mpls-router)# ldp router-id interface loopback 0 force PE2(config-mpls-router)# exit PE2(config)# interface GigabitEthernet 1/1 PE2(config-if-GigabitEthernet 1/1)# ip address 59.10.11.2 255.255.255.0 PE2(config-if-GigabitEthernet 1/1)# label-switching PE2(config-if-GigabitEthernet 1/1)# mpls ip PE2(config-if-GigabitEthernet 1/1)# end PE2# configure terminal PE2(config)# router ospf 1 PE2(config-router)# network 59.10.11.0 0.0.0.255 area 0 PE2(config-router)# network 2.2.2.2 0.0.0.0 area 0 PE2(config-router)# end
CE2	CE2# configure terminal CE2(config)# interface GigabitEthernet 1/1 CE2(config-if-GigabitEthernet 1/1)# ip address 191.10.10.1 255.255.255.0 CE2# configure terminal CE2(config)# router ospf 1 CE2(config-router)# network 191.10.10.0 0.0.0.255 area 0 CE2(config-router)# network 172.11.0.0 0.0.0.255 area 0 CE2(config-router)# exit
PE3	PE3# configure terminal PE3(config)# interface loopback 0 PE3(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 PE3# configure terminal PE3(config)# ip vrf VPN1 PE3(config-vrf)# rd 1:100 PE3(config-vrf)# route-target both 1:100

```
PE3(config-vrf)# end
PE3# configure terminal
PE3(config)# interface GigabitEthernet 1/2
PE3(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPN1
PE3(config-if-GigabitEthernet 1/2)# ip address 192.10.10.2 255.255.255.0
PE3(config-if-GigabitEthernet 1/2)# end
PE3# configure terminal
PE3(config)# router ospf 10 vrf VPN1
PE3(config-router)# network 192.10.10.0 0.0.0.255 area 0
PE3(config-router)# default-information originate
PE3(config-router)# redistribute bgp subnets
PE3(config-router)# exit
PE3# configure terminal
PE3(config)# router bgp 1
PE3(config-router)# neighbor 1.1.1.1 remote-as 1
PE3(config-router)# neighbor 1.1.1.1 update-source loopback 0
PE3(config-router)# address-family vpnv4
PE3(config-router-af)# neighbor 1.1.1.1 activate
PE3(config-router-af)# end
PE3(config)# router bgp 1
PE3(config-router)# address-family ipv4 vrf VPN1
PE3(config-router-af)# redistribute ospf 10
PE3(config-router-af)# end
PE3# configure terminal
PE3(config)# mpls ip
PE3(config)# mpls router ldp
PE3(config-mpls-router)# ldp router-id interface loopback 0 force
PE3(config-mpls-router)# exit
PE3(config)# interface GigabitEthernet 1/1
PE3(config-if-GigabitEthernet 1/1)# ip address 59.10.10.2 255.255.255.0
PE3(config-if-GigabitEthernet 1/1)# label-switching
PE3(config-if-GigabitEthernet 1/1)# mpls ip
```

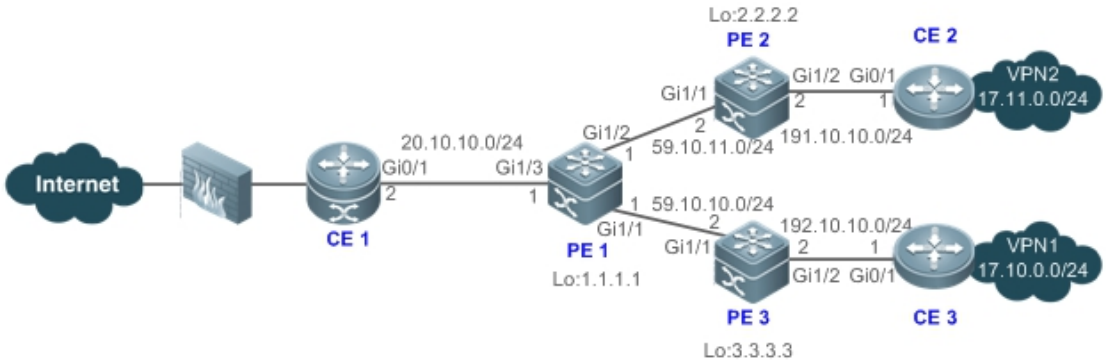
	PE3(config-if-GigabitEthernet 1/1)# end PE3# configure terminal PE3(config)# router ospf 1 PE3(config-router)# network 59.10.10.0 0.0.0.255 area 0 PE3(config-router)# network 3.3.3.3 0.0.0.0 area 0 PE3(config-router)# end CE3 CE3# configure terminal CE3(config)# interface GigabitEthernet 1/1 CE3(config-if-GigabitEthernet 1/1)# ip address 192.10.10.1 255.255.255.0 CE3# configure terminal CE3(config)# router ospf 1 CE3(config-router)# network 192.10.10.0 0.0.0.255 area 0 CE3(config-router)# network 172.10.0.0 0.0.0.255 area 0 CE3(config-router)# exit
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1# configure terminal PE1(config)# interface GigabitEthernet 1/3 PE1(config)# switch mode trunk PE1(config)# switch trunk vlan allow remove vlan 1-9,11-19,21-4094 PE1# configure terminal PE1(config)# ip vrf vrf_out PE1(config-vrf)# rd 1:300 PE1(config-vrf)# route-target export 1:100 PE1(config-vrf)# route-target export 1:200 PE1(config-vrf)# end PE1# configure terminal PE1(config)#interface vlan 20 PE1(config-if-Vlan 20)# ip vrf forwarding vrf_out

```
PE1(config-if-Vlan 20)# ip address 20.10.10.1 255.255.255.0
PE1# configure terminal
PE1(config)# ip route vrf vrf_out 0.0.0.0 0.0.0.0 vlan 20 20.10.10.2
PE1# configure terminal
PE1(config)# ip vrf vrf_in
PE1(config-vrf)# rd 1:400
PE1(config-vrf)# route-target import 1:100
PE1(config-vrf)# route-target import 1:200
PE1(config-vrf)# end
PE1# configure terminal
PE1(config)#interface vlan 10
PE1(config-if-Vlan 10)# ip vrf forwarding vrf_in
PE1(config-if-Vlan 10)# ip address 30.10.10.1 255.255.255.0
PE1# configure terminal
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf vrf_in
PE1(config-router-af)# neighbor 30.10.10.2 remote-as 100
PE1# configure terminal
PE1(config)# router bgp 1
PE1(config-router)# neighbor 2.2.2.2 remote-as 1
PE1(config-router)# neighbor 2.2.2.2 update-source loopback 0
PE1(config-router)# neighbor 3.3.3.3 remote-as 1
PE1(config-router)# neighbor 3.3.3.3 update-source loopback 0
PE1(config-router)# address-family vpnv4
PE1(config-router-af)# neighbor 2.2.2.2 activate
PE1(config-router-af)# neighbor 3.3.3.3 activate
PE1(config-router-af)# end
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf vrf_out
PE1(config-router-af)# default-information originate
PE1(config-router-af)# redistribute static
PE1(config-router-af)# end
```

	PE1# configure terminal PE1(config)# mpls ip PE1(config)# mpls router ldp PE1(config-mpls-router)# ldp router-id interface loopback 0 force PE1(config-mpls-router)# exit PE1(config)# interface GigabitEthernet 1/2 PE1(config-if-GigabitEthernet 1/2)# ip address 59.10.11.1 255.255.255.0 PE1(config-if-GigabitEthernet 1/2)# label-switching PE1(config-if-GigabitEthernet 1/2)# mpls ip PE1(config-if-GigabitEthernet 1/2)# end PE1(config)# interface GigabitEthernet 1/1 PE1(config-if-GigabitEthernet 1/1)# ip address 59.10.10.1 255.255.255.0 PE1(config-if-GigabitEthernet 1/1)# label-switching PE1(config-if-GigabitEthernet 1/1)# mpls ip PE1(config-if-GigabitEthernet 1/1)# end PE1# configure terminal PE1(config)# router ospf 1 PE1(config-router)# network 59.10.11.0 0.0.0.255 area 0 PE1(config-router)# network 59.10.10.0 0.0.0.255 area 0 PE1(config-router)# network 1.1.1.1 0.0.0.0 area 0 PE1(config-router)# end
CE1	CE1# configure terminal CE1(config)# interface GigabitEthernet 0/1.10 CE1(config-if-GigabitEthernet 0/1.10)# encapsulation dot1q 10 CE1(config-if-GigabitEthernet 0/1.10)# ip address 30.10.10.2 255.255.255.0 CE1(config)# exit CE1(config)# interface GigabitEthernet 0/1.20 CE1(config-if-GigabitEthernet 0/1.20)# encapsulation dot1q 20 CE1(config-if-GigabitEthernet 0/1.20)# ip address 20.10.10.2 255.255.255.*0* CE1# configure terminal CE1(config)# router bgp 100

	CE1(config)# neighbor 30.10.10.1 remote-as 1 CE1(config)#access-list 1 deny 17.0.0.0 0.255.255.255 CE1(config)# access-list 1 permit any CE1(config)# interface GigabitEthernet 0/1.10 CE1(config-if-GigabitEthernet 0/1.10)# ip access-group 1 out
Verification	● Run the ping command for connection checking.
PE2	PE1 and PE3 are pingable.
CE2	CE1 is pingable, whereas CE3 is unpingable.
PE3	PE1 and PE2 are pingable.
CE3	CE1 is pingable, whereas CE2 is unpingable.
PE1	PE2 and PE3 are pingable.
CE1	CE2 and CE3 are pingable.

Distributed Isolation Control at the Internet Egress

Figure 2- 22	 Requirements: Disable inter-VPN communication, and configure VPNs to access the Internet over a single device. In Figure 2- , VPN1 and VPN2 must access the Internet over PE1, but the two VPNs are not allowed to communicate with each other. Two isolation methods are available: centralized isolation control and distributed isolation control. In distributed isolation control, to allow new VPN sites to access the Internet over a single egress, you need to add filter rules on the CEs for the original VPN sites that also access the Internet over the same egress. This method has poor scalability. The advantage is that traffic is discarded on the CEs for VPN sites, which saves network bandwidth.
Configuration Steps	● On PE1, configure a loopback interface, create a VRF table named vrf_net, define RD and RT values, and associate the VRF table with the corresponding interface. Configure a default route destined for the Internet. Configure the capability of exchanging routes with CE1 over EBGP. Configure an IBGP neighbor relationship with PE2 and PE3 respectively. Configure the route exchange function of the vrf_net VRF table. Configure backbone MPLS

	signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On CE1, configure an EBGP neighbor relationship with PE1. ● On PE2, configure a loopback interface, create a VRF table named VPN2, define RD and RT values, and associate the VRF table with the corresponding interface. Configure the capability of exchanging routes with CE2 over OSPF. Configure an IBGP neighbor relationship with PE1 for IP route distribution. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On CE2, create an ACL, configure ACL application on a sub-interface, configure the capability of exchanging routes with PE2, and configure a default static route. ● On PE3, configure a loopback interface and create a VRF table named VPN1. Assume that the export and import RT attributes of VPN1 are both 1: 200. Define RD and RT values, and associate the VRF table with the corresponding interface. Configure the capability of exchanging routes with CE3 over OSPF. Configure an IBGP neighbor relationship with PE1 for IP route distribution. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On CE3, create an ACL, configure ACL application on a sub-interface, configure the capability of exchanging routes with PE3, and configure a default static route.
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1# configure terminal PE1(config)# ip vrf vrf_net PE1(config-vrf)# rd 1:300 PE1(config-vrf)# route-target import 1:100 PE1(config-vrf)# route-target import 1:200 PE1(config-vrf)# route-target export 1:100 PE1(config-vrf)# route-target export 1:200 PE1# configure terminal PE1(config)# interface GigabitEthernet 1/3 PE1(config-if-GigabitEthernet 1/3)# no switchport PE1(config-if-GigabitEthernet 1/3)# ip vrf forwarding vrf_net PE1(config-if-GigabitEthernet 1/3)# ip address 20.10.10.1 255.255.255.0 PE1(config)#ip route vrf vrf_net 0.0.0.0 0.0.0.0 GigabitEthernet 1/3 20.10.10.2 PE1# configure terminal PE1(config)# router bgp 1 PE1(config-router)# address-family ipv4 vrf vrf_net PE1(config-router-af)# neighbor 20.10.10.2 remote-as 100

```
PE1# configure terminal
PE1(config)# router bgp 1
PE1(config-router)# neighbor 2.2.2.2 remote-as 1
PE1(config-router)# neighbor 2.2.2.2 update-source loopback 0
PE1(config-router)# neighbor 3.3.3.3 remote-as 1
PE1(config-router)# neighbor 3.3.3.3 update-source loopback 0
PE1(config-router)# address-family vpnv4
PE1(config-router-af)# neighbor 2.2.2.2 activate
PE1(config-router-af)# neighbor 3.3.3.3 activate
PE1(config-router-af)# end
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf vrf_net
PE1(config-router-af)# default-information originate
PE1(config-router-af)# redistribute static
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# mpls ip
PE1(config)# mpls router ldp
PE1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1(config-mpls-router)# exit
PE1(config)# interface GigabitEthernet 1/2
PE1(config-if-GigabitEthernet 1/2)# ip address 59.10.11.1 255.255.255.0
PE1(config-if-GigabitEthernet 1/2)# label-switching
PE1(config-if-GigabitEthernet 1/2)# mpls ip
PE1(config-if-GigabitEthernet 1/2)# end
PE1(config)# interface GigabitEthernet 1/1
PE1(config-if-GigabitEthernet 1/1)# ip address 59.10.10.1 255.255.255.0
PE1(config-if-GigabitEthernet 1/1)# label-switching
PE1(config-if-GigabitEthernet 1/1)# mpls ip
PE1(config-if-GigabitEthernet 1/1)# end
PE1# configure terminal
PE1(config)# router ospf 1
```

	PE1(config-router)# network 59.10.11.0 0.0.0.255 area 0 PE1(config-router)# network 59.10.10.0 0.0.0.255 area 0 PE1(config-router)# network 1.1.1.1 0.0.0.0 area 0 PE1(config-router)# end
CE1	CE1(config)# interface GigabitEthernet 0/1 CE1(config-if-GigabitEthernet 0/1)# ip address 20.10.10.2 255.255.255.0 CE1# configure terminal CE1(config)# router bgp 100 CE1(config-router)# neighbor 20.10.10.1 remote-as 1
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 PE2# configure terminal PE2(config)# ip vrf VPN1 PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target both 1:100 PE2(config-vrf)# end PE2# configure terminal PE2(config)# interface GigabitEthernet 1/2 PE2(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPN1 PE2(config-if-GigabitEthernet 1/2)# ip address 191.10.10.2 255.255.255.0 PE2(config-if-GigabitEthernet 1/2)# end PE2# configure terminal PE2(config)# router ospf 10 vrf VPN1 PE2(config-router)# network 191.10.10.0 0.0.0.255 area 0 PE2(config-router)# default-information originate PE2(config-router)# redistribute bgp subnets PE2(config-router)# exit PE2# configure terminal PE2(config)# router bgp 1 PE2(config-router)# neighbor 1.1.1.1 remote-as 1

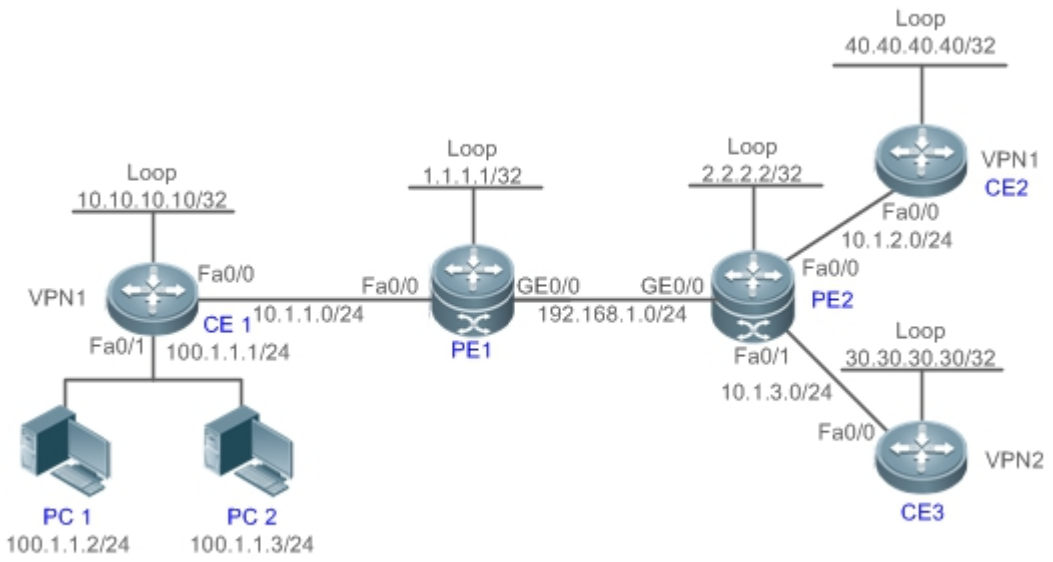
	PE2(config-router)# neighbor 1.1.1.1 update-source loopback 0 PE2(config-router)# address-family vpnv4 PE2(config-router-af)# neighbor 1.1.1.1 activate PE2(config-router-af)# end PE2(config)# router bgp 1 PE2(config-router)# address-family ipv4 vrf VPN1 PE2(config-router-af)# redistribute ospf 10 PE2(config-router-af)# end PE2# configure terminal PE2(config)# mpls ip PE2(config)# mpls router ldp PE2(config-mpls-router)# ldp router-id interface loopback 0 force PE2(config-mpls-router)# exit PE2(config)# interface GigabitEthernet 1/1 PE2(config-if-GigabitEthernet 1/1)# ip address 59.10.11.2 255.255.255.0 PE2(config-if-GigabitEthernet 1/1)# label-switching PE2(config-if-GigabitEthernet 1/1)# mpls ip PE2(config-if-GigabitEthernet 1/1)# end PE2# configure terminal PE2(config)# router ospf 1 PE2(config-router)# network 59.10.11.0 0.0.0.255 area 0 PE2(config-router)# network 2.2.2.2 0.0.0.0 area 0 PE2(config-router)# end
CE2	CE2# configure terminal CE2(config)# access-list 2000 deny ip any 17.10.0.0 0.0.255.255 CE2(config)# access-list 2000 permit ip any any CE2(config)# interface GigabitEthernet 0/1 CE2(config-if-GigabitEthernet 0/1)# ip access-group 2000 out CE2# configure terminal CE2(config)# router ospf 1 CE2(config-router)# network 191.10.10.0 0.0.0.255 area 0

	CE2(config-router)# network 17.11.0.0 0.0.0.255 area 0 CE2(config-router)# exit
PE3	PE3# configure terminal PE3(config)# interface loopback 0 PE3(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 PE3# configure terminal PE3(config)# ip vrf VPN1 PE3(config-vrf)# rd 1:200 PE3(config-vrf)# route-target both 1:200 PE3(config-vrf)# end PE3# configure terminal PE3(config)# interface GigabitEthernet 1/2 PE3(config-if-GigabitEthernet 1/2)# ip vrf forwarding VPN1 PE3(config-if-GigabitEthernet 1/2)# ip address 192.10.10.2 255.255.255.0 PE3(config-if-GigabitEthernet 1/2)# end PE3# configure terminal PE3(config)# router ospf 10 vrf VPN1 PE3(config-router)# network 192.10.10.0 0.0.0.255 area 0 PE3(config-router)# default-information originate PE3(config-router)# redistribute bgp subnets PE3(config-router)# exit PE3# configure terminal PE3(config)# router bgp 1 PE3(config-router)# neighbor 1.1.1.1 remote-as 1 PE3(config-router)# neighbor 1.1.1.1 update-source loopback 0 PE3(config-router)# address-family vpnv4 PE3(config-router-af)# neighbor 1.1.1.1 activate PE3(config-router-af)# end PE3(config)# router bgp 1 PE3(config-router)# address-family ipv4 vrf VPN1 PE3(config-router-af)# redistribute ospf 10

	PE3(config-router-af)# end PE3# configure terminal PE3(config)# mpls ip PE3(config)# mpls router ldp PE3(config-mpls-router)# ldp router-id interface loopback 0 force PE3(config-mpls-router)# exit PE3(config)# interface GigabitEthernet 1/1 PE3(config-if-GigabitEthernet 1/1)# ip address 59.10.10.2 255.255.255.0 PE3(config-if-GigabitEthernet 1/1)# label-switching PE3(config-if-GigabitEthernet 1/1)# mpls ip PE3(config-if-GigabitEthernet 1/1)# end PE3# configure terminal PE3(config)# router ospf 1 PE3(config-router)# network 59.10.10.0 0.0.0.255 area 0 PE3(config-router)# network 3.3.3.3 0.0.0.0 area 0 PE3(config-router)# end
CE3	CE3# configure terminal CE3(config)# access-list 2000 deny ip any 17.11.0.0 0.0.255.255 CE3(config)# access-list 2000 permit ip any any CE3(config)# interface GigabitEthernet 0/1 CE3(config-if-GigabitEthernet 0/1)# ip access-group 2000 out CE3# configure terminal CE3(config)# router ospf 1 CE3(config-router)# network 192.10.10.0 0.0.0.255 area 0 CE3(config-router)# network 17.10.0.0 0.0.0.255 area 0 CE3(config-router)# exit
Verification	● Run the ping command for connection checking.
PE1	PE2 and PE3 are pingable.
CE1	CE2 and CE3 are pingable.
PE2	PE1 and PE3 are pingable.

CE2	CE1 is pingable, whereas CE3 is unpingable.
PE3	PE1 and PE2 are pingable.
CE3	CE1 is pingable, whereas CE2 is unpingable.

Multi-role Host Configuration

Scenario **Figure 2- 23**	 Common hosts can only access other hosts of the same VPN site, whereas multi-role hosts can access other hosts of multiple VPNs. In Figure 2- , PC1 and PC2 are located at a site of VPN1. PC2 is a multi-role host able to access the hosts at the sites of VPN1 and VPN2, whereas PC1 can only access hosts at the same site of VPN1.
Configuration Steps	● Configure an MPLS backbone network. ● Configure an MPLS L3VPN network. ● Configure a multi-role host to access the VRF tables named VPN_MR for VPN1 and VPN2 respectively, and advertise the route of the multi-role host to other VPNs. ● Configure policy-based routing (PBR) to redirect the access packets sent by the multi-role host to the VPN_MR VRF table. ● Configure a default route on CE1.
PE1	PE1#configure terminal PE1(config)#interface loopback 0 PE1(config-if-Loopback 0)#ip address 1.1.1.1 255.255.255.255 PE1(config-if-Loopback 0)#exit PE1(config)#interface GigabitEthernet 0/0

```
PE1(config-if-GigabitEthernet 0/0)#ip address 192.168.1.1 255.255.255.0
PE1(config-if-GigabitEthernet 0/0)#exit
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0
PE1(config-router)#network 192.168.1.0 0.0.0.255 area 0
PE1(config-router)#end
PE1#configure terminal
PE1(config)#mpls ip
PE1(config)#interface GigabitEthernet 0/0
PE1(config-if-GigabitEthernet 0/0)#mpls ip
PE1(config-if-GigabitEthernet 0/0)#label-switching
PE1(config-if-GigabitEthernet 0/0)#exit
PE1(config)#mpls router ldp
PE1(config-mpls-router)#ldp router-id interface loopback 0 force
PE1(config-mpls-router)#end
PE1#configure terminal
PE1(config)#ip vrf VPN1
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config-vrf)#exit
PE1(config)#interface FastEthernet 0/0
PE1(config-if-FastEthernet 0/0)#ip vrf forwarding VPN1
PE1(config-if-FastEthernet 0/0)#ip address 10.1.1.1 255.255.255.0
PE1(config-if-FastEthernet 0/0)#exit
PE1(config)#router ospf 10 vrf VPN1
PE1(config-router)#network 10.1.1.0 0.0.0.255 area 0
PE1(config-router)#end
PE1#configure terminal
PE1(config)#router bgp 100
PE1(config-router)#neighbor 2.2.2.2 remote-as 100
PE1(config-router)#neighbor 2.2.2.2 update-source Loopback 0
PE1(config-router)#address-family vpnv4
```

```
PE1(config-router-af)#neighbor 2.2.2.2 activate
PE1(config-router-af)#exit
PE1(config-router)#address-family ipv4 vrf VPN1
PE1(config-router-af)#redistribute ospf 10
PE1(config-router-af)#exit
PE1(config-router)#exit
PE1(config)#router ospf 10 vrf VPN1
PE1(config-router)#redistribute bgp subnets
PE1(config-router)#end
PE1#configure terminal
PE1(config)#ip vrf VPN_MR
PE1(config-vrf)#rd 200:1
PE1(config-vrf)#route-target export 100:2
PE1(config-vrf)#route-target import 100:1
PE1(config-vrf)#route-target import 100:2
PE1(config-vrf)#end
PE1#configure terminal
PE1(config)#ip route vrf VPN_MR 100.1.1.3 255.255.255.255 FastEthernet 0/0
PE1(config)#router bgp 100
PE1(config-router)#address-family ipv4 vrf VPN_MR
PE1(config-router-af)#redistribute static
PE1(config-router-af)#end
PE1#configure terminal
PE1(config)#ip access-list standard VRF_ACL
PE1(config-std-nacl)#permit host 100.1.1.3
PE1(config-std-nacl)#deny any
PE1(config-std-nacl)#exit
PE1(config)#route-map VRF_MAP permit 10
PE1 (config-route-map)#match ip address VRF_ACL
PE1 (config-route-map)#set vrf VPN_MR
PE1 (config-route-map)#exit
PE1(config)#interface FastEthernet 0/0
```

	PE1(config-if-FastEthernet 0/0)#ip policy route-map VRF_MAP PE1(config-if-FastEthernet 0/0)#end
PE2	PE2#configure terminal PE2(config)#interface loopback 0 PE2(config-if-Loopback 0)#ip address 2.2.2.2 255.255.255.255 PE2(config-if-Loopback 0)#exit PE2(config)#interface GigabitEthernet 0/0 PE2(config-if-GigabitEthernet 0/0)#ip address 192.168.1.2 255.255.255.0 PE2(config-if-GigabitEthernet 0/0)#exit PE2(config)#router ospf 1 PE2(config-router)#network 2.2.2.2 0.0.0.0 area 0 PE2(config-router)#network 192.168.1.0 0.0.0.255 area 0 PE2(config-router)#end PE2#configure terminal PE2(config)#mpls ip PE2(config)#interface GigabitEthernet 0/0 PE2(config-if-GigabitEthernet 0/0)#mpls ip PE2(config-if-GigabitEthernet 0/0)#label-switching PE2(config-if-GigabitEthernet 0/0)#exit PE2(config)#mpls router ldp PE2(config-mpls-router)#ldp router-id interface loopback 0 force PE2(config-mpls-router)#end PE2#configure terminal PE2(config)#ip vrf VPN1 PE2(config-vrf)#rd 101:1 PE2(config-vrf)#route-target both 100:1 PE2(config-vrf)#exit PE2(config)#ip vrf VPN2 PE2(config-vrf)#rd 101:2 PE2(config-vrf)#route-target both 100:2 PE2(config-vrf)#exit

```
PE2(config)#interface FastEthernet 0/0
PE2(config-if-FastEthernet 0/0)#ip vrf forwarding VPN1
PE2(config-if-FastEthernet 0/0)#ip address 10.1.2.1 255.255.255.0
PE2(config-if-FastEthernet 0/0)#exit
PE2(config)#interface FastEthernet 0/1
PE2(config-if-FastEthernet 0/1)#ip vrf forwarding VPN2
PE2(config-if-FastEthernet 0/1)#ip address 10.1.3.1 255.255.255.0
PE2(config-if-FastEthernet 0/1)#exit
PE2(config)#router ospf 10 vrf VPN1
PE2(config-router)#network 10.1.2.0 0.0.0.255 area 0
PE2(config-router)#exit
PE2(config)#router ospf 20 vrf VPN2
PE2(config-router)#network 10.1.3.0 0.0.0.255 area 0
PE2(config-router)#end
PE2#configure terminal
PE2(config)#router bgp 100
PE2(config-router)#neighbor 1.1.1.1 remote-as 100
PE2(config-router)#neighbor 1.1.1.1 update-source Loopback 0
PE2(config-router)#address-family vpnv4
PE2(config-router-af)#neighbor 1.1.1.1 activate
PE2(config-router-af)#exit
PE2(config-router)#address-family ipv4 vrf VPN1
PE2(config-router-af)#redistribute ospf 10
PE2(config-router-af)#exit
PE2(config-router)#address-family ipv4 vrf VPN2
PE2(config-router-af)#redistribute ospf 20
PE2(config-router-af)#exit
PE2(config-router)#exit
PE2(config)#router ospf 10 vrf VPN1
PE2(config-router)#redistribute bgp subnets
PE2(config-router)#exit
PE2(config)#router ospf 20 vrf VPN2
```

	PE2(config-router)#redistribute bgp subnets PE2(config-router)#end
CE1	CE1#configure terminal CE1(config)#interface Loopback 0 CE1(config-if-Loopback 0)#ip address 10.10.10.10 255.255.255.255 CE1(config-if-Loopback 0)#exit CE1(config)#interface FastEthernet 0/0 CE1(config-if-FastEthernet 0/0)#ip address 10.1.1.2 255.255.255.0 CE1(config-if-FastEthernet 0/0)#exit CE1(config)#interface FastEthernet 0/1 CE1(config-if-FastEthernet 0/1)#ip address 100.1.1.1 255.255.255.0 CE1(config-if-FastEthernet 0/1)#exit CE1(config)#router ospf 1 CE1(config-router)#network 10.1.1.0 0.0.0.255 area 0 CE1(config-router)#network 100.1.1.0 0.0.0.255 area 0 CE1(config-router)#network 10.10.10.10 0.0.0.0 area 0 CE1(config-router)#end CE1#configure terminal CE1(config)#ip route 0.0.0.0 0.0.0.0 10.1.1.1
CE2	CE2#configure terminal CE2(config)#interface Loopback 0 CE2(config-if-Loopback 0)#ip address 20.20.20.20 255.255.255.255 CE2(config-if-Loopback 0)#exit CE2(config)#interface FastEthernet 0/0 CE2(config-if-FastEthernet 0/0)#ip address 10.1.2.2 255.255.255.0 CE2(config-if-FastEthernet 0/0)#exit CE2(config)#router ospf 1 CE2(config-router)#network 10.1.2.0 0.0.0.255 area 0 CE2(config-router)#network 20.20.20.20 0.0.0.0 area 0 CE2(config-router)#end
CE3	CE3#configure terminal

	CE3(config)#interface Loopback 0 CE3(config-if-Loopback 0)#ip address 30.30.30.30 255.255.255.255 CE3(config-if-Loopback 0)#exit CE3(config)#interface FastEthernet 0/0 CE3(config-if-FastEthernet 0/0)#ip address 10.1.3.2 255.255.255.0 CE3(config-if-FastEthernet 0/0)#exit CE3(config)#router ospf 1 CE3(config-router)#network 10.1.3.0 0.0.0.255 area 0 CE3(config-router)#network 30.30.30.30 0.0.0.0 area 0 CE3(config-router)#end
Verification	● Run the ping command for connection checking.

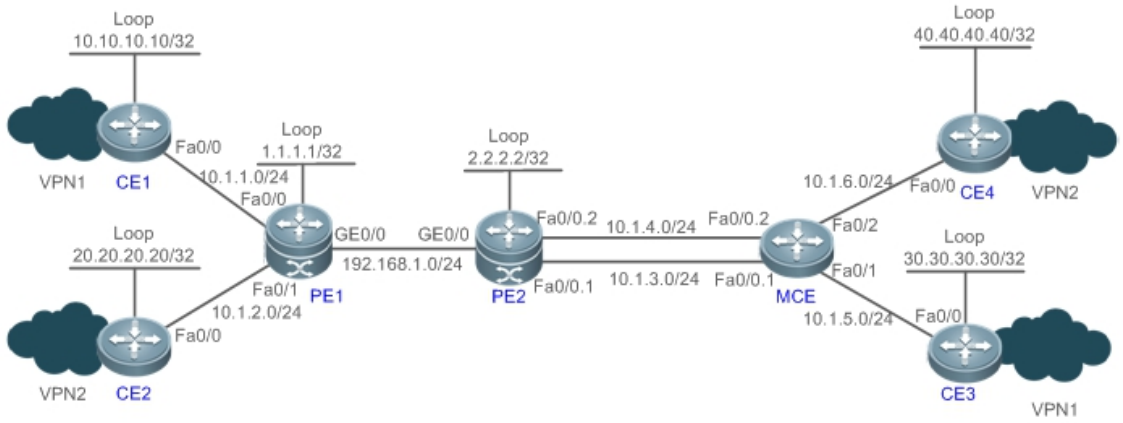
PC1	$ ping 20.20.20.20
	PING 20.20.20.20 (20.20.20.20): 56 data bytes
	64 bytes from 20.20.20.20: icmp_seq=0 ttl=61 time=15 ms
	64 bytes from 20.20.20.20: icmp_seq=1 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=2 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=3 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=4 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=5 ttl=61 time=0 ms
	
	----20.20.20.20 PING Statistics----
	6 packets transmitted, 6 packets received, 0.0% packet loss
	round-trip (ms) min/avg/max/med = 0/3/15/0
	
	$ ping 30.30.30.30
	PING 30.30.30.30 (30.30.30.30): 56 data bytes
	
	----30.30.30.30 PING Statistics----
	3 packets transmitted, 0 packets received, 100.0% packet loss

PC2	$ ping 20.20.20.20
	PING 20.20.20.20 (20.20.20.20): 56 data bytes
	64 bytes from 20.20.20.20: icmp_seq=0 ttl=61 time=15 ms
	64 bytes from 20.20.20.20: icmp_seq=1 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=2 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=3 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=4 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=5 ttl=61 time=0 ms
	64 bytes from 20.20.20.20: icmp_seq=6 ttl=61 time=0 ms
	
	----20.20.20.20 PING Statistics----
	7 packets transmitted, 7 packets received, 0.0% packet loss
	round-trip (ms) min/avg/max/med = 0/2/15/0
	
	$ ping 30.30.30.30
	PING 30.30.30.30 (30.30.30.30): 56 data bytes
	64 bytes from 30.30.30.30: icmp_seq=0 ttl=61 time=15 ms
	64 bytes from 30.30.30.30: icmp_seq=1 ttl=61 time=0 ms
	64 bytes from 30.30.30.30: icmp_seq=2 ttl=61 time=0 ms
	64 bytes from 30.30.30.30: icmp_seq=3 ttl=61 time=0 ms
	64 bytes from 30.30.30.30: icmp_seq=4 ttl=61 time=0 ms
	64 bytes from 30.30.30.30: icmp_seq=5 ttl=61 time=0 ms
	64 bytes from 30.30.30.30: icmp_seq=6 ttl=61 time=0 ms
	64 bytes from 30.30.30.30: icmp_seq=7 ttl=61 time=0 ms
	
	----30.30.30.30 PING Statistics----
	8 packets transmitted, 8 packets received, 0.0% packet loss
	round-trip (ms) min/avg/max/med = 0/2/15/0

MCE-based Hierarchy VPN Configuration

Scenario

Figure 2- 24



A multi-VPN-instance customer edge (MCE) is deployed to reconstruct the single-plane MPLS L3VPN network into a hierarchical network. Figure 2- shows the network topology. The MCE stores the routes of the connected VPN sites and the default routes distributed by PE2, but does not store the routes of other VPN sites. This poses relatively low requirements on capacity and performance. PE1 and PE2 must store all VPN routes, which poses relatively high requirements on capacity and performance.

Configuration Steps

- (MPLS network configuration) Configure interface addresses and OSPF on PE1 and PE2, and configure the MPLS feature.
- (MPLS L3VPN network configuration) Configure CE-to-PE access on PE1, PE2, CE1, and CE2. Configure MP-IBGP on PE1 and PE2.
- Configure the MCE and the connected CEs.
- Configure default route distribution for the MCE.

PE1

```
PE1#configure terminal
PE1(config)#interface loopback 0
PE1(config-if-Loopback 0)#ip address 1.1.1.1 255.255.255.255
PE1(config-if-Loopback 0)#exit
PE1(config)#interface GigabitEthernet 0/0
PE1(config-if-GigabitEthernet 0/0)#ip address 192.168.1.1 255.255.255.0
PE1(config-if-GigabitEthernet 0/0)#exit
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0
PE1(config-router)#network 192.168.1.0 0.0.0.255 area 0
PE1(config-router)#end
PE1#configure terminal
PE1(config)#mpls ip
```

```
PE1(config)#interface GigabitEthernet 0/0
PE1(config-if-GigabitEthernet 0/0)#mpls ip
PE1(config-if-GigabitEthernet 0/0)#label-switching
PE1(config-if-GigabitEthernet 0/0)#exit
PE1(config)#mpls router ldp
PE1(config-mpls-router)#ldp router-id interface loopback 0 force
PE1(config-mpls-router)#end
PE1#configure terminal
PE1(config)#ip vrf VPN1
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config-vrf)#exit
PE1(config)#ip vrf VPN2
PE1(config-vrf)#rd 100:2
PE1(config-vrf)#route-target both 100:2
PE1(config-vrf)#exit
PE1(config)#interface FastEthernet 0/0
PE1(config-if-FastEthernet 0/0)#ip vrf forwarding VPN1
PE1(config-if-FastEthernet 0/0)#ip address 10.1.1.1 255.255.255.0
PE1(config-if-FastEthernet 0/0)#exit
PE1(config)#interface FastEthernet 0/1
PE1(config-if-FastEthernet 0/1)#ip vrf forwarding VPN2
PE1(config-if-FastEthernet 0/1)#ip address 10.1.2.1 255.255.255.0
PE1(config-if-FastEthernet 0/1)#exit
PE1(config)#router ospf 10 vrf VPN1
PE1(config-router)#network 10.1.1.0 0.0.0.255 area 0
PE1(config-router)#exit
PE1(config)#router ospf 20 vrf VPN2
PE1(config-router)#network 10.1.2.0 0.0.0.255 area 0
PE1(config-router)#end
PE1#configure terminal
PE1(config)#router bgp 100
```

	PE1(config-router)#neighbor 2.2.2.2 remote-as 100 PE1(config-router)#neighbor 2.2.2.2 update-source Loopback 0 PE1(config-router)#address-family vpnv4 PE1(config-router-af)#neighbor 2.2.2.2 activate PE1(config-router-af)#exit PE1(config-router)#address-family ipv4 vrf VPN1 PE1(config-router-af)#redistribute ospf 10 PE1(config-router-af)#exit PE1(config-router)#address-family ipv4 vrf VPN2 PE1(config-router-af)#redistribute ospf 20 PE1(config-router-af)#exit PE1(config-router)#exit PE1(config)#router ospf 10 vrf VPN1 PE1(config-router)#redistribute bgp subnets PE1(config-router)#exit PE1(config)#router ospf 20 vrf VPN2 PE1(config-router)#redistribute bgp subnets PE1(config-router)#end
PE2	PE2#configure terminal PE2(config)#interface loopback 0 PE2(config-if-Loopback 0)#ip address 2.2.2.2 255.255.255.255 PE2(config-if-Loopback 0)#exit PE2(config)#interface GigabitEthernet 0/0 PE2(config-if-GigabitEthernet 0/0)#ip address 192.168.1.2 255.255.255.0 PE2(config-if-GigabitEthernet 0/0)#exit PE2(config)#router ospf 1 PE2(config-router)#network 2.2.2.2 0.0.0.0 area 0 PE2(config-router)#network 192.168.1.0 0.0.0.255 area 0 PE2(config-router)#end PE2#configure terminal PE2(config)#mpls ip

```
PE2(config)#interface GigabitEthernet 0/0
PE2(config-if-GigabitEthernet 0/0)#mpls ip
PE2(config-if-GigabitEthernet 0/0)#label-switching
PE2(config-if-GigabitEthernet 0/0)#exit
PE2(config)#mpls router ldp
PE2(config-mpls-router)#ldp router-id interface loopback 0 force
PE2(config-mpls-router)#end
PE2#configure terminal
PE2(config)#ip vrf VPN1
PE2(config-vrf)#rd 101:1
PE2(config-vrf)#route-target both 100:1
PE2(config-vrf)#exit
PE2(config)#ip vrf VPN2
PE2(config-vrf)#rd 101:2
PE2(config-vrf)#route-target both 100:2
PE2(config-vrf)#exit
PE2(config)#interface FastEthernet 0/0.1
PE2(config-if-FastEthernet 0/0.1)#encapsulation dot1Q 1
PE2(config-if-FastEthernet 0/0.1)#ip vrf forwarding VPN1
PE2(config-if-FastEthernet 0/0.1)#ip address 10.1.3.1 255.255.255.0
PE2(config-if-FastEthernet 0/0.1)#exit
PE2(config)#interface FastEthernet 0/0.2
PE2(config-if-FastEthernet 0/0.2)#encapsulation dot1Q 2
PE2(config-if-FastEthernet 0/0.2)#ip vrf forwarding VPN2
PE2(config-if-FastEthernet 0/0.2)#ip address 10.1.4.1 255.255.255.0
PE2(config-if-FastEthernet 0/0.2)#exit
PE2(config)#router ospf 10 vrf VPN1
PE2(config-router)#network 10.1.3.0 0.0.0.255 area 0
PE2(config-router)#exit
PE2(config)#router ospf 20 vrf VPN2
PE2(config-router)#network 10.1.4.0 0.0.0.255 area 0
PE2(config-router)#end
```

	PE2#configure terminal PE2(config)#router bgp 100 PE2(config-router)#neighbor 1.1.1.1 remote-as 100 PE2(config-router)#neighbor 1.1.1.1 update-source Loopback 0 PE2(config-router)#address-family vpnv4 PE2(config-router-af)#neighbor 1.1.1.1 activate PE2(config-router-af)#exit PE2(config-router)#address-family ipv4 vrf VPN1 PE2(config-router-af)#redistribute ospf 10 PE2(config-router-af)#exit PE2(config-router)#address-family ipv4 vrf VPN2 PE2(config-router-af)#redistribute ospf 20 PE2(config-router-af)#exit PE2(config-router)#end PE2#configure terminal PE2(config)#ip route vrf VPN1 0.0.0.0 0.0.0.0 null 0 PE2(config)#ip route vrf VPN2 0.0.0.0 0.0.0.0 null 0 PE2(config)#router ospf 10 vrf VPN1 PE2(config-router)#default-information originate PE2(config-router)#exit PE2(config)#router ospf 20 vrf VPN2 PE2(config-router)#default-information originate PE2(config-router)#end
CE1	CE1#configure terminal CE1(config)#interface loopback 0 CE1(config-if-Loopback 0)#ip address 10.10.10.10 255.255.255.255 CE1(config-if-Loopback 0)#exit CE1(config)#interface FastEthernet 0/0 CE1(config-if-FastEthernet 0/0)#ip address 10.1.1.2 255.255.255.0 CE1(config-if-FastEthernet 0/0)#exit CE1(config)#router ospf 1

	CE1(config-router)#network 10.1.1.0 0.0.0.255 area 0 CE1(config-router)#network 10.10.10.10 0.0.0.0 area 0 CE1(config-router)#end
CE2	CE2#configure terminal CE2(config)#interface loopback 0 CE2(config-if-Loopback 0)#ip address 20.20.20.20 255.255.255.255 CE2(config-if-Loopback 0)#exit CE2(config)#interface FastEthernet 0/0 CE2(config-if-FastEthernet 0/0)#ip address 10.1.2.2 255.255.255.0 CE2(config-if-FastEthernet 0/0)#exit CE2(config)#router ospf 1 CE2(config-router)#network 10.1.2.0 0.0.0.255 area 0 CE2(config-router)#network 20.20.20.20 0.0.0.0 area 0 CE2(config-router)#end
CE3	CE3#configure terminal CE3(config)#interface loopback 0 CE3(config-if-Loopback 0)#ip address 30.30.30.30 255.255.255.255 CE3(config-if-Loopback 0)#exit CE3(config)#interface FastEthernet 0/0 CE3(config-if-FastEthernet 0/0)#ip address 10.1.5.2 255.255.255.0 CE3(config-if-FastEthernet 0/0)#exit CE3(config)#router ospf 1 CE3(config-router)#network 10.1.5.0 0.0.0.255 area 0 CE3(config-router)#network 30.30.30.30 0.0.0.0 area 0 CE3(config-router)#end
CE4	CE4#configure terminal CE4(config)#interface loopback 0 CE4(config-if-Loopback 0)#ip address 40.40.40.40 255.255.255.255 CE4(config-if-Loopback 0)#exit CE4(config)#interface FastEthernet 0/0 CE4(config-if-FastEthernet 0/0)#ip address 10.1.6.2 255.255.255.0

	CE4(config-if-FastEthernet 0/0)#exit CE4(config)#router ospf 1 CE4(config-router)#network 10.1.6.0 0.0.0.255 area 0 CE4(config-router)#network 40.40.40.40 0.0.0.0 area 0 CE4(config-router)#end
MCE	MCE#configure terminal MCE(config)#ip vrf VPN1 MCE(config-vrf)#rd 102:1 MCE(config-vrf)#exit MCE(config)#ip vrf VPN2 MCE(config-vrf)#rd 102:2 MCE(config-vrf)#exit MCE(config)#interface FastEthernet 0/0.1 MCE(config-if-FastEthernet 0/0.1)#encapsulation dot1Q 1 MCE(config-if-FastEthernet 0/0.1)#ip vrf forwarding VPN1 MCE(config-if-FastEthernet 0/0.1)#ip address 10.1.3.2 255.255.255.0 MCE(config-if-FastEthernet 0/0.1)#exit MCE(config)#interface FastEthernet 0/0.2 MCE(config-if-FastEthernet 0/0.2)#encapsulation dot1Q 2 MCE(config-if-FastEthernet 0/0.2)#ip vrf forwarding VPN2 MCE(config-if-FastEthernet 0/0.2)#ip address 10.1.4.2 255.255.255.0 MCE(config-if-FastEthernet 0/0.2)#exit MCE(config)#interface FastEthernet 0/1 MCE(config-if-FastEthernet 0/1)#ip vrf forwarding VPN1 MCE(config-if-FastEthernet 0/1)#ip address 10.1.5.1 255.255.255.0 MCE(config-if-FastEthernet 0/1)#exit MCE(config)#interface FastEthernet 0/2 MCE(config-if-FastEthernet 0/2)#ip vrf forwarding VPN2 MCE(config-if-FastEthernet 0/2)#ip address 10.1.6.1 255.255.255.0 MCE(config-if-FastEthernet 0/2)#exit MCE(config)#router ospf 10 vrf VPN1

	MCE(config-router)#network 10.1.3.0 0.0.0.255 area 0 MCE(config-router)#network 10.1.5.0 0.0.0.255 area 0 MCE(config-router)#exit MCE(config)#router ospf 20 vrf VPN2 MCE(config-router)#network 10.1.4.0 0.0.0.255 area 0 MCE(config-router)#network 10.1.6.0 0.0.0.255 area 0 MCE(config-router)#end
Verification	● Run the **show ip route** command to display VPN routes. ● Check that hosts at the same VPN site are pingable between each other, whereas hosts at different VPN sites are unpingable between each other.
MCE	MCE#show ip route vrf VPN1 Routing Table: VPN1 Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is 10.1.3.1 to network 0.0.0.0 O*E2 0.0.0.0/0 [110/1] via 10.1.3.1, 00:38:35, FastEthernet 0/0.1 C 10.1.3.0/24 is directly connected, FastEthernet 0/0.1 C 10.1.3.2/32 is local host. C 10.1.5.0/24 is directly connected, FastEthernet 0/1 C 10.1.5.1/32 is local host. O 30.30.30.30/32 [110/1] via 10.1.5.2, 00:40:05, FastEthernet 0/1 MCE#show ip route vrf VPN2 Routing Table: VPN2

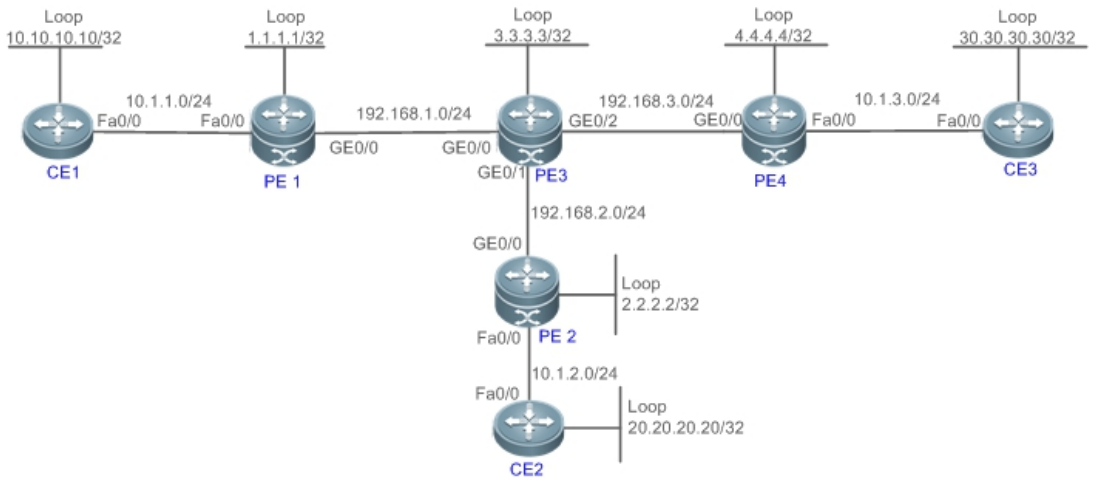
	Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is 10.1.4.1 to network 0.0.0.0 O*E2 0.0.0.0/0 [110/1] via 10.1.4.1, 00:38:35, FastEthernet 0/0.2 C 10.1.4.0/24 is directly connected, FastEthernet 0/0.2 C 10.1.4.2/32 is local host. C 10.1.6.0/24 is directly connected, FastEthernet 0/2 C 10.1.6.1/32 is local host. O 40.40.40.40/32 [110/1] via 10.1.6.2, 00:40:07, FastEthernet 0/2
CE3	CE3#ping 10.10.10.10 source 30.30.30.30 Sending 5, 100-byte ICMP Echoes to 10.10.10.10, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/10/20 ms CE3#ping 20.20.20.20 source 30.30.30.30 Sending 5, 100-byte ICMP Echoes to 20.20.20.20, timeout is 2 seconds: < press Ctrl+C to break > ..... Success rate is 0 percent (0/5)
CE4	CE4#ping 20.20.20.20 source 40.40.40.40 Sending 5, 100-byte ICMP Echoes to 20.20.20.20, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/10/20 ms

```
CE3#ping 10.10.10.10 source 40.40.40.40

Sending 5, 100-byte ICMP Echoes to 10.10.10.10, timeout is 2 seconds:

  < press Ctrl+C to break >

.....

Success rate is 0 percent (0/5)
```

Hierarchy VPN Configuration Based on BGP Routing Policies

Scenario

Figure 2- 25



BGP routing policies are implemented to reconstruct the single-plane MPLS L3VPN network into a hierarchical network. Figure 2- shows the network topology. PE1 and PE2 are lower-layer PEs that store only the routes of associated VPN sites and the default routes advertised by the upper-layer PEs. PE3 and PE4 are upper-layer PEs that store all VPN routes. PE3 is the upper-layer PE for PE1 and PE2.

Configuration Steps

- (MPLS network configuration) Configure interface addresses and OSPF on PE1, PE2, PE3, and PE4, and configure the MPLS feature.
- (MPLS L3VPN network configuration, with PE3 functioning as an RR) Configure CE-to-PE access on the PEs and CEs. Configure corresponding VRF tables on PE3, because PE3 needs to store VPN routes though it is not connected to any VPN site. Configure MP-IBGP on PE1, PE2, PE3, and PE4.
- Configure routing policies on PE3 to restrict VPN route advertisement to PE1 and PE2.
- Configure default VPN routes on PE3 and advertise the routes to PE1 and PE2 (do not advertise the routes to PE4).
- Configure VRF OSPF on PE1 and PE2 for default route advertisement.

PE1

```
PE1#configure terminal

PE1(config)#interface loopback 0

PE1(config-if-Loopback 0)#ip address 1.1.1.1 255.255.255.255

PE1(config-if-Loopback 0)#exit

PE1(config)#interface GigabitEthernet 0/0
```

```
PE1(config-if-GigabitEthernet 0/0)#ip address 192.168.1.1 255.255.255.0
PE1(config-if-GigabitEthernet 0/0)#exit
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0
PE1(config-router)#network 192.168.1.0 0.0.0.255 area 0
PE1(config-router)#end
PE1#configure terminal
PE1(config)#mpls ip
PE1(config)#interface GigabitEthernet 0/0
PE1(config-if-GigabitEthernet 0/0)#mpls ip
PE1(config-if-GigabitEthernet 0/0)#label-switching
PE1(config-if-GigabitEthernet 0/0)#exit
PE1(config)#mpls router ldp
PE1(config-mpls-router)#ldp router-id interface loopback 0 force
PE1(config-mpls-router)#end
PE1#configure terminal
PE1(config)#ip vrf VPN1
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config-vrf)#exit
PE1(config)#interface FastEthernet 0/0
PE1(config-if-FastEthernet 0/0)#ip vrf forwarding VPN1
PE1(config-if-FastEthernet 0/0)#ip address 10.1.1.1 255.255.255.0
PE1(config-if-FastEthernet 0/0)#exit
PE1(config)#router ospf 10 vrf VPN1
PE1(config-router)#network 10.1.1.0 0.0.0.255 area 0
PE1(config-router)#end
PE1#configure terminal
PE1(config)#router bgp 100
PE1(config-router)#neighbor 3.3.3.3 remote-as 100
PE1(config-router)#neighbor 3.3.3.3 update-source Loopback 0
PE1(config-router)#address-family vpnv4
```

	PE1(config-router-af)#neighbor 3.3.3.3 activate PE1(config-router-af)#exit PE1(config-router)#address-family ipv4 vrf VPN1 PE1(config-router-af)#redistribute ospf 10 PE1(config-router-af)#exit PE1(config-router)#exit PE1(config)#router ospf 10 vrf VPN1 PE1(config-router)#redistribute bgp subnets PE1(config-router)#end PE1#configure terminal PE1(config)#router ospf 10 vrf VPN1 PE1(config-router)#default-information originate PE1(config-router)#end
PE2	PE2#configure terminal PE2(config)#interface loopback 0 PE2(config-if-Loopback 0)#ip address 2.2.2.2 255.255.255.255 PE2(config-if-Loopback 0)#exit PE2(config)#interface GigabitEthernet 0/0 PE2(config-if-GigabitEthernet 0/0)#ip address 192.168.2.1 255.255.255.0 PE2(config-if-GigabitEthernet 0/0)#exit PE2(config)#router ospf 1 PE2(config-router)#network 2.2.2.2 0.0.0.0 area 0 PE2(config-router)#network 192.168.2.0 0.0.0.255 area 0 PE2(config-router)#end PE2#configure terminal PE2(config)#mpls ip PE2(config)#interface GigabitEthernet 0/0 PE2(config-if-GigabitEthernet 0/0)#mpls ip PE2(config-if-GigabitEthernet 0/0)#label-switching PE2(config-if-GigabitEthernet 0/0)#exit PE2(config)#mpls router ldp

```
PE2(config-mpls-router)#ldp router-id interface loopback 0 force
PE2(config-mpls-router)#end
PE2#configure terminal
PE2(config)#ip vrf VPN1
PE2(config-vrf)#rd 100:2
PE2(config-vrf)#route-target both 100:1
PE2(config-vrf)#exit
PE2(config)#interface FastEthernet 0/0
PE2(config-if-FastEthernet 0/0)#ip vrf forwarding VPN1
PE2(config-if-FastEthernet 0/0)#ip address 10.1.2.1 255.255.255.0
PE2(config-if-FastEthernet 0/0)#exit
PE2(config)#router ospf 10 vrf VPN1
PE2(config-router)#network 10.1.2.0 0.0.0.255 area 0
PE2(config-router)#end
PE2#configure terminal
PE2(config)#router bgp 100
PE2(config-router)#neighbor 3.3.3.3 remote-as 100
PE2(config-router)#neighbor 3.3.3.3 update-source Loopback 0
PE2(config-router)#address-family vpnv4
PE2(config-router-af)#neighbor 3.3.3.3 activate
PE2(config-router-af)#exit
PE2(config-router)#address-family ipv4 vrf VPN1
PE2(config-router-af)#redistribute ospf 10
PE2(config-router-af)#exit
PE2(config-router)#exit
PE2(config)#router ospf 10 vrf VPN1
PE2(config-router)#redistribute bgp subnets
PE2(config-router)#end
PE2#configure terminal
PE2(config)#router ospf 10 vrf VPN1
PE2(config-router)#default-information originate
PE2(config-router)#end
```

PE3	PE3#configure terminal PE3(config)#interface loopback 0 PE3(config-if-Loopback 0)#ip address 3.3.3.3 255.255.255.255 PE3(config-if-Loopback 0)#exit PE3(config)#interface GigabitEthernet 0/0 PE3(config-if-GigabitEthernet 0/0)#ip address 192.168.1.2 255.255.255.0 PE3(config-if-GigabitEthernet 0/0)#exit PE3(config)#interface GigabitEthernet 0/1 PE3(config-if-GigabitEthernet 0/1)#ip address 192.168.2.2 255.255.255.0 PE3(config-if-GigabitEthernet 0/1)#exit PE3(config)#interface GigabitEthernet 0/2 PE3(config-if-GigabitEthernet 0/2)#ip address 192.168.3.2 255.255.255.0 PE3(config-if-GigabitEthernet 0/2)#exit PE3(config)#router ospf 1 PE3(config-router)#network 3.3.3.3 0.0.0.0 area 0 PE3(config-router)#network 192.168.1.0 0.0.0.255 area 0 PE3(config-router)#network 192.168.2.0 0.0.0.255 area 0 PE3(config-router)#network 192.168.3.0 0.0.0.255 area 0 PE3(config-router)#end PE3#configure terminal PE3(config)#mpls ip PE3(config)#interface GigabitEthernet 0/0 PE3(config-if-GigabitEthernet 0/0)#mpls ip PE3(config-if-GigabitEthernet 0/0)#label-switching PE3(config-if-GigabitEthernet 0/0)#exit PE3(config)#interface GigabitEthernet 0/1 PE3(config-if-GigabitEthernet 0/1)#mpls ip PE3(config-if-GigabitEthernet 0/1)#label-switching PE3(config-if-GigabitEthernet 0/1)#exit PE3(config)#interface GigabitEthernet 0/2 PE3(config-if-GigabitEthernet 0/2)#mpls ip

```
PE3(config-if-GigabitEthernet 0/2)#label-switching
PE3(config-if-GigabitEthernet 0/2)#exit
PE3(config)#mpls router ldp
PE3(config-mpls-router)#ldp router-id interface loopback 0 force
PE3(config-mpls-router)#end
PE3#configure terminal
PE3(config)#ip vrf VPN1
PE3(config-vrf)#rd 100:3
PE3(config-vrf)#route-target both 100:1
PE3(config-vrf)#end
PE3#configure terminal
PE3(config)#router bgp 100
PE3(config-router)#neighbor 1.1.1.1 remote-as 100
PE3(config-router)#neighbor 1.1.1.1 update-source Loopback 0
PE3(config-router)#neighbor 2.2.2.2 remote-as 100
PE3(config-router)#neighbor 2.2.2.2 update-source Loopback 0
PE3(config-router)#neighbor 4.4.4.4 remote-as 100
PE3(config-router)#neighbor 4.4.4.4 update-source Loopback 0
PE3(config-router)#address-family vpnv4
PE3(config-router-af)#neighbor 1.1.1.1 activate
PE3(config-router-af)#neighbor 1.1.1.1 route-reflector-client
PE3(config-router-af)#neighbor 2.2.2.2 activate
PE3(config-router-af)#neighbor 2.2.2.2 route-reflector-client
PE3(config-router-af)#neighbor 4.4.4.4 activate
PE3(config-router-af)#neighbor 4.4.4.4 route-reflector-client
PE3(config-router-af)#exit
PE3(config-router)#address-family ipv4 vrf VPN1
PE3(config-router-af)#end
PE3#configure terminal
PE3(config)#ip access-list standard UPE_FILT_ACL
PE3(config-std-nacl)#permit 0.0.0.0 0.0.0.0
PE3(config-std-nacl)#deny any
```

	PE3(config-std-nacl)#exit PE3(config)#route-map UPE_FILT_RMP permit 10 PE3(config-route-map)#match ip address UPE_FILT_ACL PE3(config-route-map)#exit PE3(config)#router bgp 100 PE3(config-router)#address-family vpnv4 unicast PE3(config-router-af)#neighbor 1.1.1.1 route-map UPE_FILT_RMP out PE3(config-router-af)#neighbor 2.2.2.2 route-map UPE_FILT_RMP out PE3(config-router-af)#end PE3#configure terminal PE3(config)#ip access-list standard PE_FILT_ACL PE3(config-std-nacl)#deny 0.0.0.0 0.0.0.0 PE3(config-std-nacl)#permit any PE3(config-std-nacl)#exit PE3(config)#route-map PE_FILT_RMP permit 10 PE3(config-route-map)#match ip address PE_FILT_ACL PE3(config-route-map)#exit PE3(config)#ip route vrf VPN1 0.0.0.0 0.0.0.0 Null 0 PE3(config)#router bgp 100 PE3(config-router)#address-family vpnv4 unicast PE3(config-router-af)#neighbor 4.4.4.4 route-map PE_FILT_RMP out PE3(config-router-af)#exit PE3(config-router)#address-family ipv4 vrf VPN1 PE3(config-router-af)#network 0.0.0.0 PE3(config-router-af)#end
PE4	PE4#configure terminal PE4(config)#interface loopback 0 PE4(config-if-Loopback 0)#ip address 4.4.4.4 255.255.255.255 PE4(config-if-Loopback 0)#exit PE4(config)#interface GigabitEthernet 0/0 PE4(config-if-GigabitEthernet 0/0)#ip address 192.168.3.1 255.255.255.0

```
PE4(config-if-GigabitEthernet 0/0)#exit
PE4(config)#router ospf 1
PE4(config-router)#network 4.4.4.4 0.0.0.0 area 0
PE4(config-router)#network 192.168.3.0 0.0.0.255 area 0
PE4(config-router)#end
PE4#configure terminal
PE4(config)#mpls ip
PE4(config)#interface GigabitEthernet 0/0
PE4(config-if-GigabitEthernet 0/0)#mpls ip
PE4(config-if-GigabitEthernet 0/0)#label-switching
PE4(config-if-GigabitEthernet 0/0)#exit
PE4(config)#mpls router ldp
PE4(config-mpls-router)#ldp router-id interface loopback 0 force
PE4(config-mpls-router)#end
PE4#configure terminal
PE4(config)#ip vrf VPN1
PE4(config-vrf)#rd 100:4
PE4(config-vrf)#route-target both 100:1
PE4(config-vrf)#exit
PE4(config)#interface FastEthernet 0/0
PE4(config-if-FastEthernet 0/0)#ip vrf forwarding VPN1
PE4(config-if-FastEthernet 0/0)#ip address 10.1.3.1 255.255.255.0
PE4(config-if-FastEthernet 0/0)#exit
PE4(config)#router ospf 10 vrf VPN1
PE4(config-router)#network 10.1.3.0 0.0.0.255 area 0
PE4(config-router)#end
PE4#configure terminal
PE4(config)#router bgp 100
PE4(config-router)#neighbor 3.3.3.3 remote-as 100
PE4(config-router)#neighbor 3.3.3.3 update-source Loopback 0
PE4(config-router)#address-family vpnv4
PE4(config-router-af)#neighbor 3.3.3.3 activate
```

	PE4(config-router-af)#exit PE4(config-router)#address-family ipv4 vrf VPN1 PE4(config-router-af)#redistribute ospf 10 PE4(config-router-af)#exit PE4(config-router)#exit PE4(config)#router ospf 10 vrf VPN1 PE4(config-router)#redistribute bgp subnets PE4(config-router)#end
CE1	CE1#configure terminal CE1(config)#interface loopback 0 CE1(config-if-Loopback 0)#ip address 10.10.10.10 255.255.255.255 CE1(config-if-Loopback 0)#exit CE1(config)#interface FastEthernet 0/0 CE1(config-if-FastEthernet 0/0)#ip address 10.1.1.2 255.255.255.0 CE1(config-if-FastEthernet 0/0)#exit CE1(config)#router ospf 1 CE1(config-router)#network 10.1.1.0 0.0.0.255 area 0 CE1(config-router)#network 10.10.10.10 0.0.0.0 area 0 CE1(config-router)#end
CE2	CE2#configure terminal CE2(config)#interface loopback 0 CE2(config-if-Loopback 0)#ip address 20.20.20.20 255.255.255.255 CE2(config-if-Loopback 0)#exit CE2(config)#interface FastEthernet 0/0 CE2(config-if-FastEthernet 0/0)#ip address 10.1.2.2 255.255.255.0 CE2(config-if-FastEthernet 0/0)#exit CE2(config)#router ospf 1 CE2(config-router)#network 10.1.2.0 0.0.0.255 area 0 CE2(config-router)#network 20.20.20.20 0.0.0.0 area 0 CE2(config-router)#end
CE3	CE3#configure terminal

	CE3(config)#interface loopback 0 CE3(config-if-Loopback 0)#ip address 30.30.30.30 255.255.255.255 CE3(config-if-Loopback 0)#exit CE3(config)#interface FastEthernet 0/0 CE3(config-if-FastEthernet 0/0)#ip address 10.1.3.2 255.255.255.0 CE3(config-if-FastEthernet 0/0)#exit CE3(config)#router ospf 1 CE3(config-router)#network 10.1.3.0 0.0.0.255 area 0 CE3(config-router)#network 30.30.30.30 0.0.0.0 area 0 CE3(config-router)#end
Verification	● Run the **show ip route** command to display VPN routes. ● Check that VPN sites are pingable between each other.
PE1	PE1#show ip route vrf VPN1 Routing Table: VPN1 Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is 3.3.3.3 to network 0.0.0.0 B* 0.0.0.0/0 [200/0] via 3.3.3.3, 00:00:09 C 10.1.1.0/24 is directly connected, FastEthernet 0/0 C 10.1.1.1/32 is local host. O 10.10.10.10/32 [110/2] via 10.1.1.2, 01:36:19, FastEthernet 0/0
PE3	PE3#show ip route vrf VPN1

	Routing Table: VPN1 Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is 0.0.0.0 to network 0.0.0.0 S* 0.0.0.0/0 is directly connected, Null 0 B 10.1.1.0/24 [200/1] via 1.1.1.1, 01:54:23 B 10.1.2.0/24 [200/1] via 2.2.2.2, 01:54:27 B 10.1.3.0/24 [200/1] via 4.4.4.4, 01:54:29 B 10.10.10.10/32 [200/1] via 1.1.1.1, 01:54:23 B 20.20.20.20/32 [200/2] via 2.2.2.2, 01:54:27 B 30.30.30.30/32 [200/1] via 4.4.4.4, 01:54:29
CE1	CE1#ping 20.20.20.20 source 10.10.10.10 Sending 5, 100-byte ICMP Echoes to 20.20.20.20, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/10/20 ms CE1#ping 30.30.30.30 source 10.10.10.10 Sending 5, 100-byte ICMP Echoes to 30.30.30.30, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/10/20 ms
CE2	CE2#ping 10.10.10.10 source 20.20.20.20 Sending 5, 100-byte ICMP Echoes to 10.10.10.10, timeout is 2 seconds: < press Ctrl+C to break > !!!!!

	Success rate is 100 percent (5/5), round-trip min/avg/max = 1/10/20 ms CE2#ping 30.30.30.30 source 20.20.20.20 Sending 5, 100-byte ICMP Echoes to 30.30.30.30, timeout is 2 seconds: < press Ctrl+C to break > !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 1/10/20 ms

Common Errors

- The router ID does not consist of 32 bits, causing the failure to establish LDP and BGP neighbor relationships.

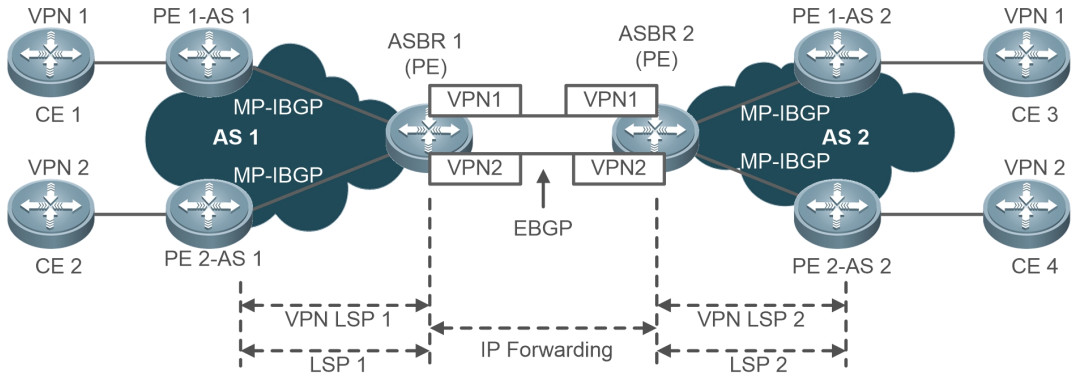
2.4.2 Configuring an Inter-AS VPN Service Model: Option A

Configuration Effect

Build the BGP/MPLS VPN service across ASs.

In the Option A solution, the ASBR in each AS creates a VRF table for every VPN that requires inter-AS communication and binds an interface to the VRF table. The VRF tables between ASBRs exchange VPN routes over the bound interfaces.

Figure 2- 26 VRF-to-VRF Inter-AS VPN



The Option A solution requires that ASBRs configure and bind an interface (a logical sub-interface) to every VPN that requires inter-AS communication. The number of bound interfaces must be at least equal to the number of the VPNs. In addition, VPNs must be configured one by one on the ASBRs. This causes scalability issues. The creation of sub-interfaces for individual VPNs also poses additional requirements on ASBRs. Therefore, the Option A solution is applicable to networks with few VPNs that require inter-AS communication.

Notes

- The LDP router ID must consist of 32 bits.
- The BGP router ID must consist of 32 bits.

- ASBRs must configure and bind an interface to every VPN that requires inter-AS communication.

Verification

- Verify that CE3 is pingable but CE2 and CE4 are unpingable from CE1.
- Verify that CE4 is pingable but CE1 and CE3 are unpingable from CE2.
- Verify that CE1 is pingable but CE2 and CE4 are unpingable from CE3.
- Verify that CE2 is pingable but CE1 and CE3 are unpingable from CE4.

Configuration Steps

For how to configure the Option A solution, see section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Common Errors

No routing protocol is enabled between PE and CE if no VRF table is bound to an interface using the **ip vrf forwarding** *vrf_name* command. As a result, the **show ip route vrf** *vrf_name* command output does not contain any route at the CE end.

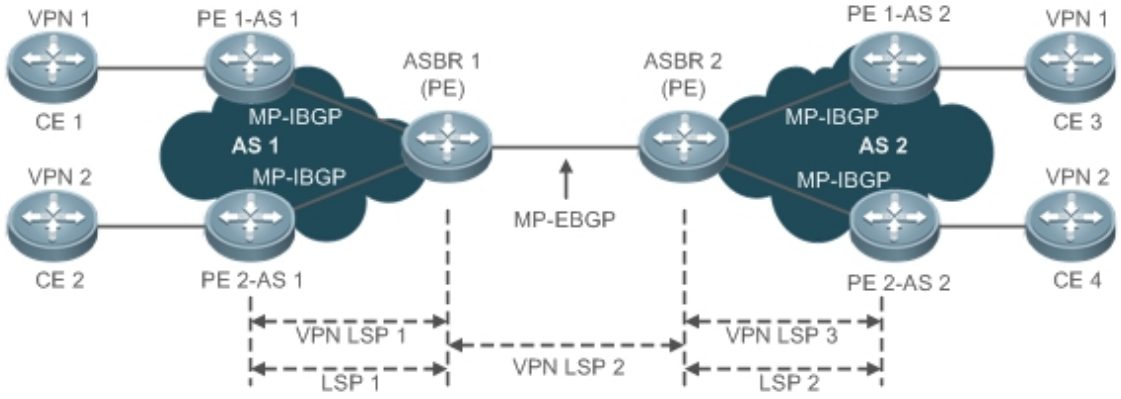
2.4.3 Configuring an Inter-AS VPN Service Model: Option B

Configuration Effect

Build the BGP/MPLS VPN service across ASs.

In the Option B solution, a single-hop MP-EBGP session is established between ASBRs for direct distribution of VPN routes.

Figure 2- 27 Option B Inter-AS VPN



The Option B solution removes the need to allocate sub-interfaces to sites of each VPN and the need to establish inter-AS LSPs. Instead, VPN routes are transmitted directly between single-hop MP-EBGP neighbors. Because VPN routes are stored and propagated by ASBRs between ASs, the ASBRs are under great burden when handling a large amount of VPN routes. The ASBRs also need to forward public IP packets. This poses high requirements on ASBRs. Because ASBRs disable the RT filter function for handling incoming VPN routes, the VPN routes on a PE may be propagated to the ASBR in a different AS, which may cause VPN route leakage. To address this problem, SPs that exchange VPN routes must implement a confidence-based routing protocol, enabling ASBRs to trust each other

and implement route filtering policies. The Option B solution is applicable to networks with many VPNs that require inter-AS communication.

Notes

- The LDP router ID and BGP router ID must consist of 32 bits.
- ASBRs establish a single-hop EBGP session to exchange VPN routes. ASBRs may or may not change the next hop in VPN routes sent to the peer.

Implementation

The Option B solution is implemented by the following methods:

- The ASBR retains the next hop in incoming VPN routes.
- The ASBR changes the next hop in incoming VPN routes.

The two methods are described in the following.

2.4.3.1 Method 1: The ASBR Retains the Next Hop in Incoming VPN Routes

When an ASBR receives a VPN route from the peer ASBR in a different AS, the ASBR sends the VPN route with the next hop unchanged to its MP-IBGP neighbors in the local AS. The PEs and the ASBR in the same AS exchange VPN routes over an MP-IBGP session, and two ASBRs exchange VPN routes directly over an MP-EBGP session. Because the ASBR does not change the next hop in the VPN route received from the MP-EBGP peer, the PEs (MP-IBGP neighbors) in the local AS must have the route destined for the next hop (an ASBR in the other AS). To do this, the ASBR redistributes this route to IGP in the local AS, in order to make the destination ASBR reachable. Then an LSP is established over LDP.

Configuration Steps

- (Mandatory) Configure PEs and CEs to exchange routes.
- (Mandatory) Configure IGP and MPLS signaling for ASs.
- (Mandatory) Disable the default RT filter function on ASBRs.
- (Mandatory) Configure the PEs and the ASBR in the same AS to exchange VPN routes.
- (Mandatory) Configure ASBRs to establish an MP-EBGP session.
- (Optional) Configure a route map for filtering VPN routes.
- (Mandatory) Configure IGP redistribution of the route destined for the ASBR in the other AS.

Configuring PEs and CEs to Exchange Routes

For details, see "Configuring PEs and CEs to Exchange Routes" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring IGP and MPLS Signaling for ASs

For details, see "Configuring an MPLS Network" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Disabling the Default RT Filter Function on ASBRs

By default, a PE rejects the VPN route received from the peer PE or an ASBR if the route is not imported to any of its VRF tables. To enable ASBRs to accept all VPN routes from other PEs or ASBRs without checking whether the VPN routes are imported to local VRF tables, you need to disable the RT filter function on ASBRs.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-number*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **no bgp default route-target filter**	Disables the RT filter function.
FS(config-router)# **show running-config**	Displays existing configurations.

Configuring the PEs and the ASBR in the Same AS to Exchange VPN Routes

For details, see "Configuring PEs to Exchange VPN Routes" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring ASBRs to Establish an MP-EBGP Session

Configure inter-AS ASBRs to establish a single-hop MP-EBGP session in order to distribute VPN routes.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-number*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *asbr-address* **remote-as** *asbr-asn-number*	Configures an EBGP session between ASBRs.
FS(config-router)# **address-family vpnv4**	Enters BGP VPN address family configuration mode.
FS(config-router-af)**# neighbor** *asbr-address* **activate**	Enables the capability of exchanging VPN routes between peers.
FS(config-router-af)# **show running-config**	Displays existing configurations.

To forward MPLS packets over the link between ASBRs, run the **label-switching** command to enable MPLS packet forwarding on the interfaces connected between the ASBRs.

If ASBRs use the loopback interface address with a 32-bit mask (instead of a direct link address) as the source address for establishing an MP-EBGP session, do the following: (1) run the **neighbor ebgp-multi-hop** command to enable multi-hop EGBP; (2) configure a loopback static route destined for the peer end between ASBRs; (3) enable LDP or configure static FTN (set the outgoing label to 3, indicating that the ASBR is the penultimate hop).

(Optional) Configuring a Route Map for Filtering VPN Routes

For AS security consideration, ASBRs are configured to send or receive only particular VPN routes based on specified policies. This security mechanism can be implemented using the RT extended community attribute for VPN route filtering. For ASBRs that disable the default RT filter function and therefore store all VPN routes, you can configure VPN routing policies to enable the ASBRs to receive only the inter-AS VPN routes from the PEs in the local AS, thus reducing the capacity burden on the ASBRs.

Configure filter policies in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **ip extcommunity-list standard** *extcommunity-name* | *extcommunity-number* {**permit|deny**} **rt** *rt-value*	Creates an extcommunity attribute list.
FS(config)# **show ip** extcommunity-list [*list-number*|*list-name*]	Verifies the configured extcommunity attribute list.
FS(config)# **route-map** *route-map-name* **permit** [*number*]	Creates a route map and enters route map configuration mode.
FS(config-route-map)# **match extcommunity** *extcommunity-name*|*extcommunity-number*	Configures RT matching rules for the route map.
FS(config-route-map)# **show route-map** *route-map-name*	Displays the route map.
FS(config-route-map)# **exit**	Exits route map configuration mode.
FS(config)# **router bgp** *as-num*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **address-family vpnv4**	Enters VPN address family configuration mode.
FS(config-router-af)# **neighbor** *peer-address* **route-map** *route-map-name* **in**	Filters the VPN routes received from the ASBR in the other AS.
FS(config-router-af)# **neighbor** *peer-address* **route-map** *route-map-name* **out**	Filters the VPN routes sent to the ASBR in the other AS.
FS(config-router-af)# **show running-config**	Displays existing configurations.

Configuring IGP Redistribution of the Route Destined for the ASBR in the Other AS

If the ASBR in one AS retains the next hop in incoming VPN routes, the IBGP peers (PEs) in the AS learn the VPN routes with the next hop still being the ASBR in the other AS. Therefore, the PEs must learn the route destined for the next hop. On the ASBRs that establish a single-hop MP-EBGP session enabled with BGP label distribution to IPv4 or VPN routes, MP-BGP automatically generates a direct host route (with a 32-bit mask) destined for the peer ASBR and an FTN entry (the outgoing label is 3), so that the tunnel egress is not terminated on the local ASBR. To allow the PEs to learn the route destined for the ASBR in the other AS, you only need to configure the ASBR in the local AS to redistribute the route to local IGP.

Perform configuration in privileged mode according to the following table.

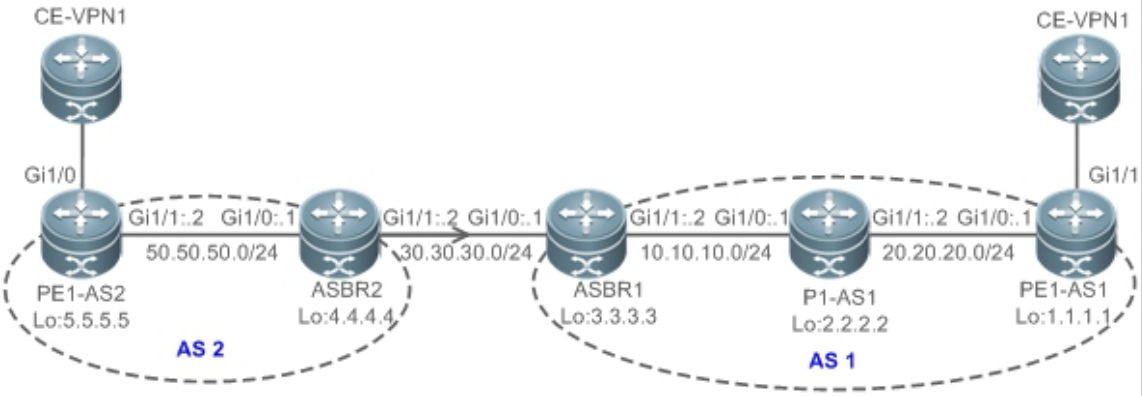
Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router** *igp*	Enters IGP (OSPF, RIP, or IS-IS) configuration mode.
FS(config-router)# **redistribute connected subnets**	Redistributes the direct segment route.
FS(config-router)# **show running-config**	Displays existing configurations.

Verification

- Verify that CE3 is pingable but CE2 and CE4 are unpingable from CE1.
- Verify that CE4 is pingable but CE1 and CE3 are unpingable from CE2.
- Verify that CE1 is pingable but CE2 and CE4 are unpingable from CE3.
- Verify that CE2 is pingable but CE1 and CE3 are unpingable from CE4.

Configuration Example

Inter-VPN Option B: Next Hop Unchanged

Figure 2- 28	 Requirements: A VPN has sites in two ASs. Enable inter-AS communication between the sites.
Configuration Steps	● On PE1-AS1, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure BGP, and establish an MP-IBGP session with ASBR1. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On PE1-AS2, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure BGP, and establish an MP-IBGP session with ASBR2. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On P1-AS1, configure a loopback interface, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol. ● On ASBR1, configure a loopback interface, configure BGP, and disable the BGP RT filter function. Establish a neighbor relationship with P1-AS1 and ASBR2 respectively. Configure MPLS signaling and enable MPLS on the interface connected to the public network. Configure OSPF for the backbone network to transmit routes, and redistribute the direct segment route. Configure the IP address of the interface connected to ASBR2, and enable label forwarding on the interface. ● On ASBR2, configure a loopback interface, configure BGP, and disable the BGP RT filter function. Establish a neighbor relationship with PE1-AS2 and ASBR1 respectively. Configure MPLS signaling and enable MPLS on the interface connected to the public network. Configure OSPF for the backbone network to transmit routes, and redistribute the direct segment route. Configure the IP address of the interface connected to ASBR1, and enable label forwarding on the interface.
PE1-AS1	PE1-AS1# configure terminal PE1-AS1(config)# interface loopback 0 PE1-AS1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1-AS1# configure terminal

```
PE1-AS1(config)# ip vrf VPN1
PE1-AS1(config-vrf)# rd 1:100
PE1-AS1(config-vrf)# route-target both 1:100
PE1-AS1(config-vrf)# end
PE1-AS1# configure terminal
PE1-AS1(config)# interface GigabitEthernet 1/1
PE1-AS1(config-if-GigabitEthernet 1/1)# ip vrf forwarding VPN1
PE1-AS1(config-if-GigabitEthernet 1/1)# ip address 192.168.16.2 255.255.255.0
PE1-AS1(config-if-GigabitEthernet 1/1)# end
PE1-AS1# configure terminal
PE1-AS1(config)# router bgp 1
PE1-AS1(config-router)# neighbor 3.3.3.3 remote-as 1
PE1-AS1(config-router)# neighbor 3.3.3.3 update-source loopback 0
PE1-AS1(config-router)# address-family vpnv4
PE1-AS1(config-router-af)# neighbor 3.3.3.3 activate
PE1-AS1(config-router-af)# end
PE1-AS1# configure terminal
PE1-AS1(config)# router bgp 1
PE1-AS1(config)# address-family ipv4 vrf VPN1
PE1-AS1(config-router-af)# neighbor 192.168.16.1 remote-as 65002
PE1-AS1(config-router-af)# end
PE1-AS1# configure terminal
PE1-AS1(config)# mpls ip
PE1-AS1(config)# mpls router ldp
PE1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1-AS1(config-mpls-router)# exit
PE1-AS1(config)# interface GigabitEthernet 1/0
PE1-AS1(config-if-GigabitEthernet 1/0)# ip address 20.20.20.1 255.255.255.0
PE1-AS1(config-if-GigabitEthernet 1/0)# label-switching
PE1-AS1(config-if-GigabitEthernet 1/0)# mpls ip
PE1-AS1(config-if-GigabitEthernet 1/0)# end
PE1-AS1# configure terminal
```

	PE1-AS1(config)# router ospf 10 PE1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 PE1-AS1(config-router)# network 1.1.1.1 0.0.0.0 area 0 PE1-AS1(config-router)# end
PE1-AS2	PE1-AS2# configure terminal PE1-AS2(config)# interface loopback 0 PE1-AS2(config-if-Loopback 0)# ip address 5.5.5.5 255.255.255.255 PE1-AS2# configure terminal PE1-AS2(config)# ip vrf VPN1 PE1-AS2(config-vrf)# rd 1:100 PE1-AS2(config-vrf)# route-target both 1:100 PE1-AS2(config-vrf)# end PE1-AS2# configure terminal PE1-AS2(config)# interface GigabitEthernet 1/0 PE1-AS2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPN1 PE1-AS2(config-if-GigabitEthernet 1/0)# ip address 192.168.15.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/0)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config-router)# neighbor 4.4.4.4 remote-as 1 PE1-AS2(config-router)# neighbor 4.4.4.4 update-source loopback 0 PE1-AS2(config-router)# address-family vpnv4 PE1-AS2(config-router-af)# neighbor 4.4.4.4 activate PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config)# address-family ipv4 vrf VPN1 PE1-AS2(config-router-af)# neighbor 192.168.15.1 remote-as 65002 PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# mpls ip

	PE1-AS2(config)# mpls router ldp PE1-AS2(config-mpls-router)# ldp router-id interface loopback 0 force PE1-AS2(config-mpls-router)# exit PE1-AS2(config)# interface GigabitEthernet 1/1 PE1-AS2(config-if-GigabitEthernet 1/1)# ip address 50.50.50.2.1 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/1)# label-switching PE1-AS2(config-if-GigabitEthernet 1/1)# mpls ip PE1-AS2(config-if-GigabitEthernet 1/1)# end PE1-AS2# configure terminal PE1-AS2(config)# router ospf 10 PE1-AS2(config-router)# network 50.50.50.0 0.0.0.255 area 0 PE1-AS2(config-router)# network 5.5.5.5 0.0.0.0 area 0 PE1-AS2(config-router)# end
P1-AS1	P1-AS1# configure terminal P1-AS1(config)# interface loopback 0 P1-AS1(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 P1-AS1# configure terminal P1-AS1(config)# mpls ip P1-AS1(config)# mpls router ldp P1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force P1-AS1(config-mpls-router)# exit P1-AS1(config)# interface GigabitEthernet 1/0 P1-AS1(config-if-GigabitEthernet 1/0)# ip address 10.10.10.1 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/0)# label-switching P1-AS1(config-if-GigabitEthernet 1/0)# mpls ip P1-AS1(config-if-GigabitEthernet 1/0)# exit P1-AS1(config)# interface GigabitEthernet 1/1 P1-AS1(config-if-GigabitEthernet 1/1)# ip address 20.20.20.2 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/1)# label-switching P1-AS1(config-if-GigabitEthernet 1/1)# mpls ip P1-AS1# configure terminal

	P1-AS1(config)# router ospf 10 P1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 P1-AS1(config-router)# network 10.10.10.0 0.0.0.255 area 0 P1-AS1(config-router)# network 2.2.2.2 0.0.0.0 area 0 P1-AS1(config-router)# end
ASBR1	ASBR1# configure terminal ASBR1(config)# interface loopback 0 ASBR1(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 ASBR1# configure terminal ASBR1(config)# router bgp 1 ASBR1(config-router)# no bgp default route-target filter ASBR1(config-router)# neighbor 1.1.1.1 remote-as 1 ASBR1(config-router)# neighbor 1.1.1.1 update-source loopback 0 ASBR1(config-router)# neighbor 30.30.30.2 remote-as 2 ASBR1(config-router)# address-family vpnv4 unicast ASBR1(config-router-af)# neighbor 1.1.1.1 activate ASBR1(config-router-af)# neighbor 30.30.30.2 activate ASBR1(config-router-af)# end ASBR1# configure terminal ASBR1(config)# mpls ip ASBR1(config)# mpls router ldp ASBR1(config-mpls-router)# ldp router-id interface loopback 0 force ASBR1(config-mpls-router)# exit ASBR1(config)# interface GigabitEthernet 1/1 ASBR1(config-if-GigabitEthernet 1/1)# ip address 10.10.10.2 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/1)# label-switching ASBR1(config-if-GigabitEthernet 1/1)# mpls ip ASBR1(config-if-GigabitEthernet 1/1)# end ASBR1# configure terminal ASBR1(config)# router ospf 10 ASBR1(config-router)# network 10.10.10.0 0.0.0.255 area 0

	ASBR1(config-router)# network 3.3.3.3 0.0.0.0 area 0 ASBR1(config-router)# redistribute connected subnets ASBR1(config-router)# end ASBR1(config)# interface GigabitEthernet 1/0 ASBR1(config-if-GigabitEthernet 1/0)# ip address 30.30.30.1 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/0)# label-switching
ASBR2	ASBR2# configure terminal ASBR2(config)# interface loopback 0 ASBR2(config-if-Loopback 0)# ip address 4.4.4.4 255.255.255.255 ASBR2# configure terminal ASBR2(config)# router bgp 1 ASBR2(config-router)# no bgp default route-target filter ASBR2(config-router)# neighbor 5.5.5.5 remote-as 1 ASBR2(config-router)# neighbor 5.5.5.5 update-source loopback 0 ASBR2(config-router)# neighbor 30.30.30.1 remote-as 2 ASBR2(config-router)# address-family vpnv4 unicast ASBR2(config-router-af)# neighbor 5.5.5.5 activate ASBR2(config-router-af)# neighbor 30.30.30.1 activate ASBR2(config-router-af)# end ASBR2# configure terminal ASBR2(config)# mpls ip ASBR2(config)# mpls router ldp ASBR2(config-mpls-router)# ldp router-id interface loopback 0 force ASBR2(config-mpls-router)# exit ASBR2(config)# interface GigabitEthernet 1/1 ASBR2(config-if-GigabitEthernet 1/1)# ip address 30.30.30.2 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/1)# label-switching ASBR2(config-if-GigabitEthernet 1/1)# mpls ip ASBR2(config-if-GigabitEthernet 1/1)# end ASBR2# configure terminal ASBR2(config)# router ospf 10

	ASBR2(config-router)# network 50.50.50.0 0.0.0.255 area 0 ASBR2(config-router)# network 4.4.4.4 0.0.0.0 area 0 ASBR2(config-router)# redistribute connected subnets ASBR2(config-router)# end ASBR2(config)# interface GigabitEthernet 1/0 ASBR2(config-if-GigabitEthernet 1/0)# ip address 50.50.50.1 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/0)# label-switching
Verification	● Run the **show mpls forwarding-table** command to display MPLS VPN routes. ● Check that sites of the same VPN can communicate with each other but sites of different VPNs cannot communicate with each other.
PE1-AS1	Assume that the loopback interface on the peer CE is 20.20.20.20. VPN routes are pingable, whereas ASBR2 and PE1-AS2 are unpingable. That is: 20.20.20.20 is pingable from the local CE. VRF VPN1 20.20.20.20 is pingable. 3.3.3.3 is pingable. 4.4.4.4 is unpingable. 5.5.5.5 is unpingable.
PE1-AS2	Assume that the loopback interface on the peer CE is 10.10.10.10. VPN routes are pingable, whereas ASBR1 and PE1-AS1 are unpingable. That is: 10.10.10.10 is pingable from the local CE. VRF VPN1 10.10.10.10 is pingable. 3.3.3.3 is unpingable. 4.4.4.4 is pingable. 5.5.5.5 is unpingable.
P1-AS1	PE1-AS1 and ASBR1 are pingable. That is: 1.1.1.1 is pingable. 3.3.3.3 is pingable.
ASBR1	PE1-AS1 is pingable. That is: 1.1.1.1 is pingable. 10.10.10.10 is unpingable.

	20.20.20.20 is unpingable.
ASBR2	PE1-AS2 is pingable. That is: 5.5.5.5 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable.

Common Errors

- LDP sessions fail to be established, or the BGP RT filter function is not disabled. The direct subnet route is not redistributed to OSPF (that is, the **redistribute connected subnets** command is not executed); as a result, sites of CE-VPN1 are unpingable between each other.

2.4.3.2 Method 2: The ASBR Changes the Next Hop in Incoming VPN Routes

When an ASBR receives a VPN route from the peer ASBR in a different AS, the ASBR changes the next hop in the incoming VPN route before sending the route to PEs in the local AS. The PEs and the ASBR in the same AS exchange VPN routes over an MP-IBGP session. The two ASBRs in different ASs exchange VPN routes over an MP-EBGP session. When receiving a VPN route from one ASBR, the other ASBR changes the route's next hop to itself before sending the route to MP-IBGP peers in the local AS.

Configuration Steps

- (Mandatory) Configure PEs and CEs to exchange routes.
- (Mandatory) Configure IGP and MPLS signaling for ASs.
- (Mandatory) Disable the default RT filter function on ASBRs.
- (Mandatory) Configure the PEs and the ASBR in the same AS to establish an MP-IBGP session and enable the ASBR to change the next hop in incoming VPN routes to itself.
- (Mandatory) Configure ASBRs to establish an MP-EBGP session.
- (Optional) Configure a route map for filtering VPN routes.

Configuring PEs and CEs to Exchange Routes

For details, see "Configuring PEs and CEs to Exchange Routes" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring IGP and MPLS Signaling for ASs

For details, see "Configuring an MPLS Network" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Disabling the Default RT Filter Function on ASBRs

For details, see "Disabling the Default RT Filter Function on ASBRs" in section 2.4.3.1 "Method 1: The ASBR Retains the Next Hop in Incoming VPN Routes."

Configuring the PEs and the ASBR in the Same AS to Establish an MP-IBGP Session and Enabling the ASBR to Change the Next Hop in Incoming VPN Routes to Itself

By default, when receiving a VPN route from the MP-EBGP peer, an ASBR does not change the next hop in the VPN route but sends the route as it is to MP-IBGP peers. You can configure the ASBR to change the next hop to itself. In this way, the PEs in the local AS do not need to know the address of the ASBR in the other AS. Compared with Method 1, Method 2 requires the additional configuration of next hop change.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-num*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *pe-address* **remote-as** *asn-num*	Establishes an IBGP session with PEs.
FS(config-router)# **neighbor** *pe-address* **update-source** *interface-name*	Specifies the local loopback interface as the source address for the IBGP session.
FS(config-router)#**address-family vpnv4**	Enters BGP VPN address family configuration mode.
FS(config-router-af)# **neighbor** *pe-address* **activate**	Enables the capability of exchanging VPN routes between peers.
FS(config-router-af)# **neighbor** *pe-address* **next-hop-self**	Configures the ASBR to change the next hop in incoming VPN routes to itself before sending the routes to IBGP peers.
FS(config-router-af)# **show running-config**	Displays existing configurations.

Configuring ASBRs to Establish an MP-EBGP Session

For details, see "Configuring ASBRs to Establish an MP-EBGP Session" in section 2.4.3.1 "Method 1: The ASBR Retains the Next Hop in Incoming VPN Routes."

(Optional) Configuring a Route Map for Filtering VPN Routes

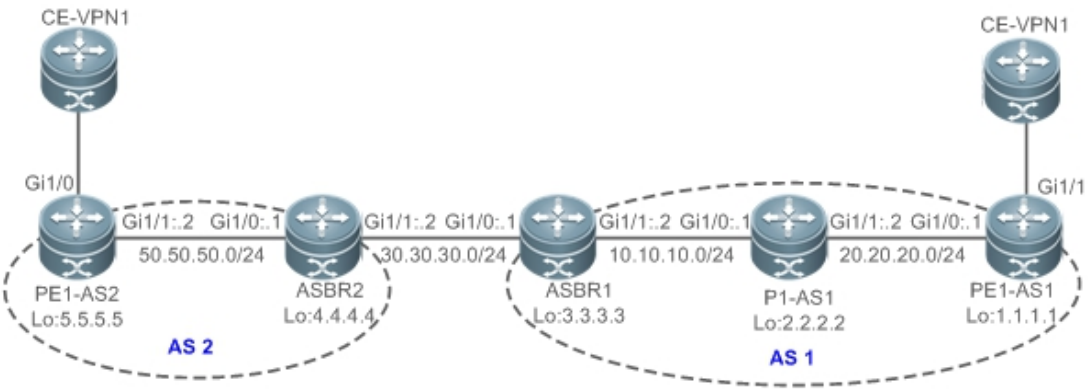
For details, see "Configuring a Route Map for Filtering VPN Routes" in section 2.4.3.1 "Method 1: The ASBR Retains the Next Hop in Incoming VPN Routes."

Verification

Run the ping command to check the connection between sites of the same VPN.

Configuration Example

Inter-VPN Option B: Next Hop Changed

Scenario **Figure 2-29**	 Requirements: A VPN has sites in two ASs. Enable inter-AS communication between the sites.
Configuration Steps	● On PE1-AS1, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure BGP, and establish an MP-IBGP session with ASBR1. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On PE1-AS2, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure BGP, and establish an MP-IBGP session with ASBR2. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On P1-AS1, configure a loopback interface, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol. ● On ASBR1, configure a loopback interface, configure BGP, and disable the BGP RT filter function. Establish a neighbor relationship with PE1-AS1 and ASBR2 respectively. Configure MPLS signaling and enable MPLS on the interface connected to the public network. Configure OSPF for the backbone network to transmit routes, and redistribute the direct segment route. Configure the IP address of the interface connected to ASBR2, and enable label forwarding on the interface. Configure ASBR1 to change the next hop in VPN routes to itself before sending the routes to IBGP neighbors. ● On ASBR2, configure a loopback interface, configure BGP, and disable the BGP RT filter function. Establish neighbor relationships with PE1-AS2 and ASBR1. Configure MPLS signaling and enable MPLS on the interface connected to the public network. Configure OSPF for the backbone network to transmit routes, and redistribute the direct segment route. Configure the IP address of the interface connected to ASBR1, and enable label forwarding on the interface. Configure ASBR2 to change the next hop in VPN routes to itself before sending the routes to IBGP neighbors.
PE1-AS1	PE1-AS1# configure terminal PE1-AS1(config)# interface loopback 0 PE1-AS1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1-AS1# configure terminal

```
PE1-AS1(config)# ip vrf VPN1
PE1-AS1(config-vrf)# rd 1:100
PE1-AS1(config-vrf)# route-target both 1:100
PE1-AS1(config-vrf)# end
PE1-AS1# configure terminal
PE1-AS1(config)# interface GigabitEthernet 1/1
PE1-AS1(config-if-GigabitEthernet 1/1)# ip vrf forwarding VPN1
PE1-AS1(config-if-GigabitEthernet 1/1)# ip address 192.168.16.2 255.255.255.0
PE1-AS1(config-if-GigabitEthernet 1/1)# end
PE1-AS1# configure terminal
PE1-AS1(config)# router bgp 1
PE1-AS1(config-router)# neighbor 3.3.3.3 remote-as 1
PE1-AS1(config-router)# neighbor 3.3.3.3 update-source loopback 0
PE1-AS1(config-router)# address-family vpnv4
PE1-AS1(config-router-af)# neighbor 3.3.3.3 activate
PE1-AS1(config-router-af)# end
PE1-AS1# configure terminal
PE1-AS1(config)# router bgp 1
PE1-AS1(config)# address-family ipv4 vrf VPN1
PE1-AS1(config-router-af)# neighbor 192.168.16.1 remote-as 65002
PE1-AS1(config-router-af)# end
PE1-AS1# configure terminal
PE1-AS1(config)# mpls ip
PE1-AS1(config)# mpls router ldp
PE1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1-AS1(config-mpls-router)# exit
PE1-AS1(config)# interface GigabitEthernet 1/0
PE1-AS1(config-if-GigabitEthernet 1/0)# ip address 20.20.20.1 255.255.255.0
PE1-AS1(config-if-GigabitEthernet 1/0)# label-switching
PE1-AS1(config-if-GigabitEthernet 1/0)# mpls ip
PE1-AS1(config-if-GigabitEthernet 1/0)# end
PE1-AS1# configure terminal
```

	PE1-AS1(config)# router ospf 10 PE1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 PE1-AS1(config-router)# network 1.1.1.1 0.0.0.0 area 0 PE1-AS1(config-router)# end
PE1-AS2	PE1-AS2# configure terminal PE1-AS2(config)# interface loopback 0 PE1-AS2(config-if-Loopback 0)# ip address 5.5.5.5 255.255.255.255 PE1-AS2# configure terminal PE1-AS2(config)# ip vrf VPN1 PE1-AS2(config-vrf)# rd 1:100 PE1-AS2(config-vrf)# route-target both 1:100 PE1-AS2(config-vrf)# end PE1-AS2# configure terminal PE1-AS2(config)# interface GigabitEthernet 1/0 PE1-AS2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPN1 PE1-AS2(config-if-GigabitEthernet 1/0)# ip address 192.168.15.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/0)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config-router)# neighbor 4.4.4.4 remote-as 1 PE1-AS2(config-router)# neighbor 4.4.4.4 update-source loopback 0 PE1-AS2(config-router)# address-family vpnv4 PE1-AS2(config-router-af)# neighbor 4.4.4.4 activate PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config)# address-family ipv4 vrf VPN1 PE1-AS2(config-router-af)# neighbor 192.168.15.1 remote-as 65002 PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# mpls ip

	PE1-AS2(config)# mpls router ldp PE1-AS2(config-mpls-router)# ldp router-id interface loopback 0 force PE1-AS2(config-mpls-router)# exit PE1-AS2(config)# interface GigabitEthernet 1/1 PE1-AS2(config-if-GigabitEthernet 1/1)# ip address 50.50.50.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/1)# label-switching PE1-AS2(config-if-GigabitEthernet 1/1)# mpls ip PE1-AS2(config-if-GigabitEthernet 1/1)# end PE1-AS2# configure terminal PE1-AS2(config)# router ospf 10 PE1-AS2(config-router)# network 50.50.50.0 0.0.0.255 area 0 PE1-AS2(config-router)# network 5.5.5.5 0.0.0.0 area 0 PE1-AS2(config-router)# end
P1-AS1	P1-AS1# configure terminal P1-AS1(config)# interface loopback 0 P1-AS1(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 P1-AS1# configure terminal P1-AS1(config)# mpls ip P1-AS1(config)# mpls router ldp P1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force P1-AS1(config-mpls-router)# exit P1-AS1(config)# interface GigabitEthernet 1/0 P1-AS1(config-if-GigabitEthernet 1/0)# ip address 10.10.10.1 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/0)# label-switching P1-AS1(config-if-GigabitEthernet 1/0)# mpls ip P1-AS1(config-if-GigabitEthernet 1/0)# exit P1-AS1(config)# interface GigabitEthernet 1/1 P1-AS1(config-if-GigabitEthernet 1/1)# ip address 20.20.20.2 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/1)# label-switching P1-AS1(config-if-GigabitEthernet 1/1)# mpls ip P1-AS1# configure terminal

	P1-AS1(config)# router ospf 10 P1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 P1-AS1(config-router)# network 10.10.10.0 0.0.0.255 area 0 P1-AS1(config-router)# network 2.2.2.2 0.0.0.0 area 0 P1-AS1(config-router)# end
ASBR1	ASBR1# configure terminal ASBR1(config)# interface loopback 0 ASBR1(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 ASBR1# configure terminal ASBR1(config)# router bgp 1 ASBR1(config-router)# no bgp default route-target filter ASBR1(config-router)# neighbor 1.1.1.1 remote-as 1 ASBR1(config-router)# neighbor 1.1.1.1 update-source loopback 0 ASBR1(config-router)# neighbor 30.30.30.2 remote-as 2 ASBR1(config-router)# address-family vpnv4 unicast ASBR1(config-router-af)# neighbor 1.1.1.1 activate ASBR1(config-router-af)# neighbor 1.1.1.1 next-hop-self ASBR1(config-router-af)# neighbor 30.30.30.2 activate ASBR1(config-router-af)# end ASBR1# configure terminal ASBR1(config)# mpls ip ASBR1(config)# mpls router ldp ASBR1(config-mpls-router)# ldp router-id interface loopback 0 force ASBR1(config-mpls-router)# exit ASBR1(config)# interface GigabitEthernet 1/1 ASBR1(config-if-GigabitEthernet 1/1)# ip address 10.10.10.2 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/1)# label-switching ASBR1(config-if-GigabitEthernet 1/1)# mpls ip ASBR1(config-if-GigabitEthernet 1/1)# end ASBR1# configure terminal ASBR1(config)# router ospf 10

	ASBR1(config-router)# network 10.10.10.0 0.0.0.255 area 0 ASBR1(config-router)# network 3.3.3.3 0.0.0.0 area 0 ASBR1(config-router)# end ASBR1(config)# interface GigabitEthernet 1/0 ASBR1(config-if-GigabitEthernet 1/0)# ip address 30.30.30.1 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/0)# label-switching
ASBR2	ASBR2# configure terminal ASBR2(config)# interface loopback 0 ASBR2(config-if-Loopback 0)# ip address 4.4.4.4 255.255.255.255 ASBR2# configure terminal ASBR2(config)# router bgp 1 ASBR2(config-router)# no bgp default route-target filter ASBR2(config-router)# neighbor 5.5.5.5 remote-as 1 ASBR2(config-router)# neighbor 5.5.5.5 update-source loopback 0 ASBR2(config-router)# neighbor 30.30.30.1 remote-as 1 ASBR2(config-router)# address-family vpnv4 unicast ASBR2(config-router-af)# neighbor 5.5.5.5 activate ASBR2(config-router-af)# neighbor 5.5.5.5 next-hop-self ASBR2(config-router-af)# neighbor 30.30.30.1 activate ASBR2(config-router-af)# end ASBR2# configure terminal ASBR2(config)# mpls ip ASBR2(config)# mpls router ldp ASBR2(config-mpls-router)# ldp router-id interface loopback 0 force ASBR2(config-mpls-router)# exit ASBR2(config)# interface GigabitEthernet 1/1 ASBR2(config-if-GigabitEthernet 1/1)# ip address 30.30.30.2 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/1)# label-switching ASBR2(config-if-GigabitEthernet 1/1)# mpls ip ASBR2(config-if-GigabitEthernet 1/1)# end ASBR2# configure terminal

	ASBR2(config)# router ospf 10 ASBR2(config-router)# network 50.50.50.0 0.0.0.255 area 0 ASBR2(config-router)# network 4.4.4.4 0.0.0.0 area 0 ASBR2(config-router)# end ASBR2(config)# interface GigabitEthernet 1/0 ASBR2(config-if-GigabitEthernet 1/0)# ip address 50.50.50.1 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/0)# label-switching
Verification	● Check that sites of the same VPN can communicate with each other but sites of different VPNs cannot communicate with each other.
PE1-AS1	Assume that the loopback interface on the peer CE is 20.20.20.20. VPN routes are pingable, whereas ASBR2 and PE1-AS2 are unpingable. That is: 20.20.20.20 is pingable from the local CE. VRF VPN1 20.20.20.20 is pingable. 3.3.3.3 is pingable. 4.4.4.4 is unpingable. 5.5.5.5 is unpingable.
PE1-AS2	Assume that the loopback interface on the peer CE is 10.10.10.10. VPN routes are pingable, whereas ASBR1 and PE1-AS1 are unpingable. That is: 10.10.10.10 is pingable from the local CE. VRF VPN1 10.10.10.10 is pingable. 3.3.3.3 is unpingable. 4.4.4.4 is pingable. 5.5.5.5 is unpingable.
P1-AS1	PE1-AS1 and ASBR1 are pingable. That is: 1.1.1.1 is pingable. 3.3.3.3 is pingable.
ASBR1	PE1-AS1 is pingable. That is: 1.1.1.1 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable.

ASBR2	PE1-AS2 is pingable. That is: 5.5.5.5 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable.

Common Errors

- LDP sessions fail to be established, or the BGP RT filter function is not disabled. The ASBRs do not change the next hop in incoming VPN routes to itself before sending the routes to IBGP neighbors; as a result, sites of CE-VPN1 are unpingable between each other.

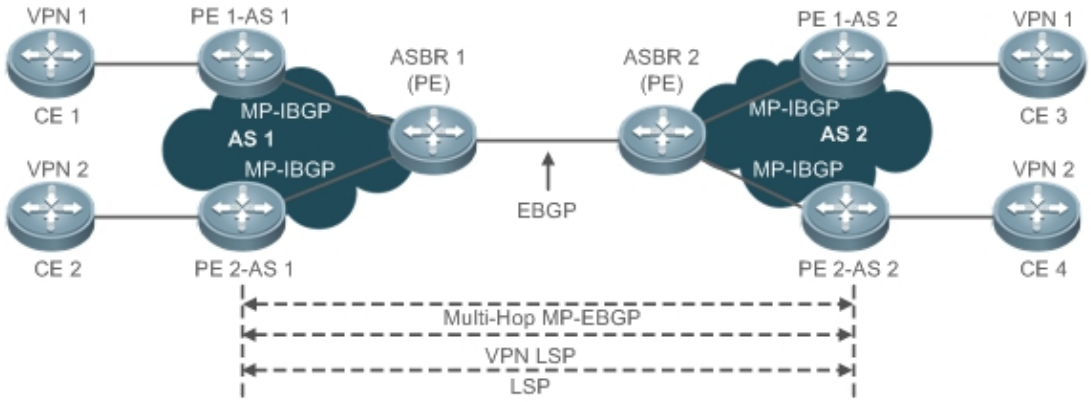
2.4.4 Configuring an Inter-AS VPN Service Model: Option C

Configuration Effect

Build the BGP/MPLS VPN service across ASs.

In the Option C solution, PEs in different ASs establish a multi-hop MP-EBGP session to directly exchange VPN routes.

Figure 2- 5 Multi-hop MP-EBGP



In multi-hop MP-EBGP mode, only PEs maintain VPN routes, whereas ASBRs do not store VPN user information. The configuration procedure is relatively complex. The solution is applicable to inter-AS VPN service provision in large scale.

Notes

- The LDP router ID must consist of 32 bits.
- The BGP router ID must consist of 32 bits.

Implementation

The Option C solution is implemented by the following methods:

- Enable IPv4 label distribution only between EBGP neighbors.

- Enable IPv4 label distribution between EBGP neighbors and between IBGP neighbors.
- Establish a multi-hop MP-EBGP session between RRs.

For Option C scalability, an RR is deployed in each AS, and the RRs between two ASs establish a multi-hop MP-EBGP session to exchange VPN routes. From the deployment perspective, Option C is also called the inter-RR multi-hop MP-EBGP solution.

The configuration of each Option C implementation method is described in the following.

2.4.4.1 Method 1: Enable IPv4 Label Switching Only Between EBGP Neighbors

In this implementation method, BGP routes must be redistributed using the IGP protocol (for example, OSPF or RIP) on ASBRs, so that each device in one AS has the route destined for each PE in the other AS. LSPs are established by distributing labels over LDP to the routes destined for PEs between ASs. IPv4 label switching is enabled between the directly connected inter-AS ASBRs. BGP assumes the role of MPLS signaling to distribute labels to the routes destined for PEs between ASs for LSP establishment.

Configuration Steps

- (Mandatory) Configure PEs and CEs in each AS to exchange routes.
- (Mandatory) Configure IGP and MPLS signaling for ASs.
- (Mandatory) Configure ASBRs to establish an EBGP session in order to distribute labels to IPv4 routes.
- (Mandatory) Configure ASBRs to redistribute to IGP the routes of PEs that are learned over EBGP.
- (Mandatory) Configure a multi-hop MP-EBGP session.

Configuring PEs and CEs to Exchange Routes

For details, see "Configuring PEs and CEs to Exchange Routes" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring IGP and MPLS Signaling for ASs

For details, see "Configuring an MPLS Network" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring ASBRs to Establish an EBGP Session in Order to Distribute Labels to IPv4 Routes

Establish an EBGP session between inter-AS ASBRs, and enable IPv4 label distribution. To introduce the routes of PEs to BGP, you can use the **network** command in BGP IPv4 address family configuration mode or the IGP redistribution command. For AS security consideration, IPv4 route distribution policies must be configured on ASBRs. You can define a route map to specify which routes are distributed to neighbors and whether the outgoing routes carry a label. Route maps can also be configured at the receive end.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-num*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *asbr-address* **remote-as** *asbr-asn-num*	Establishes an EBGP session between ASBRs.
FS(config-router)#**address-family ipv4**	Enters BGP IPv4 address family configuration mode.

Command	Description
FS(config-router-af)# **neighbor** *asbr-address* **send-label**	Configures the function of exchanging labeled IPv4 routes with the peer ASBR in the other AS.
FS(config-router-af**)# network** *pe-address* **mask** *mask*	(Optional) Configures the addresses of the PEs in the local AS that require BGP routing table injection, that is, the host routes destined for the PEs in the local AS.
FS(config-router-af)# **neighbor** *asbr-address* **route-map** *routemap-name* **out**	(Optional) Configures route distribution policies. You can define a route map to specify which routes are distributed to neighbors and whether the outgoing routes carry a label.
FS(config-router-af)# **neighbor** *asbr-address* **route-map** *routemap-name* **in**	(Optional) Configures route distribution policies. You can define a route map to specify reception of only labeled routes.
FS(config-router-af)# **show running-config**	Displays existing configurations.

To forward MPLS packets over the link between ASBRs, run the **label-switching** command to enable MPLS packet forwarding on the interfaces connected between the ASBRs.

If ASBRs use the loopback interface address with a 32-bit mask (instead of a direct link address) as the source address for establishing an MP-EBGP session, do the following: (1) run the **neighbor ebgp-multi-hop** command to enable multi-hop EGBP; (2) configure a loopback static route destined for the peer end between ASBRs; (3) enable LDP or configure static FTN (set the outgoing label to 3, indicating that the ASBR is the penultimate hop).

In actual application, the ASBRs only need to distribute labels to the routes of PEs that carry the inter-AS VPN service. You can use the **set mpls-label** rule to enable such label distribution. This rule means setting a label for routes. You can create a route map to specify sending of only the inter-AS routes of PEs to the peer ASBR and set labels for those routes, and then run the **neighbor** *peer-address* **route-map** *rmap_name* **out** command in BGP IPv4 address family configuration mode to associate the route map.

The following example creates a route map, whereby MPLS labels are distributed to routes with the 1.1.1.1/32 prefix whereas ordinary IPv4 route updates instead of labels are distributed to routes with the 1.1.1.2/32 prefix. For routes that do not match ACL1 and ACL2, route updates are not sent to neighbors.

```
Router(config)# ip access-list standard acl1
Router(config-std-nacl)# permit host 1.1.1.1
Router(config-std-nacl)# exit
Router(config)# ip access-list standard acl2
Router(config-std-nacl)# permit host 1.1.1.2
Router(config-std-nacl)# exit
Router(config)# route-map out-as permit 10
Router(config-route-map)# match ip address acl1
Router(config-route-map)# set mpls-label
Router(config-std-nacl)# exit
Router(config)# route-map out-as permit 20
```

```
Router(config-route-map)# match ip address acl2
Router(config)# router bgp 100
Router(config-router)# neighbor 30.30.30.2 remote-as 100
Router(config-router)# neighbor 30.30.30.2 route-map out-as out
```

You can use the **match mpls-label** rule to specify reception of only labeled IPv4 routes. Configure a route map, and run the **neighbor** *peer-address* **route-map** *rmap_name* **in** command to associate the route map.

The following example creates a route map to specify reception of only labeled IPv4 routes from the BGP peer 30.30.30.2 while rejecting other routes.

```
Router(config)# route-map match-mpls
Router(config-route-map)# match mpls-label
Router(config)# router bgp 100
Router(config-router)# neighbor 30.30.30.2 remote-as 100
Router(config-router)# neighbor 30.30.30.2 route-map match-mpls in
```

Configuring ASBRs to Redistribute to IGP the Routes of PEs That Are Learned over EBGP

When an ASBR learns from the peer ASBR the routes of PEs in the other AS, it needs to send the routes to the PEs in the local AS and establish an LSP destined for the PEs in the other AS. To do this, configure ASBRs to redistribute to IGP the routes learned over EBGP, and enable LDP for distributing labels to BGP routes, in order to establish an LSP destined for the PEs in the other AS.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router** *igp*	Enters IGP configuration mode.
FS(config-router)# **redistribute bgp subnets** [**route-map** *routemap-name*]	Configures the protocol for BGP route redistribution. Route-map-based filtering is optional.
FS(config-router)# **exit**	Exits IGP configuration mode.
FS(config)# **mpls router ldp**	Enters LDP configuration mode.
FS(config-mpls-router)# **ldp router-id interface** *loopback 0* **force**	Configures the LDP router ID, which typically uses the IP address of a loopback interface.
FS(config-mpls-router)# **advertise-labels for bgp-routes** [**acl** *acl-name*]	Distributes labels to BGP routes. ACL-based filtering is optional.
FS(config-mpls-router)# **show running-config**	Displays existing configurations.

By default, LDP only distributes labels to IGP routes but not to BGP routes. To distribute labels to BGP routes, run the **advertise-labels for bgp-routes** command.

In Method 1 of the Option C solution, you can run the **redistribute bgp subnets route-map** *routemap-name* command in IGP configuration mode to specify which BGP-learned routes are redistributed to IGP. You can run the **advertise-labels for bgp-routes acl** *acl-name* command in LDP configuration mode to specify which BGP routes are distributed with a label.

Configuring a Multi-hop MP-EBGP Session

After an inter-AS LSP is established, establish a multi-hop MP-EBGP session between the two PEs in different ASs that need to carry the inter-AS VPN service, in order to allow the PEs to exchange VPN routes.

Configure the multi-hop MP-EBGP session in privileged mode according to the following table.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-num*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *ebgp-peer-address* **remote-as** *ebgp-asn-num*	Establishes a multi-hop EBGP session with the PE in the other AS.
FS(config-router)# **neighbor** *ebgp-peer-address* **update-source** *interface-name*	Specifies the loopback interface used to establish a neighbor relationship with the EBGP peer.
FS(config-router)# **neighbor** *ebgp-peer-address* **ebgp-multi-hop**	Configures multi-hop attributes.
FS(config-router)#**address-family vpnv4**	Enters BGP VPN address family configuration mode.
FS(config-router-af)# **neighbor** *ebgp-peer-address* **activate**	Enables the capability of exchanging VPN routes between peers.
FS(config-router-af)# **exit**	Exits BGP VPN address family configuration mode.
FS(config-router)# **address-family ipv4**	Enters BGP IPv4 address family configuration mode.
FS(config-router-af)# **no neighbor** *ebgp-peer-address* **activate**	Disables the capability of exchanging IPv4 routes.
FS(config-router)# **show running-config**	Displays existing configurations.

Avoid the situation where the multi-hop MP-EBGP peers exchange IPv4 routes or the routes corresponding to the two addresses used for BGP session establishment. If such routes are exchanged over the multi-hop MP-EBGP session, one PE has two routes destined for the PE in the other AS. One route is advertised by the ASBR in the local AS, and the other is advertised over the multi-hop MP-EBGP session. Because EBGP routes take precedence over IBGP routes, BGP preferentially selects the route advertised over the multi-hop MP-EBGP session, causing constant flapping of the route destined for the peer PE. As a result, VPN routes become unreachable.

Verification

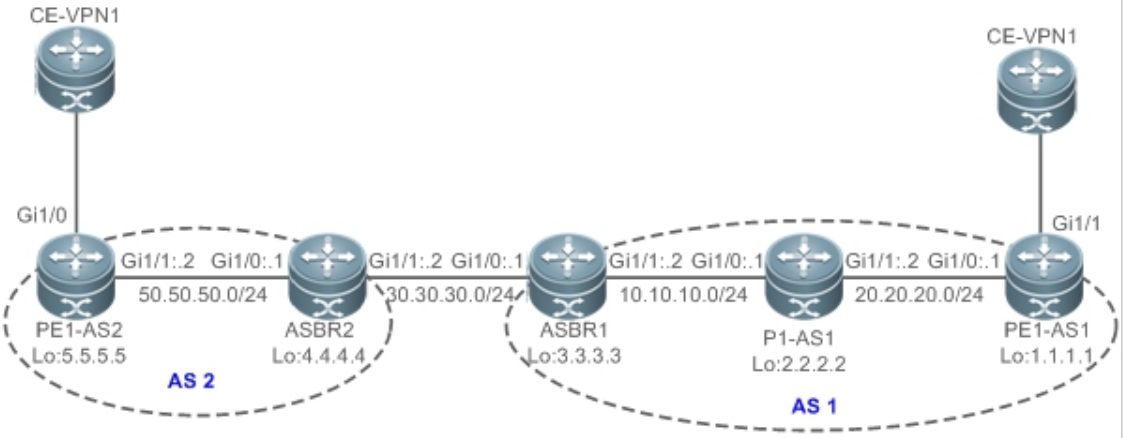
- Check that sites of the same VPN can communicate with each other but sites of different VPNs cannot communicate with each other.

Configuration Example

Inter-AS VPN Option C: Enabling IPv4 Label Switching Between EBGP Neighbors

Scenario

Figure 2- 31



Requirements: A VPN has sites in two ASs. Enable inter-AS communication between the sites.

Configuration Steps

- On PE1-AS1, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure a multi-hop MP-EBGP session. Disable the capability of exchanging IPv4 routes over EBGP. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On PE1-AS2, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure a multi-hop MP-EBGP session. Disable the capability of exchanging IPv4 routes over EBGP. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On P1-AS1, configure a loopback interface, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol.
- On ASBR1, configure a loopback interface, and configure an ACL and a route map to distribute or set labels only for compliant routes. Establish an EBGP session with ASBR2. Configure a route map to set labels only for the compliant routes of PEs. (The route map is optional and used to specify which routes are distributed with a label over BGP.) Configure statically the routes of PEs in the local AS. Configure MPLS. Configure an ACL to distribute labels only to specified BGP routes. (The ACL is optional and used to reduce the number of undesired route entries.) Configure a backbone routing protocol, which redistributes only BGP routes compliant with the route map. (The route map is optional and used to reduce the number of undesired route entries.) Configure the IP address of the interface connected to ASBR2, and enable label forwarding on the interface.
- On ASBR2, configure a loopback interface, and configure an ACL and a route map to distribute or set labels only for compliant routes. Establish an EBGP session with ASBR1. Configure a route map to set labels only for the compliant routes of PEs. (The route map is optional and used to specify which routes are distributed with a label over BGP.) Configure statically the routes of PEs in the local AS. Configure MPLS. Configure an ACL to distribute labels only to specified BGP routes. (The ACL is optional and used to reduce the number of undesired route entries.) Configure a backbone routing protocol, which redistributes only BGP routes compliant with the route map. (The route map is optional and used to reduce the number of undesired route entries.) Configure the IP

	address of the interface connected to ASBR1, and enable label forwarding on the interface.
PE1-AS1	PE1-AS1# configure terminal PE1-AS1(config)# interface loopback 0 PE1-AS1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1-AS1# configure terminal PE1-AS1(config)# ip vrf VPN1 PE1-AS1(config-vrf)# rd 1:100 PE1-AS1(config-vrf)# route-target both 1:100 PE1-AS1(config-vrf)# end PE1-AS1# configure terminal PE1-AS1(config)# interface GigabitEthernet 1/1 PE1-AS1(config-if-GigabitEthernet 1/1)# ip vrf forwarding VPN1 PE1-AS1(config-if-GigabitEthernet 1/1)# ip address 192.168.16.2 255.255.255.0 PE1-AS1(config-if-GigabitEthernet 1/1)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config-router)# neighbor 5.5.5.5 remote-as 2 PE1-AS1(config-router)# neighbor 5.5.5.5 update-source loopback 0 PE1-AS1(config-router)# neighbor 5.5.5.5 ebgp-multihop PE1-AS1(config-router)# address-family ipv4 PE1-AS1(config-router-af)# no neighbor 5.5.5.5 activate PE1-AS1(config-router-af)# exit PE1-AS1(config-router)# address-family vpnv4 unicast PE1-AS1(config-router-af)# neighbor 5.5.5.5 activate PE1-AS1(config-router-af)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config)# address-family ipv4 vrf VPN1 PE1-AS1(config-router-af)# neighbor 192.168.16.1 remote-as 65002 PE1-AS1(config-router-af)# end PE1-AS1# configure terminal PE1-AS1(config)# mpls ip

	PE1-AS1(config)# mpls router ldp PE1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force PE1-AS1(config-mpls-router)# exit PE1-AS1(config)# interface GigabitEthernet 1/0 PE1-AS1(config-if-GigabitEthernet 1/0)# ip address 20.20.20.1 255.255.255.0 PE1-AS1(config-if-GigabitEthernet 1/0)# label-switching PE1-AS1(config-if-GigabitEthernet 1/0)# mpls ip PE1-AS1(config-if-GigabitEthernet 1/0)# end PE1-AS1# configure terminal PE1-AS1(config)# router ospf 10 PE1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 PE1-AS1(config-router)# network 1.1.1.1 0.0.0.0 area 0 PE1-AS1(config-router)# end
PE1-AS2	PE1-AS2# configure terminal PE1-AS2(config)# interface loopback 0 PE1-AS2(config-if-Loopback 0)# ip address 5.5.5.5 255.255.255.255 PE1-AS2# configure terminal PE1-AS2(config)# ip vrf VPN1 PE1-AS2(config-vrf)# rd 1:100 PE1-AS2(config-vrf)# route-target both 1:100 PE1-AS2(config-vrf)# end PE1-AS2# configure terminal PE1-AS2(config)# interface GigabitEthernet 1/0 PE1-AS2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPN1 PE1-AS2(config-if-GigabitEthernet 1/0)# ip address 192.168.15.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/0)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config-router)# neighbor 1.1.1.1 remote-as 1 PE1-AS2(config-router)# neighbor 1.1.1.1 update-source loopback 0 PE1-AS2(config-router)# neighbor 1.1.1.1 ebgp-multihop

	PE1-AS2(config-router)# address-family ipv4 PE1-AS2(config-router-af)# no neighbor 1.1.1.1 activate PE1-AS2(config-router-af)# exit PE1-AS2(config-router)# address-family vpnv4 unicast PE1-AS2(config-router-af)# neighbor 1.1.1.1 activate PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config)# address-family ipv4 vrf VPN1 PE1-AS2(config-router-af)# neighbor 192.168.15.1 remote-as 65002 PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# mpls ip PE1-AS2(config)# mpls router ldp PE1-AS2(config-mpls-router)# ldp router-id interface loopback 0 force PE1-AS2(config-mpls-router)# exit PE1-AS2(config)# interface GigabitEthernet 1/1 PE1-AS2(config-if-GigabitEthernet 1/1)# ip address 50.50.50.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/1)# label-switching PE1-AS2(config-if-GigabitEthernet 1/1)# mpls ip PE1-AS2(config-if-GigabitEthernet 1/1)# end PE1-AS2# configure terminal PE1-AS2(config)# router ospf 10 PE1-AS2(config-router)# network 50.50.50.0 0.0.0.255 area 0 PE1-AS2(config-router)# network 5.5.5.5 0.0.0.0 area 0 PE1-AS2(config-router)# end
P1-AS1	P1-AS1# configure terminal P1-AS1(config)# interface loopback 0 P1-AS1(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 P1-AS1# configure terminal P1-AS1(config)# mpls ip

	P1-AS1(config)# mpls router ldp P1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force P1-AS1(config-mpls-router)# exit P1-AS1(config)# interface GigabitEthernet 1/0 P1-AS1(config-if-GigabitEthernet 1/0)# ip address 10.10.10.1 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/0)# label-switching P1-AS1(config-if-GigabitEthernet 1/0)# mpls ip P1-AS1(config-if-GigabitEthernet 1/0)# exit P1-AS1(config)# interface GigabitEthernet 1/1 P1-AS1(config-if-GigabitEthernet 1/1)# ip address 20.20.20.2 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/1)# label-switching P1-AS1(config-if-GigabitEthernet 1/1)# mpls ip P1-AS1# configure terminal P1-AS1(config)# router ospf 10 P1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 P1-AS1(config-router)# network 10.10.10.0 0.0.0.255 area 0 P1-AS1(config-router)# network 2.2.2.2 0.0.0.0 area 0 P1-AS1(config-router)# end
ASBR1	ASBR1# configure terminal ASBR1(config)# interface loopback 0 ASBR1(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 ASBR1# configure terminal ASBR1(config)# ip access-list extended 101 ASBR1(config-ext-nacl)# permit ip host 1.1.1.1 any ASBR1(config-ext-nacl)# exit ASBR1(config)# ip access-list extended 102 ASBR1(config-ext-nacl)# permit ip host 5.5.5.5 any ASBR1(config-ext-nacl)# exit ASBR1(config)# route-map set-mpls ASBR1(config-route-map)# match ip address 101 ASBR1(config-route-map)# set mpls-label

```
ASBR1(config-route-map)# exit
ASBR1(config)# route-map external-pe-route
ASBR1(config-route-map)# match ip address 102
ASBR1(config-route-map)# end
ASBR1# configure terminal
ASBR1(config)# router bgp 1
ASBR1(config-router)# neighbor 30.30.30.2 remote-as 2
ASBR1(config-router)# address-family ipv4
ASBR1(config-router-af)# neighbor 30.30.30.2 send-label
ASBR1(config-router-af)# neighbor 30.30.30.2 route-map set-mpls out
ASBR1(config-router-af)# network 1.1.1.1 mask 255.255.255.255
ASBR1(config-router-af)# end
ASBR1# configure terminal
ASBR1(config)# mpls ip
ASBR1(config)# mpls router ldp
ASBR1(config-mpls-router)# ldp router-id interface loopback 0 force
ASBR1(config-mpls-router)# advertise-labels for bgp-routes acl 102
ASBR1(config-mpls-router)# exit
ASBR1(config)# interface GigabitEthernet 1/1
ASBR1(config-if-GigabitEthernet 1/1)# ip address 10.10.10.2 255.255.255.0
ASBR1(config-if-GigabitEthernet 1/1)# label-switching
ASBR1(config-if-GigabitEthernet 1/1)# mpls ip
ASBR1(config-if-GigabitEthernet 1/1)# end
ASBR1# configure terminal
ASBR1(config)# router ospf 10
ASBR1(config-router)# network 10.10.10.0 0.0.0.255 area 0
ASBR1(config-router)# network 3.3.3.3 0.0.0.0 area 0
ASBR1(config-router)# redisteibute bgp subnets route-map external-pe-route
ASBR1(config-router)# end
ASBR1(config)# interface GigabitEthernet 1/0
ASBR1(config-if-GigabitEthernet 1/0)# ip address 30.30.30.1 255.255.255.0
ASBR1(config-if-GigabitEthernet 1/0)# label-switching
```

ASBR2	ASBR2# configure terminal ASBR2(config)# interface loopback 0 ASBR2(config-if-Loopback 0)# ip address 4.4.4.4 255.255.255.255 ASBR2# configure terminal ASBR2(config)# ip access-list extended 101 ASBR2(config-ext-nacl)# permit ip host 5.5.5.5 any ASBR2(config-ext-nacl)# exit ASBR2(config)# ip access-list extended 102 ASBR2(config-ext-nacl)# permit ip host 1.1.1.1 any ASBR2(config-ext-nacl)# exit ASBR2(config)# route-map set-mpls ASBR2(config-route-map)# match ip address 101 ASBR2(config-route-map)# set mpls-label ASBR2(config-route-map)# exit ASBR2(config)# route-map external-pe-route ASBR2(config-route-map)# match ip address 102 ASBR2(config-route-map)# end ASBR2# configure terminal ASBR2(config)# router bgp 1 ASBR2(config-router)# neighbor 30.30.30.1 remote-as 1 ASBR2(config-router)# address-family ipv4 ASBR2(config-router-af)# neighbor 30.30.30.1 send-label ASBR2(config-router-af)# neighbor 30.30.30.1 route-map set-mpls out ASBR2(config-router-af)# network 5.5.5.5 mask 255.255.255.255 ASBR2(config-router-af)# end ASBR2# configure terminal ASBR2(config)# mpls ip ASBR2(config)# mpls router ldp ASBR2(config-mpls-router)# ldp router-id interface loopback 0 force ASBR2(config-mpls-router)# advertise-labels for bgp-routes acl 102 ASBR2(config-mpls-router)# exit

	ASBR2(config)# interface GigabitEthernet 1/0 ASBR2(config-if-GigabitEthernet 1/0)# ip address 50.50.50.1 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/0)# label-switching ASBR2(config-if-GigabitEthernet 1/0)# mpls ip ASBR2(config-if-GigabitEthernet 1/0)# end ASBR2# configure terminal ASBR2(config)# router ospf 10 ASBR2(config-router)# network 50.50.50.0.0 0.0.0.255 area 0 ASBR2(config-router)# network 5.5.5.5 0.0.0.0 area 0 ASBR2(config-router)# redisteibute bgp subnets route-map external-pe-route ASBR2(config-router)# end ASBR2(config)# interface GigabitEthernet 1/0 ASBR2(config-if-GigabitEthernet 1/0)# ip address 30.30.30.2 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/0)# label-switching
Verification	● Run the **show mpls forwarding-table** command to display MPLS VPN routes. ● Check that different sites of the same VPN can communicate with each other.
PE1-AS1	Assume that the loopback interface on the peer CE is 20.20.20.20. 20.20.20.20 is pingable from the local CE. 5.5.5.5 is unpingable. The destination address 5.5.5.5 is pingable from the source address 1.1.1.1. 4.4.4.4 is unpingable. The destination address 4.4.4.4 is unpingable from the source address 1.1.1.1. 3.3.3.3 is pingable. 2.2.2.2 is pingable.
PE1-AS2	Assume that the loopback interface on the peer CE is 10.10.10.10. 10.10.10.10 is pingable from the local CE. 1.1.1.1 is unpingable. The destination address 1.1.1.1 is pingable from the source address 5.5.5.5. 4.4.4.4 is pingable. 3.3.3.3 is unpingable. The destination address 3.3.3.3 is unpingable from the source address 5.5.5.5.

	The destination address 2.2.2.2 is unpingable from the source address 5.5.5.5. 1.1.1.1 is unpingable. The destination address 1.1.1.1 is pingable from the source address 5.5.5.5.
P1-AS1	1.1.1.1 is pingable. 3.3.3.3 is pingable. 4.4.4.4 is unpingable. 5.5.5.5 is unpingable.
ASBR1	Only the EBGP neighbor is pingable, That is: 30.30.30.2 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 4.4.4.4 is unpingable.
ASBR2	Only the EBGP neighbor is pingable, That is: 30.30.30.1 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 1.1.1.1 is unpingable.

Common Errors

The capability of exchanging IPv4 routes over multi-hop EBGP is not disabled when a multi-hop MP-EBGP session is configured. LDP is not configured on ASBRs for distributing labels to BGP routes, causing a communication failure between VPN sites.

2.4.4.2 Method 2: Enable IPv4 Label Switching Between EBGP Neighbors and Between IBGP Neighbors

As Method 1 enables IPv4 label switching only between EBGP neighbors, IGP and LDP in one AS need to maintain the routes destined for PEs in the other AS. That is, the inter-AS routes of PEs must be propagated to all devices between ASs. In actual application, for AS security consideration, the routes of PEs in one AS should not be propagated to each device in the other AS. Rather, the routes must be maintained only by BGP and transparent to IGP and LDP in the same AS. To meet this requirement, you can enable IPv4 label switching between EBGP neighbors and between IBGP neighbors.

Different from Method 1, Method 2 does not require IGP on ASBRs to redistribute BGP routes and does not require LDP to distribute labels to BGP routes. LDP is only responsible for LSP establishment in the local AS. However, IPv4 label switching must be enabled between IBGP neighbors and between EBGP neighbors for inter-AS LSP establishment, and PEs must support Stacking of three consecutive labels.

Configuration Steps

- (Mandatory) Configure PEs and CEs in each AS to exchange routes.
- (Mandatory) Configure IGP and MPLS signaling for ASs.
- (Mandatory) Configure PEs and ASBRs to establish an IBGP session in order to distribute labels to IPv4 routes.
- (Mandatory) Configure ASBRs to establish an EBGP session in order to distribute labels to IPv4 routes.
- (Mandatory) Configure a multi-hop MP-EBGP session.

Configuring PEs and CEs to Exchange Routes

For details, see "Configuring PEs and CEs to Exchange Routes" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring IGP and MPLS Signaling for ASs

For details, see "Configuring an MPLS Network" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring PEs and ASBRs to Establish an IBGP Session in Order to Distribute Labels to IPv4 Routes

In Method 1, the routes of PEs learned over EBGP are redistributed to IGP between ASs, whereas in Method 2, the inter-AS routes of PEs are transmitted over the IBGP session between ASBR and PE, and labels are distributed to those routes over BGP.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-number*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *peer-address* **remote-as** *asn-number*	Establishes an IBGP session with the ASBR (PE).
FS(config-router)# **neighbor** *peer-address* **update-source** *interface-name*	Specifies the address of a loopback interface as the source address for the BGP session established with the peer ASBR (PE)
FS(config-router)# **address-family ipv4**	Enters IPv4 address family configuration mode.
FS(config-router-af)# **neighbor** *peer-address* **send-label**	Configures the capability of exchanging labeled IPv4 routes with the peer ASBR (PE).
FS(config-router-af)# **show running-config**	Displays existing configurations.

Before you enable IPv4 label switching over the IBGP session between IBGP peers, run the **neighbor update-source** command to specify the address of a loopback interface as the source address for the IBGP session; otherwise, an inter-AS LSP cannot be established.

Configuring ASBRs to Establish an EBGP Session in Order to Distribute Labels to IPv4 Routes

For details, see "Configuring ASBRs to Establish an EBGP Session in Order to Distribute Labels to IPv4 Routes" in section 2.4.4.1 "Method 1: Enable IPv4 Label Switching Only Between EBGP Neighbors."

Configuring a Multi-hop MP-EBGP Session

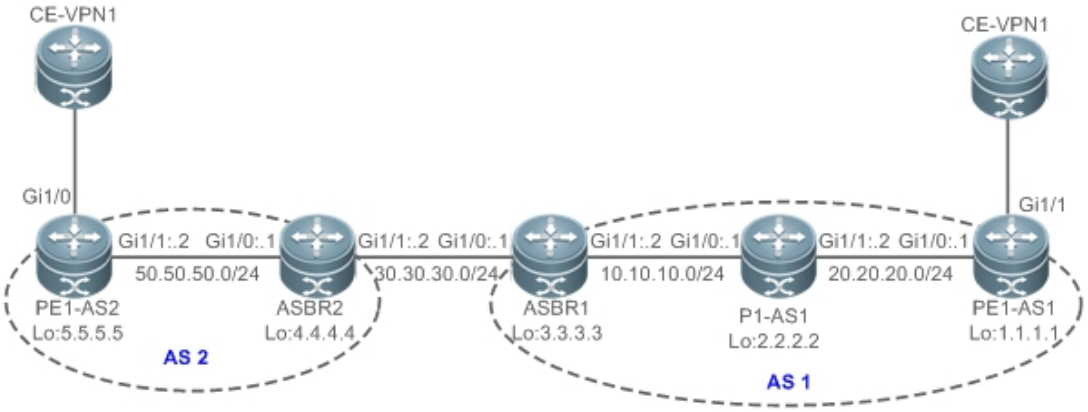
For details, see "Configuring a Multi-hop MP-EBGP Session" in section 2.4.4.1 "Method 1: Enable IPv4 Label Switching Only Between EBGP Neighbors."

Verification

- Check that sites of the same VPN can communicate with each other but sites of different VPNs cannot communicate with each other.

Configuration Example

Inter-AS VPN Option C: Enabling IPv4 Label Switching Between EBGP Neighbors and Between IBGP Neighbors

Scenario **Figure 2-32**	 Requirements: A VPN has sites in two ASs. Enable inter-AS communication between the sites.
Configuration Steps	- On PE1-AS1, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure a multi-hop MP-EBGP session. Disable the capability of exchanging IPv4 routes over EBGP. Establish an IBGP session with ASBR1. Enable IPv4 label switching. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. - On PE1-AS2, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Configure a multi-hop MP-EBGP session. Disable the capability of exchanging IPv4 routes over EBGP. Establish an IBGP session with ASBR2. Enable IPv4 label switching. Configure a CE neighbor relationship over EBGP. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. - On P1-AS1, configure a loopback interface, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol. - On ASBR1, configure a loopback interface, and configure an ACL and a route map to distribute or set labels only for compliant routes. Establish an EBGP session with ASBR2. Configure a route map to set labels only for the compliant routes of PEs. (The route map is optional and used to specify which routes are distributed with a label over BGP.) Configure statically the routes of PEs in the local AS. Configure MPLS signaling and enable MPLS on the interface connected to the public network. Configure OSPF for the backbone network to transmit routes. Configure the IP address of the interface connected to ASBR2, and enable label forwarding on the interface. - On ASBR2, configure a loopback interface, and configure an ACL and a route map to distribute or set labels only for compliant routes. Establish an EBGP session with ASBR1. Configure a route map to set labels only for the compliant routes of PEs. (The route map is optional and used to specify which routes are distributed with a label over BGP.) Configure statically the routes of PEs in the local AS. Configure MPLS signaling and enable MPLS on the

	interface connected to the public network. Configure OSPF for the backbone network to transmit routes. Configure the IP address of the interface connected to ASBR1, and enable label forwarding on the interface.
PE1-AS1	PE1-AS1# configure terminal PE1-AS1(config)# interface loopback 0 PE1-AS1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1-AS1# configure terminal PE1-AS1(config)# ip vrf VPNA PE1-AS1(config-vrf)# rd 1:100 PE1-AS1(config-vrf)# route-target both 1:100 PE1-AS1(config-vrf)# end PE1-AS1# configure terminal PE1-AS1(config)# interface GigabitEthernet 1/1 PE1-AS1(config-if-GigabitEthernet 1/1)# ip vrf forwarding VPNA PE1-AS1(config-if-GigabitEthernet 1/1)# ip address 192.168.10.2 255.255.255.0 PE1-AS1(config-if-GigabitEthernet 1/1)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config-router)# neighbor 5.5.5.5 remote-as 2 PE1-AS1(config-router)# neighbor 5.5.5.5 update-source loopback 0 PE1-AS1(config-router)# neighbor 5.5.5.5 ebgp-multihop PE1-AS1(config-router)# address-family ipv4 PE1-AS1(config-router-af)# no neighbor 5.5.5.5 activate PE1-AS1(config-router-af)# exit PE1-AS1(config-router)# address-family vpnv4 unicast PE1-AS1(config-router-af)# neighbor 5.5.5.5 activate PE1-AS1(config-router-af)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config-router)# neighbor 3.3.3.3 remote-as 1 PE1-AS1(config-router)# neighbor 3.3.3.3 update-source loopback 0 PE1-AS1(config-router)# address-family ipv4 PE1-AS1(config-router-af)# neighbor 3.3.3.3 activate

	PE1-AS1(config-router-af)# neighbor 3.3.3.3 send-label PE1-AS1(config-router-af)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config)# address-family ipv4 vrf VPNA PE1-AS1(config-router-af)# neighbor 192.168.10.1 remote-as 65002 PE1-AS1(config-router-af)# end PE1-AS1# configure terminal PE1-AS1(config)# mpls ip PE1-AS1(config)# mpls router ldp PE1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force PE1-AS1(config-mpls-router)# exit PE1-AS1(config)# interface GigabitEthernet 1/0 PE1-AS1(config-if-GigabitEthernet 1/0)# ip address 20.20.20.1 255.255.255.0 PE1-AS1(config-if-GigabitEthernet 1/0)# label-switching PE1-AS1(config-if-GigabitEthernet 1/0)# mpls ip PE1-AS1(config-if-GigabitEthernet 1/0)# end PE1-AS1# configure terminal PE1-AS1(config)# router ospf 10 PE1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 PE1-AS1(config-router)# network 1.1.1.1 0.0.0.0 area 0 PE1-AS1(config-router)# end
PE1-AS2	PE1-AS2# configure terminal PE1-AS2(config)# interface loopback 0 PE1-AS2(config-if-Loopback 0)# ip address 5.5.5.5 255.255.255.255 PE1-AS2# configure terminal PE1-AS2(config)# ip vrf VPNA PE1-AS2(config-vrf)# rd 1:100 PE1-AS2(config-vrf)# route-target both 1:100 PE1-AS2(config-vrf)# end PE1-AS2# configure terminal

```
PE1-AS2(config)# interface GigabitEthernet 1/0
PE1-AS2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPNA
PE1-AS2(config-if-GigabitEthernet 1/0)# ip address 192.168.11.2 255.255.255.0
PE1-AS2(config-if-GigabitEthernet 1/0)# end
PE1-AS2# configure terminal
PE1-AS2(config)# router bgp 1
PE1-AS2(config-router)# neighbor 1.1.1.1 remote-as 2
PE1-AS2(config-router)# neighbor 1.1.1.1 update-source loopback 0
PE1-AS2(config-router)# neighbor 1.1.1.1 ebgp-multihop
PE1-AS2(config-router)# address-family ipv4
PE1-AS2(config-router-af)# no neighbor 1.1.1.1 activate
PE1-AS2(config-router-af)# exit
PE1-AS2(config-router)# address-family vpnv4 unicast
PE1-AS2(config-router-af)# neighbor 1.1.1.1 activate
PE1-AS2(config-router-af)# end
PE1-AS2# configure terminal
PE1-AS2(config)# router bgp 1
PE1-AS2(config-router)# neighbor 4.4.4.4 remote-as 1
PE1-AS2(config-router)# neighbor 4.4.4.4 update-source loopback 0
PE1-AS2(config-router)# address-family ipv4
PE1-AS2(config-router-af)# neighbor 4.4.4.4 activate
PE1-AS2(config-router-af)# neighbor 4.4.4.4 send-label
PE1-AS2(config-router-af)# end
PE1-AS2# configure terminal
PE1-AS2(config)# router bgp 1
PE1-AS2(config)# address-family ipv4 vrf VPNA
PE1-AS2(config-router-af)# neighbor 192.168.11.1 remote-as 65002
PE1-AS2(config-router-af)# end
PE1-AS2# configure terminal
PE1-AS2(config)# mpls ip
PE1-AS2(config)# mpls router ldp
PE1-AS2(config-mpls-router)# ldp router-id interface loopback 0 force
```

	PE1-AS2(config-mpls-router)# exit PE1-AS2(config)# interface GigabitEthernet 1/1 PE1-AS2(config-if-GigabitEthernet 1/1)# ip address 50.50.50.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/1)# label-switching PE1-AS2(config-if-GigabitEthernet 1/1)# mpls ip PE1-AS2(config-if-GigabitEthernet 1/1)# end PE1-AS2# configure terminal PE1-AS2(config)# router ospf 10 PE1-AS2(config-router)# network 50.50.50.0 0.0.0.255 area 0 PE1-AS2(config-router)# network 5.5.5.5 0.0.0.0 area 0 PE1-AS2(config-router)# end
P1-AS1	P1-AS1# configure terminal P1-AS1(config)# interface loopback 0 P1-AS1(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 P1-AS1# configure terminal P1-AS1(config)# mpls ip P1-AS1(config)# mpls router ldp P1-AS1(config-mpls-router)# ldp router-id interface loopback 0 force P1-AS1(config-mpls-router)# exit P1-AS1(config)# interface GigabitEthernet 1/0 P1-AS1(config-if-GigabitEthernet 1/0)# ip address 10.10.10.1 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/0)# label-switching P1-AS1(config-if-GigabitEthernet 1/0)# mpls ip P1-AS1(config-if-GigabitEthernet 1/0)# exit P1-AS1(config)# interface GigabitEthernet 1/1 P1-AS1(config-if-GigabitEthernet 1/1)# ip address 20.20.20.2 255.255.255.0 P1-AS1(config-if-GigabitEthernet 1/1)# label-switching P1-AS1(config-if-GigabitEthernet 1/1)# mpls ip P1-AS1# configure terminal P1-AS1(config)# router ospf 10 P1-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0

	P1-AS1(config-router)# network 10.10.10.0 0.0.0.255 area 0 P1-AS1(config-router)# network 2.2.2.2 0.0.0.0 area 0 P1-AS1(config-router)# end
ASBR1	ASBR1# configure terminal ASBR1(config)# interface loopback 0 ASBR1(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 ASBR1# configure terminal ASBR1(config)# ip access-list extended 101 ASBR1(config)# permit ip host 1.1.1.1 any ASBR1(config)# exit ASBR1(config)# ip access-list extended 102 ASBR1(config)# permit ip host 5.5.5.5 any ASBR1(config)# route-map internal-mpls-route permit 10 ASBR1(config-route-map)# match ip address 101 ASBR1(config-route-map)# set mpls-label ASBR1(config-route-map)# exit ASBR1(config)# route-map external-mpls-route permit 10 ASBR1(config-route-map)# match ip address 102 ASBR1(config-route-map)# set mpls-label ASBR1(config-route-map)# end ASBR1# configure terminal ASBR1(config)# router bgp 1 ASBR1(config-router)# neighbor 30.30.30.2 remote-as 2 ASBR1(config-router)# neighbor 1.1.1.1 remote-as 1 ASBR1(config-router)# neighbor 1.1.1.1 update-source loopback 0 ASBR1(config-router)# address-family ipv4 ASBR1(config-router-af)# neighbor 30.30.30.2 send-label ASBR1(config-router-af)# neighbor 30.30.30.2 route-map internal-mpls-route out ASBR1(config-router-af)# neighbor 1.1.1.1 send-label ASBR1(config-router-af)# neighbor 1.1.1.1 route-map external-mpls-route out ASBR1(config-router-af)# network 1.1.1.1 mask 255.255.255.255

	ASBR1(config-router-af)# end ASBR1# configure terminal ASBR1(config)# mpls ip ASBR1(config)# mpls router ldp ASBR1(config-mpls-router)# ldp router-id interface loopback 0 force ASBR1(config-mpls-router)# exit ASBR1(config)# interface GigabitEthernet 1/1 ASBR1(config-if-GigabitEthernet 1/1)# ip address 10.10.10.2 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/1)# label-switching ASBR1(config-if-GigabitEthernet 1/1)# mpls ip ASBR1(config-if-GigabitEthernet 1/1)# end ASBR1# configure terminal ASBR1(config)# router ospf 10 ASBR1(config-router)# network 10.10.10.0 255.255.255.0 area 0 ASBR1(config-router)# network 3.3.3.3 0.0.0.0 area 0 ASBR1(config-router)# end ASBR1(config)# interface GigabitEthernet 1/0 ASBR1(config-if-GigabitEthernet 1/0)# ip address 30.30.30.1 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/0)# label-switching
ASBR2	ASBR2# configure terminal ASBR2(config)# interface loopback 0 ASBR2(config-if-Loopback 0)# ip address 4.4.4.4 255.255.255.255 ASBR2# configure terminal ASBR2(config)# ip access-list extended 101 ASBR2(config)# permit ip host 5.5.5.5 any ASBR2(config)# exit ASBR2(config)# ip access-list extended 102 ASBR2(config)# permit ip host 1.1.1.1 any ASBR2(config)# route-map internal-mpls-route permit 10 ASBR2(config-route-map)# match ip address 101 ASBR2(config-route-map)# set mpls-label

```
ASBR2(config-route-map)# exit
ASBR2(config)# route-map external-mpls-route permit 10
ASBR2(config-route-map)# match ip address 102
ASBR2(config-route-map)# set mpls-label
ASBR2(config-route-map)# end
ASBR2# configure terminal
ASBR2(config)# router bgp 1
ASBR2(config-router)# neighbor 30.30.30.1 remote-as 2
ASBR2(config-router)# neighbor 5.5.5.5 remote-as 1
ASBR2(config-router)# neighbor 5.5.5.5 update-source loopback 0
ASBR2(config-router)# address-family ipv4
ASBR2(config-router-af)# neighbor 30.30.30.1 send-label
ASBR2(config-router-af)# neighbor 30.30.30.1 route-map internal-mpls-route out
ASBR2(config-router-af)# neighbor 5.5.5.5 send-label
ASBR2(config-router-af)# neighbor 5.5.5.5 route-map external-mpls-route out
ASBR2(config-router-af)# network 5.5.5.5 mask 255.255.255.255
ASBR2(config-router-af)# end
ASBR2# configure terminal
ASBR2(config)# mpls ip
ASBR2(config)# mpls router ldp
ASBR2(config-mpls-router)# ldp router-id interface loopback 0 force
ASBR2(config-mpls-router)# exit
ASBR2(config)# interface GigabitEthernet 1/0
ASBR2(config-if-GigabitEthernet 1/0)# ip address 50.50.50.1 255.255.255.0
ASBR2(config-if-GigabitEthernet 1/0)# label-switching
ASBR2(config-if-GigabitEthernet 1/0)# mpls ip
ASBR2(config-if-GigabitEthernet 1/0)# end
ASBR2# configure terminal
ASBR2(config)# router ospf 10
ASBR2(config-router)# network 50.50.50.0 255.255.255.0 area 0
ASBR2(config-router)# network 4.4.4.4 0.0.0.0 area 0
ASBR2(config-router)# end
```

	ASBR2(config)# interface GigabitEthernet 1/1 ASBR2(config-if-GigabitEthernet 1/1)# ip address 30.30.30.2 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/1)# label-switching
Verification	● Run the **show mpls forwarding-table** command to display MPLS VPN routes. ● Check that different sites of the same VPN can communicate with each other.
PE1-AS1	Assume that the loopback interface on the peer CE is 20.20.20.20. 20.20.20.20 is pingable from the local CE. 5.5.5.5 is unpingable. The source address 1.1.1.1 of 5.5.5.5 is pingable. 4.4.4.4 is unpingable. The source address 1.1.1.1 of 4.4.4.4 is unpingable. 3.3.3.3 is pingable. 2.2.2.2 is pingable.
PE1-AS2	Assume that the loopback interface on the peer CE is 10.10.10.10. 10.10.10.10 is pingable from the local CE. 1.1.1.1 is unpingable. The source address 5.5.5.5 of 1.1.1.1 is pingable. 4.4.4.4 is pingable. 3.3.3.3 is unpingable. The source address 5.5.5.5 of 3.3.3.3 is unpingable. 2.2.2.2 is unpingable. The source address 5.5.5.5 of 2.2.2.2 is unpingable. 1.1.1.1 is unpingable. The source address 5.5.5.5 of 1.1.1.1 is pingable.
P1-AS1	1.1.1.1 is pingable. 3.3.3.3 is pingable. 4.4.4.4 is unpingable. 5.5.5.5 is unpingable.
ASBR1	Only the EBGP neighbor is pingable. That is: 30.30.30.2 is pingable.

	10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 4.4.4.4 is unpingable.
ASBR2	Only the EBGP neighbor is pingable. That is: 30.30.30.1 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 1.1.1.1 is unpingable.

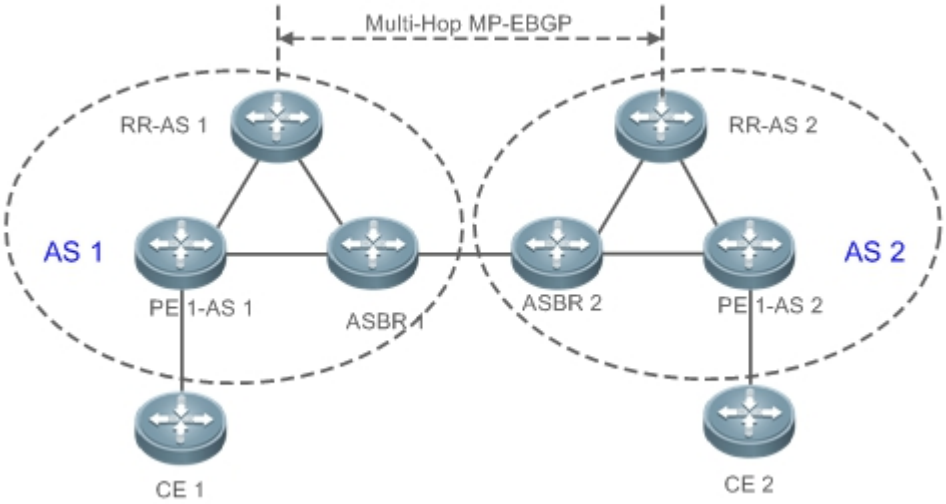
Common Errors

The capability of exchanging IPv4 routes over multi-hop EBGP is not disabled when a multi-hop MP-EBGP session is configured. IPv4 label switching is not configured between IBGP neighbors, causing a communication failure between VPN sites.

2.4.4.3Method 3: Establish Multi-hop MP-EBGP Between RRs

The Option C solution requires inter-AS VPN sites to be fully meshed. A new VPN site must set up a multi-hop MP-EBGP connection with PEs in the other AS. This hampers VPN site scalability. To address this problem, you can deploy an RR in each AS and enable the RRs in different ASs to set up a multi-hop MP-EBGP connection for exchanging VPN routes.

Figure 2- 33 Establishment of Inter-RR Multi-hop MP-EBGP for Option C



Configuration Steps

- (Mandatory) Configure PEs and CEs in each AS to exchange routes.
- (Mandatory) Configure IGP and MPLS signaling for ASs.

- (Mandatory) Configure RRs to establish an MP-IBGP session with PEs, and configure IPv4 label switching.
- (Mandatory) Configure RRs to establish an IBGP session with ASBRs, and configure IPv4 label switching.
- (Mandatory) Configure ASBRs to establish an EBGP session in order to distribute labels to IPv4 routes.
- (Mandatory) Configure a multi-hop MP-EBGP session.

Configuring PEs and CEs to Exchange Routes

For details, see "Configuring PEs and CEs to Exchange Routes" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring IGP and MPLS Signaling for ASs

For details, see "Configuring an MPLS Network" in section 2.4.1 "Configuring BGP MPLS IP VPN Basic Functions."

Configuring RRs to Establish an MP-IBGP Session with PEs, and Configuring IPv4 Label Switching

Configure RRs to establish an MP-IBGP session with PEs for exchanging VPN routes, and enable IPv4 label switching over the MP-IBGP session.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-number*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *peer-address* **remote-as** *asn-number*	Establishes an IBGP session.
FS(config-router)# **neighbor** *peer-adddress* **update-source** *interface-name*	Specifies the address of a loopback interface as the source address for the IBGP session.
FS(config-router)# **address-family ipv4**	Enters IPv4 address family configuration mode.
FS(config-router-af)# **neighbor** *peer-address* **activate**	Enables the capability of exchanging IPv4 routes.
FS(config-router-af)# **neighbor** *peer-address* **send-label**	Enables IPv4 label switching.
FS(config-router-af)# **neighbor** *peer-address* **route-reflector-client**	Configures peer PEs as the clients for reflecting IPv4 routes on the RRs.
FS(config-router-af)# **exit**	Exits IPv4 address family configuration mode.
FS(config-router)# **address-family vpnv4**	Enters VPN address family configuration mode.
FS(config-router-af)# **neighbor** *peer-address* **activate**	Enables the capability of exchanging VPN routes between peers.
FS(config-router-af)# **neighbor** *peer-adderss* **route-reflector-client**	Configures peer PEs as the clients for reflecting VPN routes on the RRs.
FS(config-router-af)# **show running-config**	Displays existing configurations.

Configuring RRs to Establish an IBGP Session with ASBRs, and Configuring IPv4 Label Switching

Configure ASBRs to establish an MP-IBGP session with RRs, so that ASBRs receive from RRs the routes destined for PEs in the local AS, and send RRs the routes destined for PEs in the other AS. You also need to enable IPv4 label switching over the MP-IBGP session.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-number*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *peer-address* **remote-as** *asn-number*	Establishes an IBGP session.
FS(config-router)# **neighbor** *peer-adddress* **update-source** *interface-name*	Specifies the address of a loopback interface as the source address for the IBGP session.
FS(config-router)**# address-family ipv4**	Enters IPv4 address family configuration mode.
FS(config-router-af)# **neighbor** *peer-address* **activate**	Enables the capability of exchanging IPv4 routes.
FS(config-router-af)# **neighbor** *peer-address* **send-label**	Enables IPv4 label switching.
FS(config-router)# **show running-config**	Displays existing configurations.

When you configure RRs to establish an IBGP session with ASBRs, you do not need to configure the ASBRs as the clients of the RRs, unless the ASBRs also function as PEs.

Configuring ASBRs to Establish an EBGP Session in Order to Distribute Labels to IPv4 Routes

For details, see "Configuring ASBRs to Establish an EBGP Session in Order to Distribute Labels to IPv4 Routes" in section 2.4.4.1 "Method 1: Enable IPv4 Label Switching Only Between EBGP Neighbors."

Configuring a Multi-hop MP-EBGP Session

Establish a multi-hop MP-EBGP session between the RRs in two different ASs for exchanging inter-AS VPN routes. You also need to disable the capability of exchanging IPv4 routes over the MP-EBGP session. Routes arriving at one PE are transmitted by the corresponding ASBR to the other AS.

Perform configuration in privileged mode according to the following table.

Command	Description
FS# **config terminal**	Enters global configuration mode.
FS(config)# **router bgp** *asn-number*	Configures BGP and enters BGP configuration mode.
FS(config-router)# **neighbor** *rr-address* **remote-as** *ebgp-asn-numbe*	Establishes an EBGP session.
FS(config-router)**# neighbor** *rr-address* **update-source** *interface-name*	Specifies the address of a loopback interface as the source address for the EBGP session.
FS(config-router)# **neighbor** *rr-address* **ebgp-multi-hop**	Configures multi-hop EBGP attributes.
FS(config-router)# **address-family ipv4**	Enters IPv4 address family configuration mode.
FS(config-router-af)# **no neighbor** *rr-address* **activate**	Disables the capability of exchanging IPv4 routes over the MP-EBGP session.
FS(config-router-af)# **exit**	Exits IPv4 address family configuration mode.

FS(config-router)# **address-family vpnv4**	Enters VPN address family configuration mode.
FS(config-router-af)# **neighbor** *rr-address* **activate**	Enables the capability of exchanging VPN routes with the RR in the other AS.
FS(config-router-af)# **neighbor** *rr-address* **next-hop-unchanged**	(Optional) Configures next hop retention for VPN routes advertised to the peer.
FS(config-router)# **show running-config**	Displays existing configurations.

By default, the next hop in VPN routes advertised to the EBGP peer is changed to the sender. Then the PEs in the other AS determine that the next hop in incoming VPN routes is an RR, causing all inter-AS VPN traffic to pass through the RR. Such a forwarding path is not the best path, and RRs require higher performance to forward traffic. To avoid the preceding situation, run the **neighbor next-hop-unchanged** command in VPNv4 address family configuration mode to configure next hop retention for VPNv4 routes sent to the BGP peer when you configure a multi-hop MP-EBGP connection on RRs.

Avoid the situation where the multi-hop MP-EBGP peers exchange IPv4 routes or the routes corresponding to the two addresses used for BGP session establishment. If such routes are exchanged over the multi-hop MP-EBGP session, one PE has two routes destined for the PE in the other AS. One route is advertised by the ASBR in the local AS, and the other is advertised over the multi-hop MP-EBGP session. Because EBGP routes take precedence over IBGP routes, BGP preferentially selects the route advertised over the multi-hop MP-EBGP session, causing constant flapping of the route destined for the peer PE. As a result, VPN routes become unreachable.

Verification

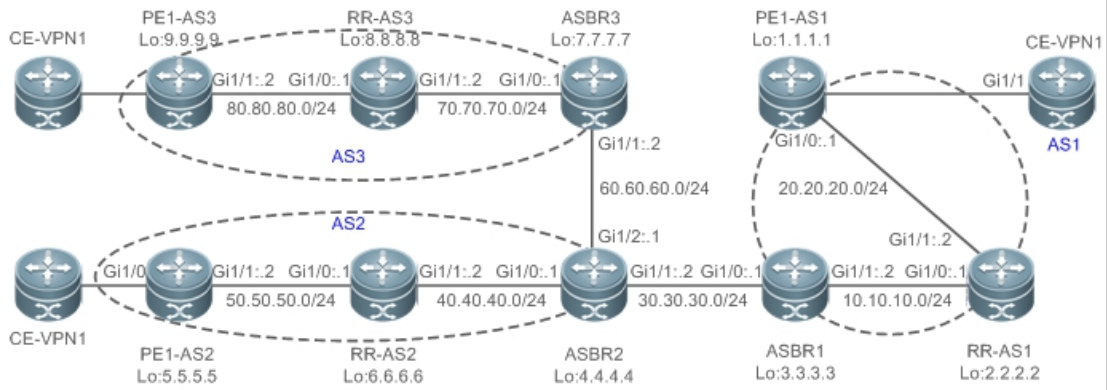
- Check that sites of the same VPN can communicate with each other but sites of different VPNs cannot communicate with each other.

Configuration Example

Inter-AS VPN Option C: RR Networking

Scenario

Figure 2- 34



When you apply the Option C solution, to ensure reachability between sites of the same VPN located in different ASs (in Figure 2- , sites of VPN1 are located in three different ASs), you need to configure PEs to be fully meshed across the ASs over BGP. A new VPN site needs to set up a BGP connection with existing VPN sites. This limits the Option C application. To ensure scalability, you can add an RR to each AS and set up a multi-hop MP-EBGP connection between RRs in different ASs to exchange inter-AS VPN routes. Then establish an MP-IBGP session between the PE and the RR in each AS.

Configuration Steps	● On PE1-AS1, PE1-AS2, and PE1-AS3, configure a loopback interface, create a VRF table named VPN1, define RD and RT values, and associate the VRF table with the corresponding interface. Establish an MP-IBGP session with the RR in the same AS, and enable IPv4 label switching. Configure a CE neighbor relationship over EBGP. ● On RR-AS1, RR-AS2, and RR-AS3, configure a loopback interface, establish an MP-IBGP session with the PE in the same AS, specify the PE as the RR client, and enable IPv4 label switching. Establish a multi-hop MP-EBGP session with the peer RR, and configure next hop retention for VPN routes exchanged with the peer RR. Disable the capability of exchanging IPv4 routes with the peer RR. Establish an IBGP session with the ASBR in the same AS. Enable IPv4 label switching. Configure MPLS, and configure OSPF for the backbone network to transmit routes. ● On ASBR1, ASBR2, and ASBR3, configure a loopback interface, and configure an ACL and a route map. Establish an EBGP session with the peer ASBR. Enable IPv4 label switching. Configure a route map to set labels only for the compliant routes of PEs. (The route map is optional and used to specify which routes are distributed with a label over BGP.) Establish an IBGP session with the RR in the same AS. Enable IPv4 label switching. Configure a route map to set labels only for the compliant inter-AS routes of PEs. Configure statically the route of the PE in the local AS. Configure MPLS signaling and enable MPLS on the interface connected to the public network. Configure OSPF for the backbone network to transmit routes. Configure the IP address of the interface connected to the peer ASBR, and enable label forwarding on the interface.
PE1-AS1	PE1-AS1# configure terminal PE1-AS1(config)# interface loopback 0 PE1-AS1(config-if-Loopback 0)# ip address 1.1.1.1 255.255.255.255 PE1-AS1# configure terminal PE1-AS1(config)# ip vrf VPN1 PE1-AS1(config-vrf)# rd 1:100 PE1-AS1(config-vrf)# route-target both 1:100 PE1-AS1(config-vrf)# end PE1-AS1# configure terminal PE1-AS1(config)# interface GigabitEthernet 1/1 PE1-AS1(config-if-GigabitEthernet 1/1)# ip vrf forwarding VPN1 PE1-AS1(config-if-GigabitEthernet 1/1)# ip address 192.168.16.2 255.255.255.0 PE1-AS1(config-if-GigabitEthernet 1/1)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config-router)# neighbor 2.2.2.2 remote-as 1 PE1-AS1(config-router)# neighbor 2.2.2.2 update-source loopback 0 PE1-AS1(config-router)# address-family vpnv4 unicast PE1-AS1(config-router-af)# neighbor 2.2.2.2 activate

	PE1-AS1(config-router-af)# exit PE1-AS1(config-router)# address-family ipv4 PE1-AS1(config-router-af)# neighbor 2.2.2.2 activate PE1-AS1(config-router-af)# neighbor 2.2.2.2 send-label PE1-AS1(config-router-af)# end PE1-AS1# configure terminal PE1-AS1(config)# router bgp 1 PE1-AS1(config)# address-family ipv4 vrf VPN1 PE1-AS1(config-router-af)# neighbor 192.168.16.1 remote-as 65002 PE1-AS1(config-router-af)# end
PE1-AS2	PE1-AS2# configure terminal PE1-AS2(config)# interface loopback 0 PE1-AS2(config-if-Loopback 0)# ip address 5.5.5.5 255.255.255.255 PE1-AS2# configure terminal PE1-AS2(config)# ip vrf VPN1 PE1-AS2(config-vrf)# rd 1:100 PE1-AS2(config-vrf)# route-target both 1:100 PE1-AS2(config-vrf)# end PE1-AS2# configure terminal PE1-AS2(config)# interface GigabitEthernet 1/0 PE1-AS2(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPN1 PE1-AS2(config-if-GigabitEthernet 1/0)# ip address 192.168.17.2 255.255.255.0 PE1-AS2(config-if-GigabitEthernet 1/0)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config-router)# neighbor 6.6.6.6 remote-as 2 PE1-AS2(config-router)# neighbor 6.6.6.6 update-source loopback 0 PE1-AS2(config-router)# address-family vpnv4 unicast PE1-AS2(config-router-af)# neighbor 6.6.6.6 activate PE1-AS2(config-router-af)# exit PE1-AS2(config-router)# address-family ipv4

	PE1-AS2(config-router-af)# neighbor 6.6.6.6 activate PE1-AS2(config-router-af)# neighbor 6.6.6.6 send-label PE1-AS2(config-router-af)# end PE1-AS2# configure terminal PE1-AS2(config)# router bgp 1 PE1-AS2(config)# address-family ipv4 vrf VPN1 PE1-AS2(config-router-af)# neighbor 192.168.17.1 remote-as 65002 PE1-AS2(config-router-af)# end
PE1-AS3	PE1-AS3# configure terminal PE1-AS3(config)# interface loopback 0 PE1-AS3(config-if-Loopback 0)# ip address 9.9.9.9 255.255.255.255 PE1-AS3# configure terminal PE1-AS3(config)# ip vrf VPN1 PE1-AS3(config-vrf)# rd 1:100 PE1-AS3(config-vrf)# route-target both 1:100 PE1-AS3(config-vrf)# end PE1-AS3# configure terminal PE1-AS3(config)# interface GigabitEthernet 1/0 PE1-AS3(config-if-GigabitEthernet 1/0)# ip vrf forwarding VPN1 PE1-AS3(config-if-GigabitEthernet 1/0)# ip address 192.168.18.2 255.255.255.0 PE1-AS3(config-if-GigabitEthernet 1/0)# end PE1-AS3# configure terminal PE1-AS3(config)# router bgp 1 PE1-AS3(config-router)# neighbor 8.8.8.8 remote-as 3 PE1-AS3(config-router)# neighbor 8.8.8.8 update-source loopback 0 PE1-AS3(config-router)# address-family vpnv4 unicast PE1-AS3(config-router-af)# neighbor 8.8.8.8 activate PE1-AS3(config-router-af)# exit PE1-AS3(config-router)# address-family ipv4 PE1-AS3(config-router-af)# neighbor 8.8.8.8 activate PE1-AS3(config-router-af)# neighbor 8.8.8.8 send-label

	PE1-AS3(config-router-af)# end PE1-AS3# configure terminal PE1-AS3(config)# router bgp 1 PE1-AS3(config)# address-family ipv4 vrf VPN1 PE1-AS3(config-router-af)# neighbor 192.168.18.1 remote-as 65002 PE1-AS3(config-router-af)# end
RR-AS1	RR-AS1# configure terminal RR-AS1(config)# interface loopback 0 RR-AS1(config-if-Loopback 0)# ip address 2.2.2.2 255.255.255.255 RR-AS1# configure terminal RR-AS1(config)# router bgp 1 RR-AS1(config-router)# neighbor 1.1.1.1 remote-as 1 RR-AS1(config-router)# neighbor 1.1.1.1 update-source loopback 0 RR-AS1(config-router)# address-family vpnv4 unicast RR-AS1(config-router-af)# neighbor 1.1.1.1 activate RR-AS1(config-router-af)# neighbor 1.1.1.1 route-reflector-client RR-AS1(config-router-af)# exit RR-AS1(config-router)# address-family ipv4 RR-AS1(config-router-af)# neighbor 1.1.1.1 activate RR-AS1(config-router-af)# neighbor 1.1.1.1 send-label RR-AS1(config-router-af)# neighbor 1.1.1.1 route-reflector-client RR-AS1(config-router-af)# end RR-AS1# configure terminal RR-AS1(config)# router bgp 1 RR-AS1(config-router)# neighbor 6.6.6.6 remote-as 2 RR-AS1(config-router)# neighbor 6.6.6.6 update-source loopback 0 RR-AS1(config-router)# neighbor 6.6.6.6 ebgp-multihop RR-AS1(config-router)# neighbor 8.8.8.8 remote-as 3 RR-AS1(config-router)# neighbor 8.8.8.8 update-source loopback 0 RR-AS1(config-router)# neighbor 8.8.8.8 ebgp-multihop RR-AS1(config-router)# address-family ipv4

```
RR-AS1(config-router-af)# no neighbor 6.6.6.6 activate
RR-AS1(config-router-af)# no neighbor 8.8.8.8 activate
RR-AS1(config-router-af)# exit-address-family
RR-AS1(config-router)# address-family vpnv4 unicast
RR-AS1(config-router-af)# neighbor 6.6.6.6 activate
RR-AS1(config-router-af)# neighbor 6.6.6.6 next-hop-unchanged
RR-AS1(config-router-af)# neighbor 8.8.8.8 activate
RR-AS1(config-router-af)# neighbor 8.8.8.8 next-hop-unchanged
RR-AS1(config-router-af)# end
RR-AS1# configure terminal
RR-AS1(config)# router bgp 1
RR-AS1(config-router)# neighbor 3.3.3.3 remote-as 1
RR-AS1(config-router)# neighbor 3.3.3.3 update-source loopback 0
RR-AS1(config-router)# address-family ipv4
RR-AS1(config-router-af)# neighbor 3.3.3.3 activate
RR-AS1(config-router-af)# neighbor 3.3.3.3 send-label
RR-AS1(config-router-af)# end
RR-AS1# configure terminal
RR-AS1(config)# mpls ip
RR-AS1(config)# mpls router ldp
RR-AS1(config-mpls-router)# ldp router-id interface loopback 0 force
RR-AS1(config-mpls-router)# exit
RR-AS1(config)# interface GigabitEthernet 1/1
RR-AS1(config-if-GigabitEthernet 1/1)# ip address 20.20.20.2 255.255.255.0
RR-AS1(config-if-GigabitEthernet 1/1)# label-switching
RR-AS1(config-if-GigabitEthernet 1/1)# mpls ip
RR-AS1(config-if-GigabitEthernet 1/1)# exit
RR-AS1(config)# interface GigabitEthernet 1/0
RR-AS1(config-if-GigabitEthernet 1/0)# ip address 10.10.10.1 255.255.255.0
RR-AS1(config-if-GigabitEthernet 1/0)# label-switching
RR-AS1(config-if-GigabitEthernet 1/0)# mpls ip
RR-AS1(config-if-GigabitEthernet 1/0)# end
```

	RR-AS1# configure terminal RR-AS1(config)# router ospf 10 RR-AS1(config-router)# network 20.20.20.0 0.0.0.255 area 0 RR-AS1(config-router)# network 2.2.2.2 0.0.0.0 area 0 RR-AS1(config-router)# network 10.10.10.0 0.0.0.255 area 0 RR-AS1(config-router)# end
RR-AS2	RR-AS2# configure terminal RR-AS2(config)# interface loopback 0 RR-AS2(config-if-Loopback 0)# ip address 6.6.6.6 255.255.255.255 RR-AS2# configure terminal RR-AS2(config)# router bgp 1 RR-AS2(config-router)# neighbor 5.5.5.5 remote-as 2 RR-AS2(config-router)# neighbor 5.5.5.5 update-source loopback 0 RR-AS2(config-router)# address-family vpnv4 unicast RR-AS2(config-router-af)# neighbor 5.5.5.5 activate RR-AS2(config-router-af)# neighbor 5.5.5.5 route-reflector-client RR-AS2(config-router-af)# exit RR-AS2(config-router)# address-family ipv4 RR-AS2(config-router-af)# neighbor 5.5.5.5 activate RR-AS2(config-router-af)# neighbor 5.5.5.5 send-label RR-AS2(config-router-af)# neighbor 5.5.5.5 route-reflector-client RR-AS2(config-router-af)# end RR-AS2# configure terminal RR-AS2(config)# router bgp 1 RR-AS2(config-router)# neighbor 2.2.2.2 remote-as 1 RR-AS2(config-router)# neighbor 2.2.2.2 update-source loopback 0 RR-AS2(config-router)# neighbor 2.2.2.2 ebgp-multihop RR-AS2(config-router)# neighbor 8.8.8.8 remote-as 3 RR-AS2(config-router)# neighbor 8.8.8.8 update-source loopback 0 RR-AS2(config-router)# neighbor 8.8.8.8 ebgp-multihop RR-AS2(config-router)# address-family ipv4

```
RR-AS2(config-router-af)# no neighbor 2.2.2.2 activate
RR-AS2(config-router-af)# no neighbor 8.8.8.8 activate
RR-AS2(config-router-af)# exit-address-family
RR-AS2(config-router)# address-family vpnv4 unicast
RR-AS2(config-router-af)# neighbor 2.2.2.2 activate
RR-AS2(config-router-af)# neighbor 2.2.2.2 next-hop-unchanged
RR-AS2(config-router-af)# neighbor 8.8.8.8 activate
RR-AS2(config-router-af)# neighbor 8.8.8.8 next-hop-unchanged
RR-AS2(config-router-af)# end
RR-AS2# configure terminal
RR-AS2(config)# router bgp 1
RR-AS2(config-router)# neighbor 4.4.4.4 remote-as 2
RR-AS2(config-router)# neighbor 4.4.4.4 update-source loopback 0
RR-AS2(config-router)# address-family ipv4
RR-AS2(config-router-af)# neighbor 4.4.4.4 activate
RR-AS2(config-router-af)# neighbor 4.4.4.4 send-label
RR-AS2(config-router-af)# end
RR-AS2# configure terminal
RR-AS2(config)# mpls ip
RR-AS2(config)# mpls router ldp
RR-AS2(config-mpls-router)# ldp router-id interface loopback 0 force
RR-AS2(config-mpls-router)# exit
RR-AS2(config)# interface GigabitEthernet 1/1
RR-AS2(config-if-GigabitEthernet 1/1)# ip address 40.40.40.2 255.255.255.0
RR-AS2(config-if-GigabitEthernet 1/1)# label-switching
RR-AS2(config-if-GigabitEthernet 1/1)# mpls ip
RR-AS2(config-if-GigabitEthernet 1/1)# exit
RR-AS2(config)# interface GigabitEthernet 1/0
RR-AS2(config-if-GigabitEthernet 1/0)# ip address 50.50.50.1 255.255.255.0
RR-AS2(config-if-GigabitEthernet 1/0)# label-switching
RR-AS2(config-if-GigabitEthernet 1/0)# mpls ip
RR-AS2(config-if-GigabitEthernet 1/0)# end
```

	RR-AS2# configure terminal RR-AS2(config)# router ospf 10 RR-AS2(config-router)# network 50.50.50.0 0.0.0.255 area 0 RR-AS2(config-router)# network 6.6.6.6 0.0.0.0 area 0 RR-AS2(config-router)# network 40.40.40.0 0.0.0.255 area 0 RR-AS2(config-router)# end
RR-AS3	RR-AS3# configure terminal RR-AS3(config)# interface loopback 0 RR-AS3(config-if-Loopback 0)# ip address 8.8.8.8 255.255.255.255 RR-AS3# configure terminal RR-AS3(config)# router bgp 1 RR-AS3(config-router)# neighbor 9.9.9.9 remote-as 3 RR-AS3(config-router)# neighbor 9.9.9.9 update-source loopback 0 RR-AS3(config-router)# address-family vpnv4 unicast RR-AS3(config-router-af)# neighbor 9.9.9.9 activate RR-AS3(config-router-af)# neighbor 9.9.9.9 route-reflector-client RR-AS3(config-router-af)# exit RR-AS3(config-router)# address-family ipv4 RR-AS3(config-router-af)# neighbor 9.9.9.9 activate RR-AS3(config-router-af)# neighbor 9.9.9.9 send-label RR-AS3(config-router-af)# neighbor 9.9.9.9 route-reflector-client RR-AS3(config-router-af)# end RR-AS3# configure terminal RR-AS3(config)# router bgp 1 RR-AS3(config-router)# neighbor 6.6.6.6 remote-as 2 RR-AS3(config-router)# neighbor 6.6.6.6 update-source loopback 0 RR-AS3(config-router)# neighbor 6.6.6.6 ebgp-multihop RR-AS3(config-router)# neighbor 2.2.2.2 remote-as 1 RR-AS3(config-router)# neighbor 2.2.2.2 update-source loopback 0 RR-AS3(config-router)# neighbor 2.2.2.2 ebgp-multihop RR-AS3(config-router)# address-family ipv4

```
RR-AS3(config-router-af)# no neighbor 6.6.6.6 activate
RR-AS3(config-router-af)# no neighbor 2.2.2.2 activate
RR-AS3(config-router-af)# exit-address-family
RR-AS3(config-router)# address-family vpnv4 unicast
RR-AS3(config-router-af)# neighbor 6.6.6.6 activate
RR-AS3(config-router-af)# neighbor 6.6.6.6 next-hop-unchanged
RR-AS3(config-router-af)# neighbor 2.2.2.2 activate
RR-AS3(config-router-af)# neighbor 2.2.2.2 next-hop-unchanged
RR-AS3(config-router-af)# end
RR-AS3# configure terminal
RR-AS3(config)# router bgp 1
RR-AS3(config-router)# neighbor 7.7.7.7 remote-as 3
RR-AS3(config-router)# neighbor 7.7.7.7 update-source loopback 0
RR-AS3(config-router)# address-family ipv4
RR-AS3(config-router-af)# neighbor 7.7.7.7 activate
RR-AS3(config-router-af)# neighbor 7.7.7.7 send-label
RR-AS3(config-router-af)# end
RR-AS3# configure terminal
RR-AS3(config)# mpls ip
RR-AS3(config)# mpls router ldp
RR-AS3(config-mpls-router)# ldp router-id interface loopback 0 force
RR-AS3(config-mpls-router)# exit
RR-AS3(config)# interface GigabitEthernet 1/1
RR-AS3(config-if-GigabitEthernet 1/1)# ip address 70.70.70.2 255.255.255.0
RR-AS3(config-if-GigabitEthernet 1/1)# label-switching
RR-AS3(config-if-GigabitEthernet 1/1)# mpls ip
RR-AS3(config-if-GigabitEthernet 1/1)# exit
RR-AS3(config)# interface GigabitEthernet 1/0
RR-AS3(config-if-GigabitEthernet 1/0)# ip address 80.80.80.1 255.255.255.0
RR-AS3(config-if-GigabitEthernet 1/0)# label-switching
RR-AS3(config-if-GigabitEthernet 1/0)# mpls ip
RR-AS3(config-if-GigabitEthernet 1/0)# end
```

	RR-AS3# configure terminal RR-AS3(config)# router ospf 10 RR-AS3(config-router)# network 70.70.70.0 0.0.0.255 area 0 RR-AS3(config-router)# network 8.8.8.8 0.0.0.0 area 0 RR-AS3(config-router)# network 80.80.80.0 0.0.0.255 area 0 RR-AS3(config-router)# end
ASBR1	ASBR1# configure terminal ASBR1(config)# interface loopback 0 ASBR1(config-if-Loopback 0)# ip address 3.3.3.3 255.255.255.255 ASBR1# configure terminal ASBR1(config)# ip access-list extended 101 ASBR1(config-ext-nacl)# permit ip host 1.1.1.1 any ASBR1(config-ext-nacl)# exit ASBR1(config)# ip access-list extended 102 ASBR1(config-ext-nacl)# permit ip host 5.5.5.5 any ASBR1(config-ext-nacl)# permit ip host 9.9.9.9 any ASBR1(config-ext-nacl)# exit ASBR1(config)# route-map internal-mpls-route permit 10 ASBR1(config-route-map)# match ip address 101 ASBR1(config-route-map)# set mpls-label ASBR1(config-route-map)# exit ASBR1(config)# route-map external-mpls-route permit 10 ASBR1(config-route-map)# match ip address 102 ASBR1(config-route-map)# set mpls-label ASBR1(config-route-map)# end ASBR1# configure terminal ASBR1(config)# router bgp 1 ASBR1(config-router)# neighbor 30.30.30.2 remote-as 2 ASBR1(config-router)# neighbor 2.2.2.2 remote-as 1 ASBR1(config-router)# neighbor 2.2.2.2 update-source loopback 0 ASBR1(config-router)# address-family ipv4

	ASBR1(config-router-af)# neighbor 30.30.30.2 send-label ASBR1(config-router-af)# neighbor 30.30.30.2 route-map internal-mpls-route out ASBR1(config-router-af)# neighbor 2.2.2.2 send-label ASBR1(config-router-af)# neighbor 2.2.2.2 route-map external-mpls-route out ASBR1(config-router-af)# network 1.1.1.1 mask 255.255.255.255 ASBR1(config-router-af)# end ASBR1# configure terminal ASBR1(config)# mpls ip ASBR1(config)# mpls router ldp ASBR1(config-mpls-router)# ldp router-id interface loopback 0 force ASBR1(config-mpls-router)# exit ASBR1(config)# interface GigabitEthernet 1/1 ASBR1(config-if-GigabitEthernet 1/1)#ip address 10.10.10.2 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/1)# label-switching ASBR1(config-if-GigabitEthernet 1/1)# mpls ip ASBR1(config-if-GigabitEthernet 1/1)# end ASBR1# configure terminal ASBR1(config)# router ospf 10 ASBR1(config-router)# network 10.10.10.0 0.0.0.255 area 0 ASBR1(config-router)# network 3.3.3.3 0.0.0.0 area 0 ASBR1(config-router)# end ASBR1(config)# interface GigabitEthernet 1/0 ASBR1(config-if-GigabitEthernet 1/0)# ip address 30.30.30.1 255.255.255.0 ASBR1(config-if-GigabitEthernet 1/0)# label-switching
ASBR2	ASBR2# configure terminal ASBR2(config)# interface loopback 0 ASBR2(config-if-Loopback 0)# ip address 4.4.4.4 255.255.255.255 ASBR2# configure terminal ASBR2(config)# ip access-list extended 101 ASBR2(config-ext-nacl)# permit ip host 5.5.5.5 any ASBR2(config-ext-nacl)# exit

```
ASBR2(config)# ip access-list extended 102
ASBR2(config-ext-nacl)# permit ip host 1.1.1.1 any
ASBR2(config-ext-nacl)# permit ip host 9.9.9.9 any
ASBR2(config-ext-nacl)# exit
ASBR2(config)# route-map internal-mpls-route permit 10
ASBR2(config-route-map)# match ip address 101
ASBR2(config-route-map)# set mpls-label
ASBR2(config-route-map)# exit
ASBR2(config)# route-map external-mpls-route permit 10
ASBR2(config-route-map)# match ip address 102
ASBR2(config-route-map)# set mpls-label
ASBR2(config-route-map)# end
ASBR2# configure terminal
ASBR2(config)# router bgp 1
ASBR2(config-router)# neighbor 30.30.30.1 remote-as 1
ASBR2(config-router)# neighbor 60.60.60.2 remote-as 3
ASBR2(config-router)# neighbor 6.6.6.6 remote-as 1
ASBR2(config-router)# neighbor 6.6.6.6 update-source loopback 0
ASBR2(config-router)# address-family ipv4
ASBR2(config-router-af)# neighbor 30.30.30.1 send-label
ASBR2(config-router-af)# neighbor 30.30.30.1 route-map internal-mpls-route out
ASBR2(config-router-af)# neighbor 60.60.60.2 send-label
ASBR2(config-router-af)# neighbor 60.60.60.2 route-map internal-mpls-route out
ASBR2(config-router-af)# neighbor 6.6.6.6 send-label
ASBR2(config-router-af)# neighbor 6.6.6.6 route-map external-mpls-route out
ASBR2(config-router-af)# network 5.5.5.5 mask 255.255.255.255
ASBR2(config-router-af)# end
ASBR2# configure terminal
ASBR2(config)# mpls ip
ASBR2(config)# mpls router ldp
ASBR2(config-mpls-router)# ldp router-id interface loopback 0 force
ASBR2(config-mpls-router)# exit
```

	ASBR2(config)# interface GigabitEthernet 1/1 ASBR2(config-if-GigabitEthernet 1/1)#ip address 30.30.30.2 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/1)# label-switching ASBR2(config-if-GigabitEthernet 1/1)# mpls ip ASBR2(config-if-GigabitEthernet 1/1)# exit ASBR2(config)# interface GigabitEthernet 1/2 ASBR2(config-if-GigabitEthernet 1/2)#ip address 60.60.60.1 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/2)# label-switching ASBR2(config-if-GigabitEthernet 1/2)# mpls ip ASBR2(config-if-GigabitEthernet 1/2)# end ASBR2# configure terminal ASBR2(config)# router ospf 10 ASBR2(config-router)# network 40.40.40.0 0.0.0.255 area 0 ASBR2(config-router)# network 4.4.4.4 0.0.0.0 area 0 ASBR2(config-router)# end ASBR2(config)# interface GigabitEthernet 1/0 ASBR2(config-if-GigabitEthernet 1/0)# ip address 40.40.40.1 255.255.255.0 ASBR2(config-if-GigabitEthernet 1/0)# label-switching
ASBR3	ASBR3# configure terminal ASBR3(config)# interface loopback 0 ASBR3(config-if-Loopback 0)# ip address 7.7.7.7 255.255.255.255 ASBR3# configure terminal ASBR3(config)# ip access-list extended 101 ASBR3(config-ext-nacl)# permit ip host 9.9.9.9 any ASBR3(config-ext-nacl)# exit ASBR3(config)# ip access-list extended 102 ASBR3(config-ext-nacl)# permit ip host 5.5.5.5 any ASBR3(config-ext-nacl)# permit ip host 1.1.1.1 any ASBR3(config-ext-nacl)# exit ASBR3(config)# route-map internal-mpls-route permit 10 ASBR3(config-route-map)# match ip address 101

```
ASBR3(config-route-map)# set mpls-label
ASBR3(config-route-map)# exit
ASBR3(config)# route-map external-mpls-route permit 10
ASBR3(config-route-map)# match ip address 102
ASBR3(config-route-map)# set mpls-label
ASBR3(config-route-map)# end
ASBR3# configure terminal
ASBR3(config)# router bgp 1
ASBR3(config-router)# neighbor 60.60.60.1 remote-as 2
ASBR3(config-router)# neighbor 8.8.8.8 remote-as 1
ASBR3(config-router)# neighbor 8.8.8.8 update-source loopback 0
ASBR3(config-router)# address-family ipv4
ASBR3(config-router-af)# neighbor 60.60.60.1 send-label
ASBR3(config-router-af)# neighbor 60.60.60.1 route-map internal-mpls-route out
ASBR3(config-router-af)# neighbor 8.8.8.8 send-label
ASBR3(config-router-af)# neighbor 8.8.8.8 route-map external-mpls-route out
ASBR3(config-router-af)# network 9.9.9.9 mask 255.255.255.255
ASBR3(config-router-af)# end
ASBR3# configure terminal
ASBR3(config)# mpls ip
ASBR3(config)# mpls router ldp
ASBR3(config-mpls-router)# ldp router-id interface loopback 0 force
ASBR3(config-mpls-router)# exit
ASBR3(config)# interface GigabitEthernet 1/0
ASBR3(config-if-GigabitEthernet 1/0)#ip address 70.70.70.1 255.255.255.0
ASBR3(config-if-GigabitEthernet 1/0)# label-switching
ASBR3(config-if-GigabitEthernet 1/0)# mpls ip
ASBR3(config-if-GigabitEthernet 1/0)# end
ASBR3# configure terminal
ASBR3(config)# router ospf 10
ASBR3(config-router)# network 70.70.70.0 0.0.0.255 area 0
ASBR3(config-router)# network 7.7.7.7 0.0.0.0 area 0
```

	ASBR3(config-router)# end ASBR3(config)# interface GigabitEthernet 1/1 ASBR3(config-if-GigabitEthernet 1/1)# ip address 60.60.60.2 255.255.255.0 ASBR3(config-if-GigabitEthernet 1/1)# label-switching
	● Run the **show mpls forwarding-table** command to display MPLS VPN routes. ● Check that different sites of the same VPN can communicate with each other.
PE1-AS1	Assume that the loopback interfaces on the peer CE are 20.20.20.20 and 30.30.30.30. 20.20.20.20 is pingable from the local CE. 30.30.30.30 is pingable from the local CE. 5.5.5.5 is unpingable. The source address 1.1.1.1 of 5.5.5.5 is pingable. 9.9.9.9 is unpingable. The source address 1.1.1.1 of 9.9.9.9 is pingable.
PE1-AS2	Assume that the loopback interfaces on the peer CE are 10.10.10.10 and 30.30.30.30. 10.10.10.10 is pingable from the local CE. 30.30.30.30 is pingable from the local CE. 1.1.1.1 is unpingable. The source address 5.5.5.5 of 1.1.1.1 is pingable. 9.9.9.9 is unpingable. The source address 5.5.5.5 of 9.9.9.9 is pingable.
PE1-AS3	Assume that the loopback interfaces on the peer CE are 10.10.10.10 and 0.20.20.20. 10.10.10.10 is pingable from the local CE. 20.20.20.20 is pingable from the local CE. 1.1.1.1 is unpingable. The source address 9.9.9.9 of 1.1.1.1 is pingable. 5.5.5.5 is unpingable. The source address 9.9.9.9 of 5.5.5.5 is pingable.
RR-AS1	show bgp vpnv4 unicast all The 10.10.10.10, 20.20.20.20, and 30.30.30.30 entries exist.
RR-AS2	show bgp vpnv4 unicast all

	The 10.10.10.10, 20.20.20.20, and 30.30.30.30 entries exist.
RR-AS3	show bgp vpnv4 unicast all The 10.10.10.10, 20.20.20.20, and 30.30.30.30 entries exist.
ASBR1	30.30.30.2 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 30.30.30.30 is unpingable. 4.4.4.4 is unpingable. 7.7.7.7 is unpingable.
ASBR2	30.30.30.1 is pingable. 60.60.60.2 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 30.30.30.30 is unpingable. 3.3.3.3 is unpingable. 7.7.7.7 is unpingable.
ASBR3	60.60.60.1 is pingable. 10.10.10.10 is unpingable. 20.20.20.20 is unpingable. 30.30.30.30 is unpingable. 3.3.3.3 is unpingable. 4.4.4.4 is unpingable.

Common Errors

- The capability of exchanging IPv4 routes is not disabled when a multi-hop MP-EBGP session is configured between RRs, or next hop retention is not configured for VPN routes exchanged over the multi-hop MP-EBGP session, causing a communication failure between VPN sites.

2.4.5 Configuring OSPF VPN Extended Attributes

Configuration Effect

When VPN user networks run OSPF, implement OSPF between PE and CE to simplify CE configuration and management.

Notes

- The LDP router ID must consist of 32 bits.

Configuration Steps

- (Optional) Configure a domain ID.
- (Optional) Configure a VPN route tag.
- (Optional) Configure a sham link.
- (Optional) Configure capability vrf-lite.
- (Optional) Configure extcommunity-type.

(Optional) Configuring a Domain ID

A domain ID indicates the domain to which an OSPF instance belongs. Typically, all VRF OSPF instances of the same VPN are configured with the same domain ID. The default domain ID of an OSPF instance is Null, and the default type is 0x0005.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router ospf** *ospf-id* **vrf** *vrf-name*	Creates an OSPF instance and enters OSPF configuration mode.
FS(config-router)# **domain-id** { *ip-address* [secondary] | **null** | **type** { **0005** | **0105** | **0205** | **8005** } **value** *hex-value* [**secondary**] }	Configures the domain ID of the OSPF instance.
FS(config-router)# **show running-config**	Displays existing configurations.

The **domain-id** command takes effect only for VRF OSPF instances.

A VRF OSPF instance can be configured with multiple domain IDs, but only one domain ID is the primary domain ID, and the other domain IDs are secondary domain IDs. The primary domain ID is configured using the **domain-id** *value* command, and the secondary domain IDs are configured using the **domain-id** *value* **secondary** command. The VPN route converted from an OSPF route contains only the primary domain ID.

The primary and secondary domain IDs can be configured using the **domain-id** *ip-address* command or the **domain-id type** { **0005** | **0105** | **0205** | **8005** } *value* command.

You can configure the domain IDs of different VRF OSPF instances to be the same. The VRF OSPF instances of the same VPN must be configured with the same domain ID to ensure correct route advertisement.

(Optional) Configuring a VPN Route Tag

By default, the VPN route of an OSPF instance is marked by the AS number of local BGP.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router ospf** *ospf-id* **vrf** *vrf-name*	Creates an OSPF instance and enters OSPF configuration mode.
FS(config-router)# **domain-tag** *value*	Configures a VPN route tag for the OSPF instance. The value range is <1–4294967295>.
FS(config-router)# **show running-config**	Displays existing configurations.

The **domain-tag** *value* command takes effect only for VRF OSPF instances.

If VRF-specific **domain-tag** is not configured manually, the default value is the AS number of local BGP.

In L3VPN, if a VPN site is connected to multiple PEs, the VPN route that a PE learns over MP-BGP is advertised to the VPN site over a type 5 or 7 LSA. If another PE connected to the VPN site learns and advertises the VPN route, a loop occurs. To avoid the loop, configure the same VPN route tag for the VRF OSPF instances of the PEs connected to the same VPN site. The type 5 or 7 LSA that a VRF OSPF instance sends to the VPN site carries a VPN route tag. When another PE receives the type 5 or 7 LSA, it checks whether the VPN route tag in the LSA is the same as that of the corresponding OSPF instance. If they are the same, the LSA is excluded from OSPF route calculation.

Typically, OSPF instances of the same VPN are configured with the same VPN route tag.

(Optional) Configuring a Sham Link

A sham link is applicable in the scenario where VPN sites are connected by a backdoor link. If you require that VPN data be still transmitted over the MPLS backbone network, create a sham link between the VRF OSPF instances on two PEs. In this way, the VRF OSPF instances establish an OSPF neighbor relationship over the sham link to distribute LSA packets.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router ospf** *ospf-id* **vrf** *vrf-name*	Creates an OSPF instance and enters OSPF configuration mode.
FS(config-router)# **area** *area-id* **sham-link** *source-address destination-address* [**cost** *number*] [**dead-interval** *seconds*] [**hello-interval** *seconds*] [**retransmit-interval** *seconds*] [**transmit-delay** *seconds*] [**authentication** [**message-digest | null**]] [[**authentication-key** [**0**|**7**] *key*] **|** [**message-digest-key** *key-id* **md5** [**0**|**7**] *key*]]	Configures the area ID, source address, and destination address of the sham link.
FS(config-router)# **show running-config**	Displays existing configurations.

A sham link must be configured on two PEs. A single PE does not support sham link creation.

The two PEs for sham link creation must meet the following configuration requirements:

- The PEs must be configured with the same area ID of the sham link.
- The PEs must be configured with the same <source address, destination address> combination of the sham link.
- The source address and destination address of the sham link must be the 32-bit loopback interface addresses bound to the VRF tables on PEs.

OSPF routes advertised over a sham link are not used for forwarding because the routes do not have a VPN route tag. Forwarding is still implemented by BGP VPNv4 routes. Therefore, ensure that routes learned over the sham link can also be learned over MP-BGP.

The source address used for sham link creation must be used by VPNv4 route advertisement over BGP, but the source address cannot be used by VRF OSPF instances for route calculation.

(Optional) Configuring capability vrf-lite

By default, VRF OSPF instances support the PE-CE OSPF feature (LSA conversion based on the domain ID, DN bit, and VPN route tag). You can run the **capability vrf-lite** command to disable the PE-CE OSPF feature for VRF OSPF instances.

Command	Description

FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router ospf** *ospf-id* **vrf** *vrf-name*	Creates an OSPF instance and enters OSPF configuration mode.
FS(config-router)# **capability vrf-lite** [**auto**]	Disables the PE-CE OSPF feature for VRF OSPF instances.
FS(config-router)# **show running-config**	Displays existing configurations.

The **capability vrf-lite** command takes effect only for VRF OSPF instances.

The loop detection function can be disabled for VRF OSPF instances in some application scenarios. For example, if VPN routes are exchanged between MCE and PE over OSPF, you need to run the **capability vrf-lite** command to disable the loop detection function for the VRF OSPF instance on the MCE to allow for route learning between VPN sites.

(Optional) Configuring extcommunity-type

After an OSPF route is redistributed to BGP to form a VPN route, the VPN route carries the extended community attributes of the OSPF route, including the router-id and route-type extended community attributes. By default, the router-id extended community attribute belongs to the 0x0107 type, and the route-type extended community attribute belongs to the 0x0306 type. You can use a command to manually configure the types of the router-id and route-type extended community attributes.

Command	Description
FS# **configure terminal**	Enters global configuration mode.
FS(config)# **router ospf** *ospf-id* **vrf** *vrf-name*	Creates an OSPF instance and enters OSPF configuration mode.
FS(config-router)# **extcommunity-type** { **router-id** { **0107 | 8001** } | **route-type** { **0306** | **8000** } }	Configures the OSPF router-id and route-type extended community attributes.
FS(config-router)# **show running-config**	Displays existing configurations.

The **extcommunity-type** command takes effect only for VRF OSPF instances.

The router-id type configuration is intended for compatibility with different vendors. For example, when you interwork with a vendor that only supports the 0x0107 type, you need to run the **extcommunity-type** command to change the type of router-id to 0x0107.

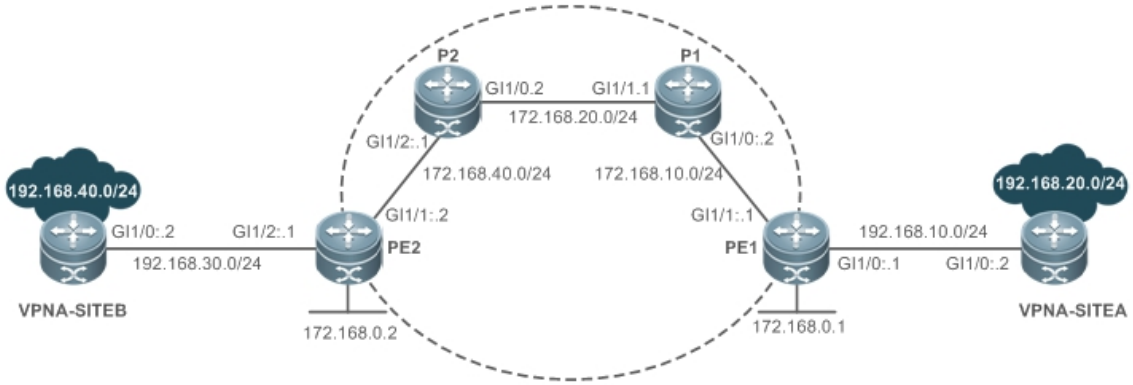
The router-type type configuration is intended for compatibility with different vendors. For example, when you interwork with a vendor that only supports the 0x8000 type, you need to run the **extcommunity-type** command to change the type of router-type to 0x8000.

Verification

- Check that VPN sites can communicate with each other.

Configuration Example

Configuring a Domain ID

Scenario **Figure 2- 39**	 Two sites of a customer exchange VPN routes over the MPLS backbone network. The sites access PEs over OSPF. The OSPF routes that are exchanged over the MPLS backbone network must be restored to the maximum extent to OSPF routes supported by the sites.
Configuration Steps	● On SiteA and SiteB, configure OSPF used for exchanging routes between PE and CE. ● On PE1 and PE2, configure a loopback interface, create a VRF table named VPNA, define RD and RT values, and associate the VRF table with the corresponding interface. Associate the CE-connected interface with the VRF table. Configure BGP, and establish an MP-IBGP session with PE2. Configure the capability of exchanging routes with the CE over OSPF. Set the domain ID of the OSPF instance to 10. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On P1 and P2, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol.
VPNA-SiteA	VPNA-SITEA# configure terminal VPNA-SITEA(config)# router ospf 10 VPNA-SITEA(config-router)# network 192.168.10.0 255.255.255.0 area 0
VPNA-SiteB	VPNA-SITEB# configure terminal VPNA-SITEB(config)# router ospf 10 VPNA-SITEB(config-router)# network 192.168.30.0 255.255.255.0 area 0
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-Loopback 0)# ip address 172.168.0.1 255.255.255.255 PE1# configure terminal PE1(config)# ip vrf VPNA PE1(config-vrf)# rd 1:100 PE1(config-vrf)# route-target both 1:100 PE1(config-vrf)# end

```
PE1# configure terminal
PE1(config)# interface GigabitEthernet 1/0
PE1(config-GigabitEthernet 1/0)# ip vrf forwarding VPNA
PE1(config-GigabitEthernet 1/0)# ip address 192.168.10.1 255.255.255.0
PE1(config-GigabitEthernet 1/0)# end
PE1# configure terminal
PE1(config)# router bgp 1
PE1(config-router)# neighbor 172.168.0.2 remote-as 1
PE1(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE1(config-router)# address-family vpnv4
PE1(config-router-af)# neighbor 172.168.0.2 activate
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# router ospf 10 vrf VPNA
PE1(config-router)# network 192.168.10.0 255.255.255.0 area 0
PE1(config-router)# redistribute bgp subnets
PE1(config-router)# domain-id 10.10.10.10
PE1(config-router)# exit
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf VPNA
PE1(config-router-af)# redistribute ospf 10
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# mpls ip
PE1(config)# mpls router ldp
PE1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1(config-mpls-router)# exit
PE1(config)# interface GigabitEthernet 1/1
PE1(config-GigabitEthernet 1/1)# ip address 172.168.10.1 255.255.255.0
PE1(config-GigabitEthernet 1/1)# label-switching
PE1(config-GigabitEthernet 1/1)# mpls ip
PE1(config-GigabitEthernet 1/1)# end
```

	PE1# configure terminal PE1(config)# router ospf 1 PE1(config-router)# network 172.168.10.0 0.0.0.255 area 0 PE1(config-router)# network 172.168.0.1 0.0.0.0 area 0 PE1(config-router)# end
P1	P1# configure terminal P1(config)# mpls ip P1(config)# interface loopback 0 P1(config-Loopback 0)# ip address 172.168.0.3 255.255.255.255 P1(config-Loopback 0)# exit P1(config)# router ospf 1 P1(config-router)# network 172.168.0.3 0.0.0.0 area 0 P1(config-router)# network 172.168.10.0 0.0.0.255 area 0 P1(config-router)# network 172.168.20.0 0.0.0.255 area 0 P1(config-router)# exit P1(config)# interface GigabitEthernet 1/0 P1(config-GigabitEthernet 1/0)# ip address 172.168.10.2 255.255.255.0 P1(config-GigabitEthernet 1/0)# label-switching P1(config-GigabitEthernet 1/0)# mpls ip P1(config-GigabitEthernet 1/0)# exit P1(config)# interface GigabitEthernet 1/1 P1(config-GigabitEthernet 1/1)# ip address 172.168.20.1 255.255.255.0 P1(config-GigabitEthernet 1/1)# label-switching P1(config-GigabitEthernet 1/1)# mpls ip P1(config-GigabitEthernet 1/1)# exit P1(config)# mpls router ldp P1(config-mpls-router)# ldp router-id interface loopback 0 force P2(config-mpls-router)# exit
P2	P2# configure terminal P2(config)# mpls ip P2(config)# interface loopback 0

	P2(config-Loopback 0)# ip address 172.168.0.4 255.255.255.255 P2(config-Loopback 0)# exit P2(config)# router ospf 1 P2(config-router)# network 172.168.0.4 0.0.0.0 area 0 P2(config-router)# network 172.168.20.0 0.0.0.255 area 0 P2(config-router)# network 172.168.40.0 0.0.0.255 area 0 P2(config-router)# exit P2(config)# interface GigabitEthernet 1/0 P2(config-GigabitEthernet 1/0)# ip address 172.168.20.2 255.255.255.0 P2(config-GigabitEthernet 1/0)# label-switching P2(config-GigabitEthernet 1/0)# mpls ip P2(config-GigabitEthernet 1/0)# exit P2(config)# interface GigabitEthernet 1/2 P2(config-GigabitEthernet 1/2)# ip address 172.168.40.1 255.255.255.0 P2(config-GigabitEthernet 1/2)# label-switching P2(config-GigabitEthernet 1/2)# mpls ip P2(config-GigabitEthernet 1/2)# exit P2(config)# mpls router ldp P2(config-mpls-router)# ldp router-id interface loopback 0 force P2(config-mpls-router)# exit
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-Loopback 0)# ip address 172.168.0.2 255.255.255.255 PE2# configure terminal PE2(config)# ip vrf VPNA PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target both 1:100 PE2(config-vrf)# exit PE2# configure terminal PE2(config)# interface GigabitEthernet 1/2 PE2(config-GigabitEthernet 1/2)# ip vrf forwarding VPNA

```
PE2(config-GigabitEthernet 1/2)# ip address 192.168.30.1 255.255.255.0
PE2(config-GigabitEthernet 1/0)# exit
PE2# configure terminal
PE2(config)# router bgp 1
PE2(config-router)# neighbor 172.168.0.1 remote-as 1
PE2(config-router)# neighbor 172.168.0.1 update-source loopback 0
PE2(config-router)# address-family vpnv4
PE2(config-router-af)# neighbor 172.168.0.1 activate
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# router ospf 10 vrf VPNA
PE2(config-router)# network 192.168.30.0 255.255.255.0 area 0
PE2(config-router)# redistribute bgp subnets
PE2(config-router)# domain-id 10.10.10.10
PE2(config-router)# exit
PE2(config)# router bgp 1
PE2(config-router)# address-family ipv4 vrf VPNA
PE2(config-router-af)# redistribute ospf 10
PE2(config-router-af)# exit
PE2# configure terminal
PE2(config)# mpls ip
PE2(config)# mpls router ldp
PE2(config-mpls-router)# ldp router-id interface loopback 0 force
PE2(config-mpls-router)# exit
PE2(config)# interface GigabitEthernet 1/1
PE2(config-GigabitEthernet 1/1)# ip address 172.168.40.2 255.255.255.0
PE2(config-GigabitEthernet 1/1)# label-switching
PE2(config-GigabitEthernet 1/1)# mpls ip
PE2(config-GigabitEthernet 1/1)# end
PE2# configure terminal
PE2(config)# router ospf 1
PE2(config-router)# network 172.168.40.0 0.0.0.255 area 0
```

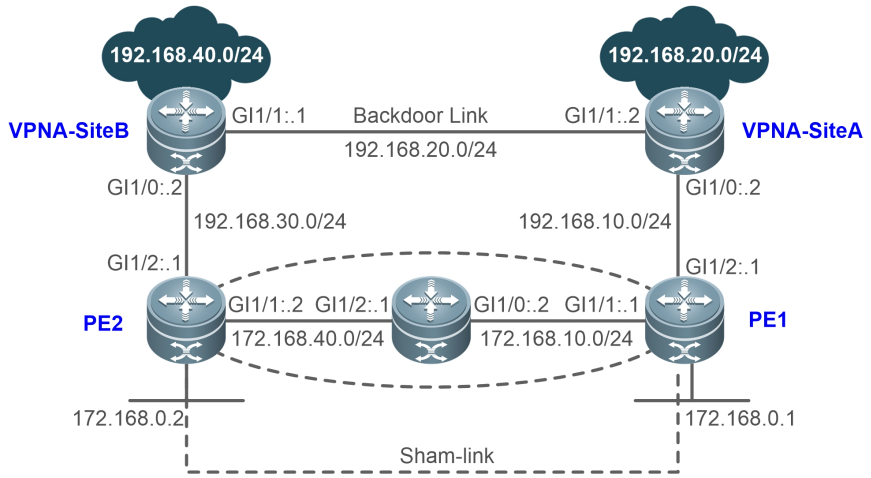
	PE2(config-router)# network 172.168.0.2 0.0.0.0 area 0 PE2(config-router)# end
Verification	● Run the **show mpls forwarding-table** command on PEs to check that the routes of VPNA-SiteA and VPNA-SiteB exist.
VPNA-SiteB	VPNA-SITEB# show ip route Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set O IA 192.168.10.0/24 [110/2] via 192.168.30.1, 00:00:36, GigabitEthernet 1/0 O IA 192.168.20.0/24 [110/2] via 192.168.30.1, 00:00:36, GigabitEthernet 1/0 C 192.168.30.0/24 is directly connected, GigabitEthernet 1/0 O 192.168.40.0/24 [110/101] via 192.168.24.2, 00:56:23, GigabitEthernet 1/1
PE2	PE2# show ip route vrf VPNA Routing Table: VPNA Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set B 192.168.10.0/24 [110/2] via 172.168.0.1, 00:00:36 B 192.168.20.0/24 [110/2] via 172.168.0.1, 00:00:36 C 192.168.30.0/24 is directly connected, GigabitEthernet 1/2

	O 192.168.40.0/24 [110/101] via 192.168.30.2, 00:56:23, GigabitEthernet 1/2
PE1	PE1# show ip route vrf VPNA Routing Table: VPNA Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.10.0/24 is directly connected, GigabitEthernet 1/0 O 192.168.20.0/24 [110/101] via 192.168.10.2, 00:56:23, GigabitEthernet 1/0 B 192.168.30.0/24 [110/2] via 172.168.0.2, 00:00:36 B 192.168.40.0/24 [110/2] via 172.168.0.2, 00:00:36
VPNA-SiteA	VPNA-SITEA# show ip route Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.10.0/24 is directly connected, GigabitEthernet 1/0 O 192.168.20.0/24 [110/101] via 192.168.23.2, 00:56:23, GigabitEthernet 1/1 O IA 192.168.30.0/24 [110/2] via 192.168.10.1, 00:00:36, GigabitEthernet 1/0 O IA 192.168.40.0/24 [110/2] via 192.168.10.1, 00:00:36, GigabitEthernet 1/0

Configuring a Sham Link

Scenario

Figure 2-40



Two sites of a customer exchange VPN routes over the MPLS backbone network. A backdoor link is established between the sites to exchange VPN routes when the MPLS backbone network fails.

Configuration Steps

- On SiteA, configure OSPF for communicating with PE1 and configure OSPF over the backdoor link for communicating with SiteB. Set the OSPF cost in interface configuration mode.
- On SiteB, configure OSPF for communicating with PE2 and configure OSPF over the backdoor link for communicating with SiteA. Set the OSPF cost in interface configuration mode.
- On PE1, configure a loopback interface, create a VRF table named VPNA, define RD and RT values, and associate the VRF table with the corresponding interface. Associate the CE-connected interface with the VRF table. Configure a loopback interface for the VRF table for establishing a sham link. Configure BGP, and establish an MP-IBGP session with PE2. Configure the capability of exchanging routes with the CE over OSPF. Establish a sham link with the OSPF instance on PE2. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On PE2, configure a loopback interface, create a VRF table named VPNA, define RD and RT values, and associate the VRF table with the corresponding interface. Associate the CE-connected interface with the VRF table. Configure a loopback interface for the VRF table for establishing a sham link. Configure BGP, and establish an MP-IBGP session with PE1. Configure the capability of exchanging routes with the CE over OSPF. Establish a sham link with PE1. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol.
- On P1, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol.

SiteA

```
SITEA# configure terminal
SITEA(config)# router ospf 10
SITEA(config-router)# network 192.168.10.0 255.255.255.0 area 0
SITEA(config-router)# network 192.168.20.0 255.255.255.0 area 0
SITEA# configure terminal
SITEA(config)# interface GigabitEthernet 1/0
```

	SITEA(config-GigabitEthernet 1/0)# ip address 192.168.10.2 255.255.255.0 SITEA(config-GigabitEthernet 1/0)# ip ospf cost 1 SITEA(config)# interface GigabitEthernet 1/1 SITEA(config-GigabitEthernet 1/1)# ip address 192.168.20.1 255.255.255.0 SITEA(config-GigabitEthernet 1/1)# ip ospf cost 200
SiteB	SITEB# configure terminal SITEB(config)# router ospf 10 SITEB(config-router)# network 192.168.30.0 255.255.255.0 area 0 SITEB(config-router)# network 192.168.20.0 255.255.255.0 area 0 SITEB# configure terminal SITEB(config)# interface GigabitEthernet 1/0 SITEB(config-GigabitEthernet 1/0)# ip address 192.168.30.2 255.255.255.0 SITEB(config-GigabitEthernet 1/0)# ip ospf cost 1 SITEB(config)# interface GigabitEthernet 1/1 SITEB(config-GigabitEthernet 1/1)# ip address 192.168.20.2 255.255.255.0 SITEB(config-GigabitEthernet 1/1)# ip ospf cost 200
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-Loopback 0)# ip address 172.168.0.1 255.255.255.255 PE1(config-Loopback 0)# end PE1# configure terminal PE1(config)# ip vrf VPNA PE1(config-vrf)# rd 1:100 PE1(config-vrf)# route-target both 1:100 PE1(config-vrf)# end PE1# configure terminal PE1(config)# interface GigabitEthernet 1/2 PE1(config-GigabitEthernet 1/2)# ip vrf forwarding VPNA PE1(config-GigabitEthernet 1/2)# ip address 192.168.10.1 255.255.255.0 PE1(config-GigabitEthernet 1/2)# end PE1# configure terminal

```
PE1(config)# interface loopback 10
PE1(config-Loopback 10)# ip vrf forwarding VPNA
PE1(config-Loopback 10)# ip address 192.168.0.1 255.255.255.255
PE1# configure terminal
PE1(config)# router bgp 1
PE1(config-router)# neighbor 172.168.0.2 remote-as 1
PE1(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE1(config-router)# address-family vpnv4
PE1(config-router-af)# neighbor 172.168.0.2 activate
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# router ospf 10 vrf VPNA
PE1(config-router)# network 192.168.10.0 255.255.255.0 area 0
PE1(config-router)# redistribute bgp subnets
PE1(config-router)# area 0 sham-link 192.168.0.1 192.168.0.2
PE1(config-router)# exit
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf VPNA
PE1(config-router-af)# redistribute ospf 10
PE1(config-router-af)# redistribute connected
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# mpls ip
PE1(config)# mpls router ldp
PE1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1(config-mpls-router)# exit
PE1(config)# interface GigabitEthernet 1/1
PE1(config-GigabitEthernet 1/1)# ip address 172.168.10.1 255.255.255.0
PE1(config-GigabitEthernet 1/1)# label-switching
PE1(config-GigabitEthernet 1/1)# mpls ip
PE1(config-GigabitEthernet 1/1)# end
PE1# configure terminal
```

	PE1(config)# router ospf 1 PE1(config-router)# network 172.168.10.0 0.0.0.255 area 0 PE1(config-router)# network 172.168.0.1 0.0.0.0 area 0 PE1(config-router)# end
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-Loopback 0)# ip address 172.168.0.2 255.255.255.255 PE2(config-Loopback 0)# end PE2# configure terminal PE2(config)# ip vrf VPNA PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target both 1:100 PE2(config-vrf)# exit PE2# configure terminal PE2(config)# interface GigabitEthernet 1/2 PE2(config-GigabitEthernet 1/2)# ip vrf forwarding VPNA PE2(config-GigabitEthernet 1/2)# ip address 192.168.30.1 255.255.255.0 PE2(config-GigabitEthernet 1/2)# exit PE2# configure terminal PE2(config)# interface loopback 10 PE2(config-Loopback 10)# ip vrf forwarding VPNA PE2(config-Loopback 10)# ip address 192.168.0.2 255.255.255.255 PE2# configure terminal PE2(config)# router bgp 1 PE2(config-router)# neighbor 172.168.0.1 remote-as 1 PE2(config-router)# neighbor 172.168.0.1 update-source loopback 0 PE2(config-router)# address-family vpnv4 PE2(config-router-af)# neighbor 172.168.0.1 activate PE2(config-router-af)# end PE2# configure terminal PE2(config)# router ospf 10 vrf VPNA

	PE2(config-router)# network 192.168.30.0 255.255.255.0 area 0 PE2(config-router)# redistribute bgp subnets PE2(config-router)# area 0 sham-link 192.168.0.2 192.168.0.1 PE2(config-router)# exit PE2(config)# router bgp 1 PE2(config-router)# address-family ipv4 vrf VPNA PE2(config-router-af)# redistribute ospf 10 PE2(config-router-af)# redistribute connected PE2(config-router-af)# exit PE2# configure terminal PE2(config)# mpls ip PE2(config)# mpls router ldp PE2(config-mpls-router)# ldp router-id interface loopback 0 force PE2(config-mpls-router)# exit PE2(config)# interface GigabitEthernet 1/1 PE2(config-GigabitEthernet 1/1)# ip address 172.168.40.2 255.255.255.0 PE2(config-GigabitEthernet 1/1)# label-switching PE2(config-GigabitEthernet 1/1)# mpls ip PE2(config-GigabitEthernet 1/1)# end PE2# configure terminal PE2(config)# router ospf 1 PE2(config-router)# network 172.168.40.0 0.0.0.255 area 0 PE2(config-router)# network 172.168.0.2 0.0.0.0 area 0 PE2(config-router)# end
P1	P1# configure terminal P1(config)# interface loopback 0 P1(config-Loopback 0)# ip address 172.168.0.3 255.255.255.255 P1(config-Loopback 0)# exit P1(config)# interface GigabitEthernet 1/0 P1(config-GigabitEthernet 1/0)# ip address 172.168.10.2 255.255.255.0 P1(config-GigabitEthernet 1/0)# mpls ip

```
P1(config-GigabitEthernet 1/0)# label-switch
P1(config-GigabitEthernet 1/0)# exit
P1(config)# interface GigabitEthernet 1/1
P1(config-GigabitEthernet 1/1)# ip address 172.168.40.1 255.255.255.0
P1(config-GigabitEthernet 1/1)# mpls ip
P1(config-GigabitEthernet 1/1)# label-switch
P1(config-GigabitEthernet 1/1)# exit
P1(config)# router ospf 1
P2(config-router)# network 172.168.40.0 0.0.0.255 area 0
P2(config-router)# network 172.168.10.0 0.0.0.255 area 0
P2(config-router)# network 172.168.0.3 0.0.0.0 area 0
P2(config-router)# exit
P1(config)# mpls ip
P1(config)# mpls router ldp
PE2(config-mpls-router)# ldp router-id interface loopback 0 force
PE2(config-mpls-router)# exit
```

Verification

- Check that the OSPF sham link route exists on PEs

PE1

```
PE1# show ip ospf 10 sham-links
Sham Link SLINK0 to address 192.168.0.2 is up
  Area 0.0.0.0 source address 192.168.0.1, Cost: 1
  Output interface is GigabitEthernet 1/1
  Nexthop address 172.16.40.2
  Transmit Delay is 1 sec, State Point-To-Point,
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:01
    Adjacency state Full

PE1# show ip ospf 10 neighbor

OSPF process 10, 1 Neighbors, 1 is Full:
Neighbor ID     Pri   State           BFD State   Dead Time   Address         Interface
```

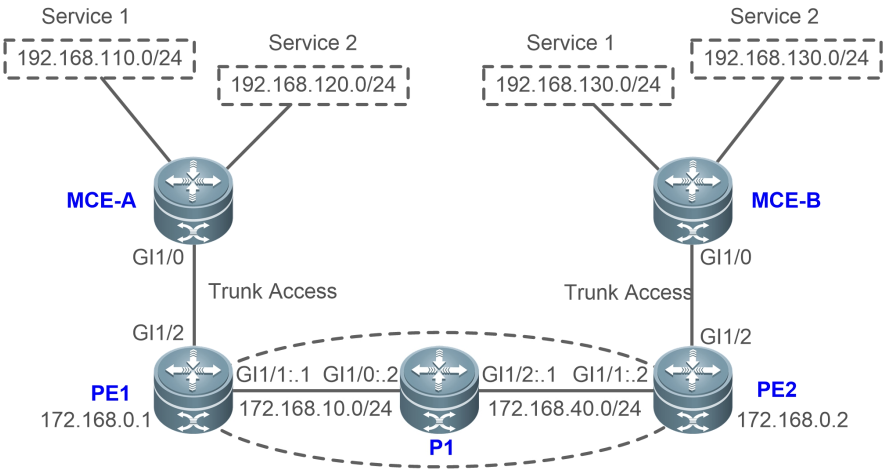
	192.168.0.2 1 Full/ - - 00:00:34 192.168.0.2 SLINK0 PE1# show ip route vrf *VPNA* Routing Table: VPNA Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.10.0/24 is directly connected, GigabitEthernet 1/2 O 192.168.20.0/24 [110/101] via 192.168.1.2, 00:56:23, GigabitEthernet 1/2 O 192.168.30.0/24 [110/2] via 172.168.0.2, 00:00:36 O 192.168.40.0/24 [110/2] via 172.168.0.2, 00:00:36
PE2	PE2# show ip ospf 10 sham-links Sham Link SLINK0 to address 192.168.0.1 is up Area 0.0.0.0 source address 192.168.0.2, Cost: 1 Output interface is GigabitEthernet 1/1 Nexthop address 172.16.10.1 Transmit Delay is 1 sec, State Point-To-Point, Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5 Hello due in 00:00:01 Adjacency state Full PE2# show ip ospf 10 neighbor OSPF process 10, 1 Neighbors, 1 is Full: Neighbor ID Pri State BFD State Dead Time Address Interface 192.168.0.1 1 Full/ - - 00:00:34 192.168.0.1 SLINK0

```
PE2# show ip route vrf VPNA
Routing Table: VPNA

Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
O        192.168.10.0/24 [110/2] via 172.168.0.1, 00:00:36
O        192.168.20.0/24 [110/2] via 172.168.0.1, 00:00:36
C        192.168.30.0/24 is directly connected, GigabitEthernet 1/2
O        192.168.40.0/24 [110/101] via 192.168.30.2, 00:56:23, GigabitEthernet 1/2
```

Configuring Multiple OSPF Instances per MCE

Scenario

Figure 2-50



The sites of a customer operate different services. The customer requires that routes of the same service be exchanged over the MPLS backbone network and different services be isolated from each other.

Configuration Steps

- On MCE-A and MCE-B, configure a trunk link between PE and CE, configure two VRF tables for different services, and bind an interface to each VRF table. Configure the two VRF tables to communicate with the PE over OSPF.
- On PE1 and PE2, configure a loopback interface, configure a trunk link between PE and CE, create two VRF tables

	named VPN1 and VPN2 for different services, and associate the VRF tables with corresponding interfaces. Associate the CE-connected interface with a VRF table. Configure BGP, and establish an MP-IBGP session with the peer PE (PE2 or PE1). Configure the capability of exchanging routes with the CE over OSPF. Configure backbone MPLS signaling. Enable MPLS on the interface connected to the public network. Configure a backbone routing protocol. ● On P1, configure backbone MPLS signaling, enable MPLS on the interface connected to the public network, and configure a backbone routing protocol. This example uses the configuration of the SVI and the 802.1Q sub-interface to describe how to connect PE1 and MCE-A. Some devices may not support the two interfaces. PE1 and MCE-A can be connected by any two links (physical or logical) based on two routing adjacencies between them. You can select a suitable connection method according to the actual condition.
MCE-A	MCE-A# configure terminal MCE-A(config)# interface GigabitEthernet 1/0 MCE-A(config-if-GigabitEthernet 1/0)# switchport mode trunk MCE-A(config-if-GigabitEthernet 1/0)# end MCE-A# configure terminal MCE-A(config)# ip vrf VPN1 MCE-A(config-vrf)# exit MCE-A(config)# vlan 10 MCE-A(config)# interface vlan 10 MCE-A(config-if-Vlan 10)# ip vrf forwarding VPN1 MCE-A(config-if-Vlan 10)# ip address 192.168.10.2 255.255.255.0 MCE-A(config-if-Vlan 10)# exit MCE-A(config)# ip vrf VPN2 MCE-A(config-vrf)# exit MCE-A(config)# vlan 20 MCE-A(config)# interface vlan 20 MCE-A(config-if-Vlan 20)# ip vrf forwarding VPN2 MCE-A(config-if-Vlan 20)# ip address 192.168.20.2 255.255.255.0 MCE-A# configure terminal MCE-A(config)# router ospf 10 vrf VPN1 MCE-A(config-router)# network 192.168.10.0 255.255.255.0 area 0 MCE-A(config-router)# capability vrf-lite

	MCE-A(config-router)# exit MCE-A(config)# router ospf 10 vrf VPN2 MCE-A(config-router)# network 192.168.20.0 255.255.255.0 area 0 MCE-A(config-router)# capability vrf-lite
PE1	PE1# configure terminal PE1(config)# interface loopback 0 PE1(config-if-Loopback 0)# ip address 172.168.0.1 255.255.255.255 PE1(config)# interface GigabitEthernet 1/2 PE1(config-if-GigabitEthernet 1/2)# switchport mode trunk PE1(config-if-GigabitEthernet 1/2)# end PE1# configure terminal PE1(config)# ip vrf VPN1 PE1(config-vrf)# rd 1:100 PE1(config-vrf)# route-target both 1:100 PE1(config-vrf)# end PE1# configure terminal PE1(config)# ip vrf VPN2 PE1(config-vrf)# rd 1:200 PE1(config-vrf)# route-target both 1:200 PE1(config-vrf)# end PE1(config)# vlan 10 PE1(config)# interface vlan 10 PE1(config-if-Vlan 10)# ip vrf forwarding VPN1 PE1(config-if-Vlan 10)# ip address 192.168.10.1 255.255.255.0 PE1(config-if-Vlan 10)# exit PE1(config)# vlan 20 PE1(config)# interface vlan 20 PE1(config-if-Vlan 20)# ip vrf forwarding VPN1 PE1(config-if-Vlan 20)# ip address 192.168.20.1 255.255.255.0 PE1# configure terminal PE1(config)# router bgp 1

```
PE1(config-router)# neighbor 172.168.0.2 remote-as 1
PE1(config-router)# neighbor 172.168.0.2 update-source loopback 0
PE1(config-router)# address-family vpnv4
PE1(config-router-af)# neighbor 172.168.0.2 activate
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# router ospf 10 vrf VPN1
PE1(config-router)# network 192.168.10.0 255.255.255.0 area 0
PE1(config-router)# exit
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf VPNA
PE1(config-router-af)# redistribute ospf 10
PE1(config-router-af)# redistribute connected
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# router ospf 20 vrf VPN2
PE1(config-router)# network 192.168.20.0 255.255.255.0 area 0
PE1(config-router)# redistribute bgp subnets
PE1(config-router)# exit
PE1(config)# router bgp 1
PE1(config-router)# address-family ipv4 vrf VPN2
PE1(config-router-af)# redistribute ospf 20
PE1(config-router-af)# redistribute connected
PE1(config-router-af)# end
PE1# configure terminal
PE1(config)# mpls ip
PE1(config)# mpls router ldp
PE1(config-mpls-router)# ldp router-id interface loopback 0 force
PE1(config-mpls-router)# exit
PE1(config)# interface GigabitEthernet 1/1
PE1(config-if-GigabitEthernet 1/1)# ip address 172.168.10.1 255.255.255.0
PE1(config-if-GigabitEthernet 1/1)# label-switching
```

	PE1(config-if-GigabitEthernet 1/1)# mpls ip PE1(config-if-GigabitEthernet 1/1)# end PE1# configure terminal PE1(config)# router ospf 1 PE1(config-router)# network 172.168.10.0 0.0.0.255 area 0 PE1(config-router)# network 172.168.0.1 0.0.0.0 area 0 PE1(config-router)# end
P1	P1(config)# interface GigabitEthernet 1/0 P1(config-if-GigabitEthernet 1/0)# 172.168.10.2 255.255.255.0 P1(config-if-GigabitEthernet 1/0)# mpls ip P1(config-if-GigabitEthernet 1/0)# label-switching P1(config-if-GigabitEthernet 1/0)# exit P1(config)# interface GigabitEthernet 1/2 P1(config-if-GigabitEthernet 1/2)# 172.168.40.1 255.255.255.0 P1(config-if-GigabitEthernet 1/2)# mpls ip P1(config-if-GigabitEthernet 1/2)# label-switching P1(config-if-GigabitEthernet 1/2)# exit P1(config)# mpls router ldp P1(config-mpls-router)# ldp router-id interface loopback 0 force P1(config-mpls-router)# exit
MCE-B	MCE-B# configure terminal MCE-B(config)# interface GigabitEthernet 1/0 MCE-B(config-GigabitEthernet 1/0)# switchport mode trunk MCE-B(config-GigabitEthernet 1/0)# end MCE-B# configure terminal MCE-B(config)# ip vrf VPN1 MCE-B(config-vrf)# exit MCE-B(config)# vlan 30 MCE-B(config)# interface vlan 30 MCE-B(config-if-Vlan 10)# ip vrf forwarding VPN1 MCE-B(config-if-Vlan 10)# ip address 192.168.30.2 255.255.255.0

	MCE-B(config-if-Vlan 10)# exit MCE-B(config)# ip vrf VPN2 MCE-B(config-vrf)# exit MCE-B(config)# vlan 40 MCE-B(config)# interface vlan 40 MCE-B(config-vlan 20)# ip vrf forwarding VPN2 MCE-B(config-if-Vlan 20)# ip address 192.168.40.2 255.255.255.0 MCE-B# configure terminal MCE-B(config)# router ospf 10 vrf VPN1 MCE-B(config-router)# network 192.168.30.0 255.255.255.0 area 0 MCE-B(config-router)# capability vrf-lite MCE-B(config-router)# exit MCE-B(config)# router ospf 10 vrf VPN2 MCE-B(config-router)# network 192.168.40.0 255.255.255.0 area 0 MCE-B(config-router)# capability vrf-lite
PE2	PE2# configure terminal PE2(config)# interface loopback 0 PE2(config-if-Loopback 0)# ip address 172.168.0.2 255.255.255.255 PE2(config)# interface GigabitEthernet 1/2 PE2(config-if-GigabitEthernet 1/2)# switchport mode trunk PE2(config-if-GigabitEthernet 1/2)# end PE2# configure terminal PE2(config)# ip vrf VPN1 PE2(config-vrf)# rd 1:100 PE2(config-vrf)# route-target both 1:100 PE2(config-vrf)# end PE2# configure terminal PE2(config)# ip vrf VPN2 PE2(config-vrf)# rd 1:200 PE2(config-vrf)# route-target both 1:200 PE2(config-vrf)# end

```
PE2(config)# vlan 30
PE2(config)# interface vlan 30
PE2(config-if-Vlan 10)# ip vrf forwarding VPN1
PE2(config-if-Vlan 10)# ip address 192.168.30.1 255.255.255.0
PE2(config)# vlan 40
PE2(config)# interface vlan 40
PE2(config-if-Vlan 20)# ip vrf forwarding VPN2
PE2(config-if-Vlan 20)# ip address 192.168.40.1 255.255.255.0
PE2# configure terminal
PE2(config)# router bgp 1
PE2(config-router)# neighbor 172.168.0.1 remote-as 1
PE2(config-router)# neighbor 172.168.0.1 update-source loopback 0
PE2(config-router)# address-family vpnv4
PE2(config-router-af)# neighbor 172.168.0.1 activate
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# router ospf 10 vrf VPN1
PE2(config-router)# network 192.168.30.0 255.255.255.0 area 0
PE2(config-router)# redistribute bgp subnets
PE2(config-router)# exit
PE2(config)# router bgp 1
PE2(config-router)# address-family ipv4 vrf VPN1
PE2(config-router-af)# redistribute ospf 10
PE2(config-router-af)# redistribute connected
PE2(config-router-af)# end
PE2# configure terminal
PE2(config)# router ospf 20 vrf VPN2
PE2(config-router)# network 192.168.40.0 255.255.255.0 area 0
PE2(config-router)# redistribute bgp subnets
PE2(config-router)# exit
PE2(config)# router bgp 1
PE2(config-router)# address-family ipv4 vrf VPN2
```

	PE2(config-router-af)# redistribute ospf 20 PE2(config-router-af)# redistribute connected PE2(config-router-af)# end PE2# configure terminal PE2(config)# mpls ip PE2(config)# mpls router ldp PE2(config-mpls-router)# ldp router-id interface loopback 0 force PE2(config-mpls-router)# exit PE2(config)# interface GigabitEthernet *1/1* PE2(config-if-GigabitEthernet 1/1)# ip address 172.168.40.2 255.255.255.0 PE2(config-if-GigabitEthernet 1/1)# label-switching PE2(config-if-GigabitEthernet 1/1)# mpls ip PE2(config-if-GigabitEthernet 1/1)# end PE2# configure terminal PE2(config)# router ospf 1 PE2(config-router)# network 172.168.40.0 0.0.0.255 area 0 PE2(config-router)# network 172.168.0.2 0.0.0.0 area 0 PE2(config-router)# end
Verification	● Run the **show ip route vrf** command to display IP entries.
MCE-A	MCE-A# show ip route vrf VPN1 Routing Table: VPN1 Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.10.0/24 is directly connected, VLAN 10

	O 192.168.110.0/24 [110/101] via 192.168.21.2, 00:56:23, GigabitEthernet 1/1 O IA 192.168.130.0/24 [110/2] via 192.168.10.1, 00:00:36, VLAN 10 MCE-A# show ip route vrf *VPN2* Routing Table: VPN2 Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.20.0/24 is directly connected, VLAN 20 O 192.168.120.0/24 [110/101] via 192.168.22.2, 00:56:23, GigabitEthernet 1/2 O IA 192.168.140.0/24 [110/2] via 192.168.20.1, 00:00:36, VLAN 20
MCE-B	MCE-B# show ip route vrf VPN1 Routing Table: VPN1 Codes: C - connected, S - static, R - RIP, B - BGP O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2 i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2 ia - IS-IS inter area, * - candidate default Gateway of last resort is no set C 192.168.30.0/24 is directly connected, VLAN 30 O 192.168.130.0/24 [110/101] via 192.168.23.2, 00:56:23, GigabitEthernet 1/1 O IA 192.168.110.0/24 [110/2] via 192.168.30.1, 00:00:36, VLAN 30

```
MCE-B# show ip route vrf VPN2
Routing Table: VPN2

Codes:   C - connected, S - static, R - RIP, B - BGP
         O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C        192.168.40.0/24 is directly connected, VLAN 40
O        192.168.140.0/24 [110/101] via 192.168.24.2, 00:56:23, GigabitEthernet 1/2
O IA     192.168.140.0/24 [110/2] via 192.168.40.1, 00:00:36, VLAN 40
```

Common Errors

An LDP session fails to be established because the router ID does not consist of 32 bits.

2.5 Monitoring

Clearing

Description	Command
Resets the VPNv4 unicast address family of BGP.	**clear bgp vpnv4 unicast** { ***** | *as-number* | *peer-address* } [**soft**] [**in** | **out**]
Clears VPNv4 route flapping information and cancels VPNv4 route suppression.	**clear bgp vpnv4 unicast dampening** [*ip-address* [*mask*]]
Resets all EBGP connections under the VPNv4 address family.	**clear bgp vpnv4 unicast external** [**soft**] [**in** | **out**]
Clears route flapping statistics under the VPNv4 address family.	**clear bgp vpnv4 unicast flap-statistics** *[ip-address* [*mask*]]
Resets the VPNv4 address family for all members of the BGP peer group.	**clear bgp vpnv4 unicast peer-group** *peer-group-name* [**soft**] [**in** | **out**]
Resets the BGP session under a VRF table.	**clear ip bgp vrf** *vrf-name* { * | *address* | *as-num* } [**soft**] [**in** | **out**]

Displaying

Description	Command
Displays the IPv4 routes with an MPLS label that are learned and sent over BGP.	**show bgp ipv4 unicast labels**
Displays the VPNv4 routes or the neighbor information that are learned and sent over BGP.	**show bgp vpnv4 unicast all** [*network* | **neighbor** [*peer-address*] | **summary** | **label**]
Displays the VPNv4 routes or the neighbor information that are learned and sent over BGP under the specified VRF table.	**show bgp vpnv4 unicast vrf** *vrf_name* [*network* | **summary** | **label**]
Displays the VPNv4 routes or the neighbor information that are learned and sent over BGP under the specified RD.	**show bgp vpnv4 unicast rd** *rd_value* [*network* | **summary** | **label**]
Displays the configurations of an extcommunity list.	**show ip extcommunity-list** [*extcommunity-list-num* | *extcommunity-list-name*]
Displays the OSPF sham link information.	**show ip ospf** [*process-id*] **sham-links** [**area** *area-id*]
Displays the configured single-protocol VRF information.	**show ip vrf** [**brief** | **detail** | **interfaces**] [*vrf-name*]
Displays brief information about VRF tables (single- and multi-protocol VRF tables).	**show vrf brief** [*vrf-name*]
Displays brief information about the IPv4 address family for VRF tables (single- and multi-protocol VRF tables).	**show vrf ipv4** [*vrf-name*]
Displays brief information about the IPv6 address family for multi-protocol VRF tables.	**show vrf ipv6** [*vrf-name*]
Displays detailed information about VRF tables (single- and multi-protocol VRF tables).	**show vrf detail** [*vrf-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Enables BGP MPLS debugging.	**debug ip bgp mpls**
Enables MPLS entry debugging.	**debug mpls**
Enables MPLS message debugging.	**debug mpls msg** [**send** | **recv**]

3 Configuring MPLS GR

The routers and router icons described in this chapter refer to general routers as well as Layer-3 switches running routing protocols. We use "router series" or "switch series" to refer to specific devices.

3.1 Overview

The Internet Engineering Task Force (IETF) extends the Multiprotocol Label Switching (MPLS) signaling protocol – Label Distribution Protocol (LDP). In this way, when LDP is restarted, a device can notify its neighbor to retain MPLS forwarding entries for a period and mark these entries as "Old". Peripheral devices assist the device in information synchronization so that the device can return to the condition before the LDP restart within a short time. If the packet forwarding path is not changed during the LDP restart, the system can forward data continuously, thereby providing reliable MPLS services.

Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS) use basic MPLS networks to establish public tunnels and use extended LDP to distribute virtual circuit (VC) labels and establish VCs. Therefore, for continuous VPWS and VPLS forwarding, extended LDP Graceful Restart (GR) must be implemented. Despite different MPLS forwarding entries, extended LDP GR and basic LDP GR are implemented in the same way. Therefore, extended LDP GR and basic LDP GR can be processed uniformly.

Layer-3 Virtual Private Network (L3VPN) GR enables continuous L3VPN forwarding. When the control plane of a device fails, VPN forwarding continues, which protects VPNs.

Protocols and Standards

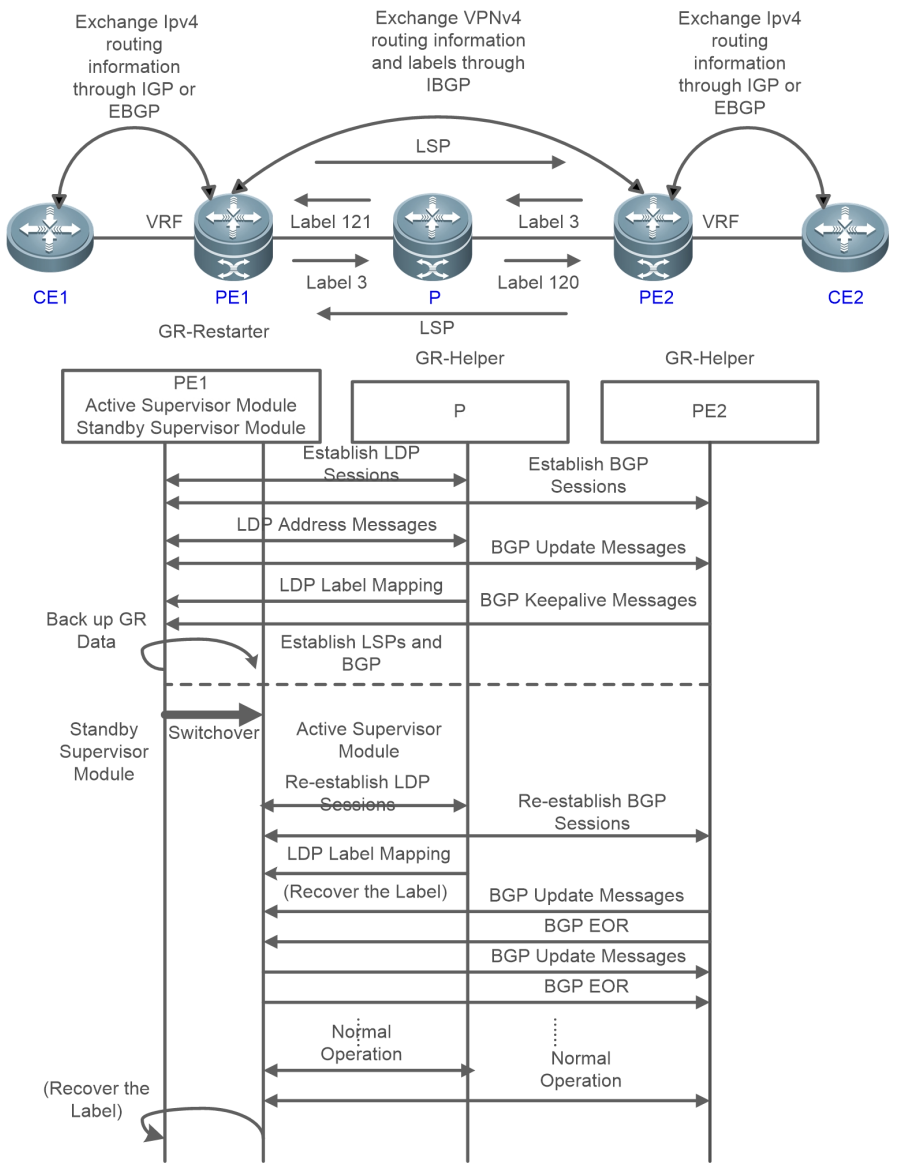
- RFC3036: LDP Specification
- RFC3037: LDP Applicability
- RFC3215: LDP State Machine
- RFC3478: Graceful Restart Mechanism for Label Distribution Protocol
- RFC3479: Fault Tolerance for the Label Distribution Protocol (LDP)
- RFC 4724: Graceful Restart Mechanism for BGP

3.2 Applications

Application	Description
L3VPN GR	Deploying GR on a device that enables MPLS L3VPN (for example, a PE device) ensures that users can forward their services in case of any fault, thereby improving reliability.

3.2.1 L3VPN GR

Figure 3- 1



Remarks	CE devices represent a customer network and run the IGP or External Border Gateway Protocol (EBGP). PE and P devices form a provider network and run the IGP. PE1, P and PE2 establish public network tunnels through LDP and establish LSPs. PE1 and PE2 establish private network tunnels through extended Internal Border Gateway Protocol(IBGP). IGP, BGP and LDP have the GR capability. PE devices are GR-Capable routers, and the P device is a GR-Aware router.

The following describes the GR process on PE1. The active and standby supervisor modules are switched over on PE1. PE1 serves as a GR Restarter, while P and PE2 serve as GR Helpers. Active and standby supervisor modules are switched over in three phases:

Before the Active/Standby Supervisor Module Switchover

PE1 negotiates with CE1 about IGP or External BGP (EBGP) GR, negotiates with P about IGP and LDP GR, and negotiates with PE2 about Internal BGP (IBGP) GR.

PE1 sends P an initialization message with FT session TLV options to establish a GR-capable LDP session. Then the devices exchange addresses and label mapping messages. In this way, GR-capable LSPs are established for data forwarding.

PE1 sends PE2 a GR-capable <AFI=IPv4, SAFI=Unicast> and <AFI=IPv4, SAFI=VPNv4> Open message to establish a GR-capable IBGP session.

When the active supervisor module works normally, you need to back up the GR data to the standby supervisor module so that after active and standby supervisor mudules are switched over, PE1 can read the data during the GR protocol restart and use it for GR.

During the Active/Standby Supervisor Module Switchover

After the GR data is backed up to the standby supervisor module, PE1 switches over active and standby supervisor modules.

P detects TCP sessions in Down state and marks the corresponding LSPs as "Old". The device enables the forwarding entry aging timer. Packet forwarding will be continued before a timeout.

When detecting disconnected TCP, PE2 immediately mark the routes learned from PE1 as "Old" and enables a restart timer for PE1. If PE2 does not receive the Open message before a timeout, it will delete the "Old" flag; if PE2 receives the Open message, it will delete the restart timer. PE1 and PE2 continue traffic forwarding through original routes.

After the Active/Standby Supervisor Module Switchover

On PE1, the standby supervisor module becomes an active supervisor module and the active supervisor module becomes a standby supervisor module. PE1 checks the GR data backed up to the new active supervisor module for forwarding entries before the protocol restart. Then devices initialize their CLIs and enable GR. During IGP, BGP, and LDP GR, the devices notify all IGP and LDP neighbors to re-establish connections.

- IGP convergence

PE1 sends P a GR-capable initialization message, and re-establishes the session upon receipt of a response to obtain topology and routing information. Then PE1 re-calculates the routing table and deletes "Old" routes. In this way, IGP converges.

- BGP processing

PE1 and CE1 exchange routing information. Then PE1 updates the routing table and forwarding entries based on route forwarding information and replaces invalid routing information. In this way, BGP converges.

PE1 and PE2 establish a new BGP session in between. First, PE1 sends PE2 an Open message with GR options. Then PE1 receives an Update message from PE2 and processes the message. The Update message includes IP address prefixes. PE1 selects BGP routes when receiving an End-of-RIB (EOR) flag. PE1 sends PE2 the Update message with IP address prefixes and then sends PE2 the EOR flag. PE2 selects routes upon receipt of the EOR flag. In this way, the network converges.

- LDP processing

PE1 sends P a GR-capable initialization message. A GR-capable LDP session is established when P receives the message. The devices re-exchange addresses and label mapping messages and delete or retain the "Old" flag for MPLS forwarding entries. After GR, the devices delete "Old" MPLS forwarding entries.

IGP GR, BGP GR, and LDP GR are processed in no particular order. Unicast routes converge first and then are advertised to LDP.

Before all protocols finish GR, PE1 only updates the RIB of the active supervisor module.

After all protocols finish GR, PE1 updates the FIB of the interface module.

Deployment

The prerequisite for implementing L3VPN GR isthatthecontinuous forwarding of routing protocols must be supported and the hardware must support 1 + 1 redundant backup formanagement modules.

continuous forwarding of routing protocols means that a unicast routing protocol must support GR; that is, the device must support Open Shortest Path First (OSPF) GR, Intermediate System to Intermediate System (IS-IS) GR, or BGP GR.

- PE1, P and PE2 create an L3VPN environment where L3VPN is configured to provide L3VPN services for users.
- IGP GR is configured for CE1, PE1, P, PE2 and CE2 to continuously forward routing protocols.
- BGP GR is configured for PE1 and PE2 to continuously forward private routes.
- LDP GR is configured for PE1, P and PE2 to continuously forward public MPLS.

3.3 Features

Basic Concepts

GR Capability-related Router

In terms of the GR capability, GR-Capable routers, GR-Aware routers, and GR-Unaware routers are available.

- GR-Capable router

A GR-Capable router has the GR capability. It is usually equipped with dual supervisor modules, which enables 1+1 redundancy. During the active/standby supervisor module switchover, the router notifies neighbors to retain its forwarding entries. After the switchover, the router re-establishes the routing table without route flapping, and the packet forwarding path is not changed. In this way, continuous forwarding is enabled.

- GR-Aware router

A GR-Aware router is aware of GR. It may not have dual supervisor modules, but it is aware of GR on neighbors and can help neighbors finish GR.

- GR-Unaware router

A GR-Unaware router is unaware of GR on neighbors and cannot help neighbors finish GR because software does not have a GR feature or the GR feature is disabled.

GR Role-related Router

During the router restart, in terms of the role, GR Restarters and GR Helpers are available.

- GR Restarter

A GR Restarter restarts because of the administrator operation or failures. It must be a GR-Capable router.

- GR Helper

A GR Helper is the neighbor of a GR Restarter. It is at least a GR-Aware router.

Overview

Feature	Description
LDP GR	When the MPLS device is faulty, the function maintains the neighbor relationship and sessions between routers, and recovers the session connection and label information subsequently to prevent common MPLS and L2VPN forwarding from being interrupted.

3.3.1 LDP GR

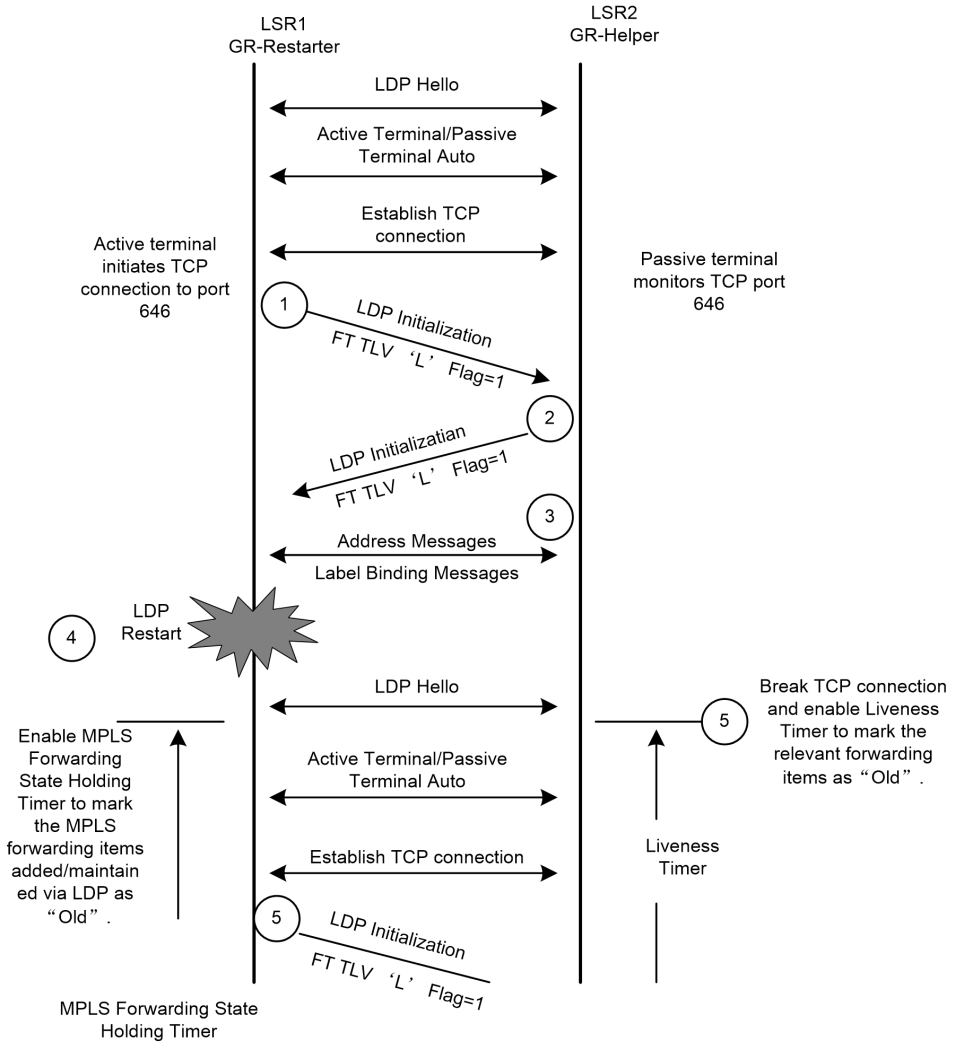
By using LDP GR together with IGP GR, you can achieve uninterruptedcommon MPLS or L2VPN forwarding. If further using BGP GR, you can achieve uninterruptedL3VPNforwarding.

To establish a GR-capable LDP session, both ends must support and enable LDP GR. If one end does not support GR, only ordinary LDP sessions can be established. If the active side supports and enables LDP GR, it will add FT Session TLV to an initialization message.

If the passive side receives an initialization message with FT Session TLV, it can decide whether to add FT Session TLV to an initialization message. If it supports and enables LDP GR, it will add FT Session TLV to an initialization message and thereby a GR-capable LDP session is established; otherwise, it will not add FT Session TLV to an initialization message and thereby an ordinary LDP session is established. If the passive side receives an initialization message without FT Session TLV, an ordinary LDP session will be established whether the passive side adds FT Session TLV to an initialization message. The following figure shows the process of LDP session establishment between two GR-capable LSRs.

Working Principle

Figure 3- 2



Related Configuration

Enabling LDP GR

By default, LDP GR is enabled on devices enabling LDP.

In LDP mode, by running the **graceful-restart** command, you can enable or disable LDP GR.

The GR capability of LDP can take effect only when LDP sessions are restarted.

Configuring LDP GR

By default, the Time To Live (TTL) of an LDP neighbor is 120 seconds; the time for re-connecting LDP sessions is 300 seconds; the LDP recovery time is 120 seconds.

- Run the **graceful-restart timer neighbor-liveness** command to configure the TTL of an LDP neighbor. During this period of time, the Helper waits for the Restarter to restart and then they re-establish the session. If it takes a long time for the Restarter to restart, you can extend the TTL.

- Run the **graceful-restart timer reconnect** command to configure the LDP session reconnection time. During this period of time, the Restarter holds the MPLS forwarding entries (corresponding to the MPLS forwarding state holding timer shown in the figure). If it takes a long time to recover a large number of MPLS forwarding entries, you can extend the time.
- Run the graceful-restart timer recovery command to configure the LDP session recovery time. During this period of time, the Helper assists the Restarter in recovering MPLS forwarding entries after the session is re-established. If it takes a long time to recover a large number of MPLS forwarding entries, you can extend the time.

3.4 Configuration

Configuration	Description and Command	
LDP GR	(Mandatory) It is used to configure LDP GR.	
	graceful-restart	Enables the GR capability of LDP.
	(Optional) It is used to configure the LDP GR timer.	
	graceful-restart timer reconnect	Configures the time for re-connecting LDP sessions.
	graceful-restarttimer neighbor-liveness	Configures the TTL of an LDP neighbor.
	graceful-restart timer recovery	Configures the time for recovering LDP sessions.

3.4.1 LDP GR

Configuration Effect

After MPLS LDP GR is configured, when MPLS devices fail, the neighbor relationship and sessions between routers can be maintained, and session connection and labels can be recovered.

Notes

The routing interface for configuring IP addresses must be a Layer-3 routing interface, or configuration will fail.

Topological changes during GR (for example, changes in the IP address or link of the Link-State Request (LSR) ID used before GR is deleted) may disable GR and interrupt service traffic.

You need to finish the following tasks before configuring MPLS LDP GR:

- Configuring IGP GR
- Configuring MPLS LDP sessions

Before configuring MPLS LDP GR, you need to configure the following:

- The LDP session reconnection time
- The TTL of an LDP neighbor
- The LDP session recovery time

Configuration Steps

Enabling LDP GR

- Mandatory.
- Run the **no graceful-restart** command to disable the GR capability of LDP.
- Enable GR by default to support process restart. Except otherwise required, do not disable LDP GR.

Enabling LDP GR will not affect LDP sessions; that is, LDP sessions will not be restarted. LDP GR takes effect only when LDP sessions ae restarted.

Configuring LDP GR

- Optional.
- Run the **graceful-restart timer neighbor-liveness** command to configure the TTL of an LDP neighbor. During this period of time, the Helper waits for the Restarter to restart and then they re-establish the session. If it takes a long time for the Restarter to restart, you can extend the TTL.
- Run the **graceful-restart timer reconnect** command to configure the LDP session reconnection time. During this period of time, the Restarter holds the MPLS forwarding entries (corresponding to the MPLS forwarding state holding timer shown in the figure). If it takes a long time to recover a large number of MPLS forwarding entries, you can extend the time.
- Run the graceful-restart timer recovery command to configure the LDP session recovery time. During this period of time, the Helper assists the Restarter in recovering MPLS forwarding entries after the session is re-established. If it takes a long time to recover a large number of MPLS forwarding entries, you can extend the time.

Verification

- After GR is configured, the sessions previously established will not be affected, and the function will take effect only when new sessions are established.
- You can run the **show mpls ldp graceful-restart** command to display GR-capable sessions.
- You can run the **show mpls ldp neighbors** command to check whether a session has the GR capability.

Related Commands

Enabling LDP GR

Command	**graceful-restart**
Parameter Description	N/A
Command Mode	This configuration is performed in config-mpls-router configuration mode.
Usage Guide	The following describes how to use this command: If LDP GR is enabled on a dual-engine device, when the active supervisor module fails (or active and standby supervisor modules are switched over manually), traffic forwarding will not be interrupted, and the MPLS forwarding state will not be changed after the LDP restart. By default, GR is enabled whether a device serves as a GR Restarter or a GR Helper.

	The changed configuration of LDP GR takes effect only when LDP sessions are restarted.

Configuring the LDP Session Reconnection Time

Command	**graceful-restart timer reconnect***seconds*
Parameter Description	*seconds:* Configures the LDP session reconnection time, which ranges from 30 to 3,600 seconds.
Command Mode	This configuration is performed in config-mpls-router configuration mode.
Usage Guide	● During GR, both GR Restarter and GR Helper use the LDP session reconnection time. ● A GR Restarter uses the LDP session reconnection time to retain "Old" entries. ● A GR Helper enables the liveness timer based on its neighbor-liveness time or received reconnection time, whichever is shorter, and retains the "Old" entries before a timeout. ● If it takes a long time to reconnect to a neighbor because of a large number of services, you need to adjust the reconnection time and neighbor-liveness time. In a typical L3VPN scenario, the time increases by 1 minute every 100 VPNs. ● The LDP session reconnection time takes effect only when LDP sessions are restarted.

Configuring the TTL of an LDP Neighbor

Command	**graceful-restarttimer neighbor-liveness***seconds*
Parameter Description	*seconds:* Configures the TTL of an LDP neighbor, which ranges from 5 to 3,600 seconds.
Command Mode	This configuration is performed in config-mpls-router configuration mode.
Usage Guide	● Only a GR Helper uses the TTL. ● A GR Helper enables the liveness timer based on its neighbor-liveness time or received reconnection time, whichever is shorter, and retains the "Old" entries before a timeout. ● The TTL of an LDP neighbor takes effect only when LDP sessions are restarted.

Configuring the LDP Session Recovery Time

Command	**graceful-restart timer recovery***seconds*
Parameter Description	*seconds:* Configures the LDP session recovery time, which ranges from 15 to 600 seconds.
Command Mode	This configuration is performed in config-mpls-router configuration mode.
Usage Guide	● Only a GR Helper uses the LDP session recovery time. ● A GR Helper enables the recovery timer based on its recovery time or received recovery time, whichever is shorter, and retains the "Old" entries before a timeout. ● The LDP session recovery time takes effect only when LDP sessions are restarted.

Displaying Configuration Effect

Command	**show mpls ldp graceful-restart** [**all** | **vrf***vrf-name*]

Parameter Description	**all:** Displays all VPN routing and forwarding (VRF, including global VRF) LDP GR sessions and session parameters. **vrf***vrf-name:* Displays specified VRF LDP GR sessions and session parameters.
Command Mode	This configuration is performed in privileged user mode.
Usage Guide	Displays LDP GR sessions and session parameters.

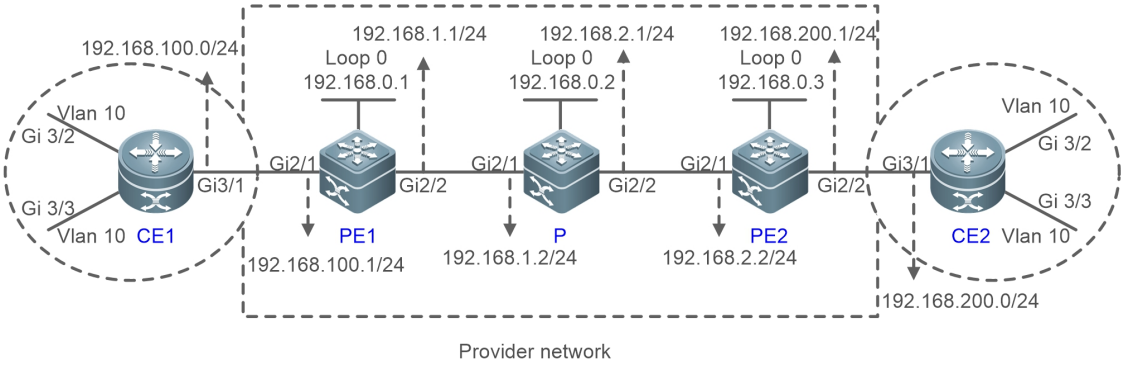
Common Errors

- The basic MPLS or L3VPN environment is incomplete, which results in forwarding failure.
- The Restarter is configured with GR, while the neighboring Helper is not.
- IGP GR or BGP GR is not configured.
- There are many routes but the timer time is not long enough, which disables GR before all entries are recovered.
- Topology or configuration changes during GR, which interrupts service traffic.

3.4.1.1 LDP GR Configuration Example

Scenario

Figure 3-3



- PEs and P form an MPLS network.
- PEs and P support LDP and have the GR capability.
- This section describes LDP GR configuration on PE1 and P. PE1 is a GR-Capable router and serves as a GR Restarter. P is a GR-Aware router and serves as a GR Helper.

Configuration Steps

- Configure the interface IP address, configure OSPF, and enable OSPF GR.
- Enable global MPLS forwarding, and enable MPLS labeled packet forwarding and LDP on an interface.
- Configure LDP to enable MPLS traffic forwarding.
- Enable LDP GR.
- Configure LDP GR.
- Restart LDP sessions.

PE1

```
FS#configure terminal
Enter configuration commands, one per line.    End with CNTL/Z.
```

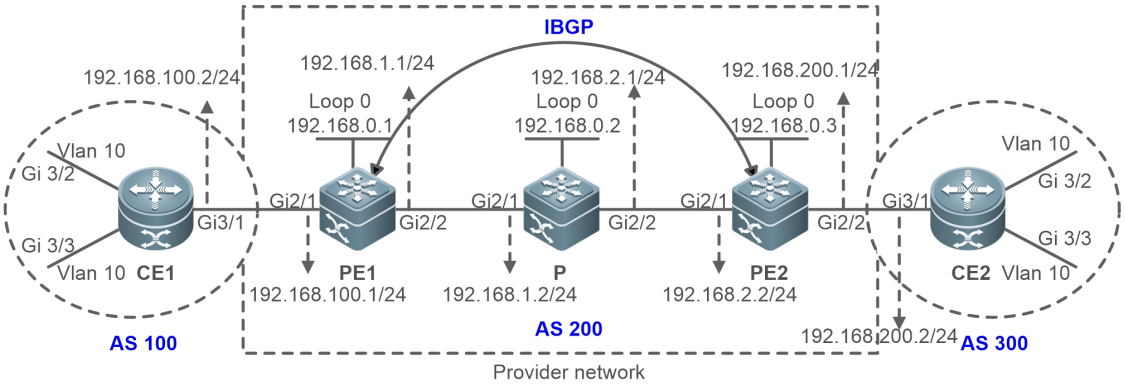
	FS(config)#interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)#ip address 192.168.100.1 255.255.255.0 FS(config-if-GigabitEthernet 2/1)#exit FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#ip address 192.168.1.1 255.255.255.0 FS(config-if-GigabitEthernet 2/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 192.168.0.1 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network 192.168.100.00.0.0.255 area 0 Router(config-router)#network 192.168.1.00.0.0.255 area 0 Router(config-router)#network 192.168.0.10.0.0.0 area 0 Router(config-router)# graceful-restart Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#label-switching FS(config-if-GigabitEthernet 2/2)#mpls ip Router(config-if-GigabitEthernet 2/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#graceful-restart FS(config-mpls-router)#graceful-restart timer reconnect 300 FS(config-mpls-router)#graceful-restart timer neighbor-liveness 120 FS(config-mpls-router)#graceful-restart timer recovery 120 FS(config-mpls-router)#end FS#clear mpls ldp neighbor all
P	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 2/1

	FS(config-if-GigabitEthernet 2/1)#ip address 192.168.1.2 255.255.255.0 FS(config-if-GigabitEthernet 2/1)#exit FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#ip address 192.168.2.1 255.255.255.0 FS(config-if-GigabitEthernet 2/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 192.168.0.2 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network 192.168.1.00.0.0.255 area 0 Router(config-router)#network 192.168.2.00.0.0.255 area 0 Router(config-router)#network 192.168.0.2 255.255.255.255 area 0 Router(config-router)# graceful-restart Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)#label-switching FS(config-if-GigabitEthernet 2/1)#mpls ip Router(config-if-GigabitEthernet 2/1)#exit FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#label-switching FS(config-if-GigabitEthernet 2/2)#mpls ip Router(config-if-GigabitEthernet 2/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#graceful-restart FS(config-mpls-router)#graceful-restart timer reconnect 300 FS(config-mpls-router)#graceful-restart timer neighbor-liveness 120 FS(config-mpls-router)#graceful-restart timer recovery 120 FS(config-mpls-router)#end FS#clear mpls ldp neighbor all
Verification	● Run the **show mpls ldp graceful-restart** command to display GR-capable sessions.

	● Run the **show mpls ldp neighbors** command to check whether a session has the GR capability.
PE1	FS#show mpls ldp graceful-restart Default VRF: LDP Graceful Restart is enabled Neighbor Liveness Timer: 120 seconds Max Recovery Time: 120 seconds Forwarding State Holding Time: 300 seconds Down Neighbor Database (0 records): Graceful Restart-enabled Sessions: Peer LDP Ident: 192.168.0.2:0, State: estab FS#show mpls ldp neighbors Default VRF: Peer LDP Ident: 192.168.0.2:0; Local LDP Ident: 192.168.0.1:0 TCP connection: 192.168.0.2.15532 - 192.168.0.1.646 State: OPERATIONAL; Msgs sent/recv: 23/27; UNSOLICITED Up time: 00:04:12 Graceful Restart enabled; Peer reconnect time (msecs): 0
P	FS#show mpls ldp graceful-restart Default VRF: LDP Graceful Restart is enabled Neighbor Liveness Timer: 120 seconds Max Recovery Time: 120 seconds Forwarding State Holding Time: 300 seconds Down Neighbor Database (0 records): Graceful Restart-enabled Sessions: Peer LDP Ident: 192.168.0.1:0, State: estab FS#show mpls ldp neighbors Default VRF: Peer LDP Ident: 192.168.0.1:0; Local LDP Ident: 192.168.0.2:0 TCP connection: 192.168.0.1.646 - 192.168.0.2.15532 State: OPERATIONAL; Msgs sent/recv: 27/23; UNSOLICITED

	Up time: 00:04:12 Graceful Restart enabled; Peer reconnect time (msecs): 0

3.4.1.2 L3VPN GR Configuration Example

Scenario **Figure 3-4**	 ● CEs represent a customer network and run IGP or EBGP. ● PEs and P form a provider network and run IGP. ● PE1, P, and PE2 establish public tunnels through LDP and establish LSPs. ● PE1 and PE2 establish private tunnels through IBGP. ● IGP, BGP, and LDP have the GR capability. ● PEs are GR-Capable routers, and P is a GR-Aware router.
Configuration Steps	● Configure VRF. ● Configure the interface IP address, configure OSPF, and enable OSPF GR. ● Enable global MPLS forwarding, and enable MPLS labeled packet forwarding and LDP on an interface. ● Configure LDP to enable MPLS traffic forwarding. ● Enable and configure LDP GR. ● Configure L3VPN. ● Enable BGP GR. ● Restart LDP sessions.
PE1	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#ip vrf 10 FS(config-vrf)#rd 1:100 FS(config-vrf)#route-target both 1:100 FS(config-vrf)#exit

```
FS(config)#interface gigabitEthernet 2/1
FS(config-if-GigabitEthernet 2/1)#ip vrf forwarding 10
FS(config-if-GigabitEthernet 2/1)#ip address 192.168.100.1 255.255.255.0
FS(config-if-GigabitEthernet 2/1)#exit
FS(config)#interface gigabitEthernet 2/2
FS(config-if-GigabitEthernet 2/2)#ip address 192.168.1.1 255.255.255.0
FS(config-if-GigabitEthernet 2/2)#exit
FS(config)#interface loopback 0
FS(config-Loopback 0)#ip address 192.168.0.1 255.255.255.255
FS(config-Loopback 0)#exit
FS(config)#router ospf 10
FS(config-router)#network 192.168.0.10.0.0.0 area 0
FS(config-router)#network 192.168.1.00.0.0.255 area 0
FS(config-router)#graceful-restart
FS(config-router)#exit
FS(config)#mpls ip
FS(config)#interface gigabitEthernet 2/2
FS(config-if-GigabitEthernet 2/2)#label-switching
FS(config-if-GigabitEthernet 2/2)#mpls ip
Router(config-if-GigabitEthernet 2/2)#exit
FS(config)#mpls router ldp
FS(config-mpls-router)#ldp router-id interface loopback 0 force
FS(config-mpls-router)#graceful-restart
FS(config-mpls-router)#graceful-restart timer reconnect 300
FS(config-mpls-router)#graceful-restart timer neighbor-liveness 120
FS(config-mpls-router)#graceful-restart timer recovery 120
FS(config-mpls-router)#exit
FS(config)#router bgp 200
FS(config-router)#address-family ipv4 vrf 10
FS(config-router-af)#neighbor 192.168.100.2 remote-as 100
FS(config-router-af)#neighbor 192.168.100.2 update-source GigabitEthernet 2/1
FS(config-router-af)#neighbor 192.168.200.2 activate
```

	FS(config-router-af)#exit-address-family FS(config-router)#address-family ipv4 FS(config-router-af)#neighbor 192.168.0.3 remote-as 200 FS(config-router-af)#neighbor 192.168.0.3 update-source loopback 0 FS(config-router-af)#neighbor 192.168.0.3 activate FS(config-router-af)#exit-address-family FS(config-router)#address-family vpnv4 unicast FS(config-router-af)#neighbor 192.168.0.3 activate FS(config-router-af)#exit-address-family FS(config-router)#bgp graceful-restart FS(config-router)#end FS#clear mpls ldp neighbor all
P	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)#ip address 192.168.1.2 255.255.255.0 FS(config-if-GigabitEthernet 2/1)#exit FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#ip address 192.168.2.1 255.255.255.0 FS(config-if-GigabitEthernet 2/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 192.168.0.2 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 10 FS(config-router)#network 192.168.1.0 0.0.0.255 area 0 FS(config-router)#network 192.168.2.0 0.0.0.255 area 0 FS(config-router)#network 192.168.0.2 0.0.0.0 area 0 FS(config-router)#graceful-restart FS(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 2/1

	FS(config-if-GigabitEthernet 2/1)#label-switching FS(config-if-GigabitEthernet 2/1)#mpls ip Router(config-if-GigabitEthernet 2/1)#exit FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#label-switching FS(config-if-GigabitEthernet 2/2)#mpls ip Router(config-if-GigabitEthernet 2/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#graceful-restart FS(config-mpls-router)#graceful-restart timer reconnect 300 FS(config-mpls-router)#graceful-restart timer neighbor-liveness 120 FS(config-mpls-router)#graceful-restart timer recovery 120 FS(config-mpls-router)#end FS#clear mpls ldp neighbor all
PE2	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#ip vrf 10 FS(config-vrf)#rd 1:100 FS(config-vrf)#route-target both 1:100 FS(config-vrf)#exit FS(config)#interface gigabitEthernet 2/2 FS(config-if-GigabitEthernet 2/2)#ip vrf forwarding 10 FS(config-if-GigabitEthernet 2/2)#ip address 192.168.200.1 255.255.255.0 FS(config-if-GigabitEthernet 2/2)#exit FS(config)#interface gigabitEthernet 2/1 FS(config-if-GigabitEthernet 2/1)#ip address 192.168.2.2 255.255.255.0 FS(config-if-GigabitEthernet 2/1)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 192.168.0.3 255.255.255.255 FS(config-Loopback 0)#exit

```
FS(config)#router ospf 10
FS(config-router)#network 192.168.0.30.0.0.0 area 0
FS(config-router)#network 192.168.2.00.0.0.255 area 0
FS(config-router)#graceful-restart
FS(config-router)#exit
FS(config)#mpls ip
FS(config)#interface gigabitEthernet 2/1
FS(config-if-GigabitEthernet 2/1)#label-switching
FS(config-if-GigabitEthernet 2/1)#mpls ip
Router(config-if-GigabitEthernet 2/1)#exit
FS(config)#mpls router ldp
FS(config-mpls-router)#ldp router-id interface loopback 0 force
FS(config-mpls-router)#graceful-restart
FS(config-mpls-router)#graceful-restart timer reconnect 300
FS(config-mpls-router)#graceful-restart timer neighbor-liveness 120
FS(config-mpls-router)#graceful-restart timer recovery 120
FS(config-mpls-router)#exit
FS(config)#router bgp 200
FS(config-router)#address-family ipv4 vrf 10
FS(config-router-af)#neighbor 192.168.200.2 remote-as 300
FS(config-router-af)#neighbor 192.168.200.2 update-source GigabitEthernet 2/2
FS(config-router-af)#neighbor 192.168.100.2 activate
FS(config-router-af)#exit-address-family
FS(config-router)#exit
FS(config-router)#address-family ipv4
FS(config-router-af)#neighbor 192.168.0.1 remote-as 200
FS(config-router-af)#neighbor 192.168.0.1 update-source loopback 0
FS(config-router-af)#neighbor 192.168.0.1 activate
FS(config-router-af)#exit-address-family
FS(config-router)#address-family vpnv4 unicast
FS(config-router-af)#neighbor 192.168.0.1 activate
FS(config-router-af)#exit-address-family
```

	FS(config-router)#bgp graceful-restart FS(config-router)#end FS#clear mpls ldp neighbor all
Verification	● Run the **show mpls ldp graceful-restart** command to display GR-capable sessions. ● Run the **show bgp vpnv4 unicast all neighbor** command to display BGP neighbors.
PE1	FS#show mpls ldp graceful-restart Default VRF: LDP Graceful Restart is enabled Neighbor Liveness Timer: 120 seconds Max Recovery Time: 120 seconds ForwardingState Holding Time: 300 seconds Down Neighbor Database (0 records): Graceful Restart-enabled Sessions: Peer LDP Ident: 192.168.0.2:0, State: estab FS#show bgp vpnv4 unicast all neighbor BGP neighbor is 192.168.0.3, remote AS 200, internal link BGP version 4, remote router ID 192.168.0.3 BGP state = Established, up for 02:49:47 Last read 00:00:47, hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received(new) Address family VPNv4 Unicast: advertised and received Graceful Restart Capabilty: advertised and received Remote Restart timer is 120 seconds Address families preserved by peer: VPNv4 Unicast
P	FS#show mpls ldp graceful-restart Default VRF: LDP Graceful Restart is enabled Neighbor Liveness Timer: 120 seconds Max Recovery Time: 120 seconds

	ForwardingState Holding Time: 300 seconds Down Neighbor Database (0 records): Graceful Restart-enabled Sessions: Peer LDP Ident: 192.168.0.1:0, State: estab Peer LDP Ident: 192.168.0.3:0, State: estab
PE1	FS#show mpls ldp graceful-restart Default VRF: LDP Graceful Restart is enabled Neighbor Liveness Timer: 120 seconds Max Recovery Time: 120 seconds ForwardingState Holding Time: 300 seconds Down Neighbor Database (0 records): Graceful Restart-enabled Sessions: Peer LDP Ident: 192.168.0.2:0, State: estab FS#show bgp vpnv4 unicast all neighbor BGP neighbor is 192.168.0.1, remote AS 200, internal link BGP version 4, remote router ID 192.168.0.1 BGP state = Established, up for 02:49:47 Last read 00:00:47, hold time is 180, keepalive interval is 60 seconds Neighbor capabilities: Route refresh: advertised and received(new) Address family VPNv4 Unicast: advertised and received Graceful Restart Capabilty: advertised and received Remote Restart timer is 120 seconds Address families preserved by peer: VPNv4 Unicast

3.5 Monitoring and Maintenance

Displaying

Description	Command

Displays LDP GR sessions and session parameters.	**show mpls ldp graceful-restart**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs LDP GR.	**debug mpls ldp graceful-restart**

4 Configuring MPLS BFD

The routers and router icons described in this chapter refer to general routers as well as Layer-3 switches running routing protocols. We use "router series" or "switch series" to refer to specific devices.

4.1 Overview

MPLS Bidirectional Forwarding Detection (BFD) is implemented based on *BFD for MPLS LSPs* defined by the IETF. As a part of BFD application, it describes a method of detecting MPLS LSPs.

For details about BFD, see *Configuring BFD*.

On an MPLS network, the following methods are used to detect LSP failures:

- MPLS Operation Administration and Maintenance (OAM) mechanism: This mechanism can detect, determine, and locate defects or failures of MPLS networks. However, the MPLS OAM mechanism is being standardized, and other OAM mechanisms have been applied to networks for only a short time. Therefore, perhaps not all devices support OAM.
- Hello packet mechanism of the MPLS signaling protocol: It takes a long time (usually several seconds) for this mechanism to detect failures. Therefore, this mechanism will cause massive traffic loss.

With the following features, MPLS BFD can solve the problems:

- Interoperability: MPLS BFD provides a uniform detection mechanism for the entire network.
- Quick detection: MPLS BFD enables light-load quick detection, enables backup forwarding paths quickly, and improves the reliability of MPLS networks.
- BFD can be used to detect failures of the data plane on an MPLS LSP, and BFD packets have fixed formats. Therefore, BFD can be implemented on hardware and traverse firewalls.

4.2 Applications

Application	Description
Detecting Static LSPs by Using BFD	Detects the connectivity of static LSPs by using Bidirectional Forwarding Detection (BFD). Correlates static LSPs with BFD. When a static LSP is in Down state, the LSP is no longer used as a forwarding path.
Detecting LDP LSPs by Using BFD	Detects the connectivity of LDP LSPs by using BFD. Correlates LDP LSPs with BFD. When an LDP LSP is in Down state, the LSP is no longer used as a forwarding path.
Detecting BGP LSPs by Using BFD	Detects the connectivity of BGP LSPs by using BFD. Correlates BGP LSPs with BFD. When a BGP LSP is in Down state, the LSP is no longer used as a forwarding path.

4.2.1 Detecting Static LSPs by Using BFD

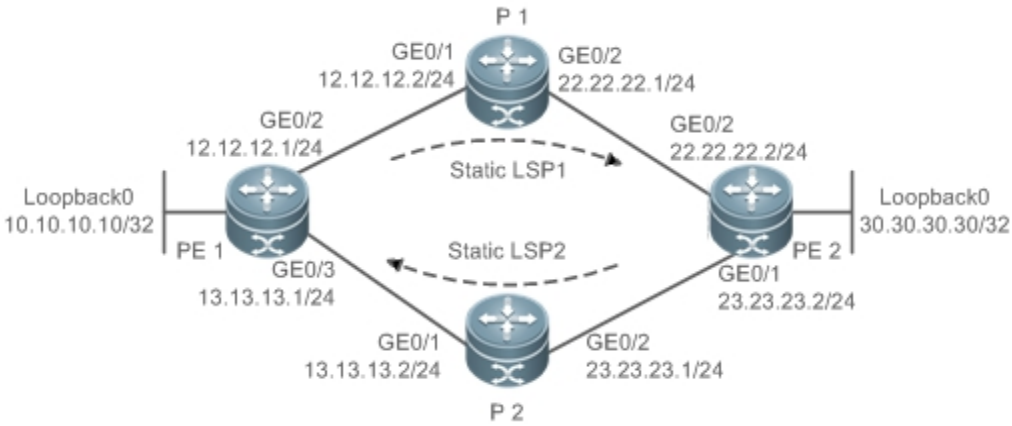
Scenario

If static LSPs are used on an MPLS network, you can detect the connectivity of static LSPs by using BFD.

Take PE1 in the following figure as an example. An LSP is unidirectional and BFD is bidirectional, so two LSPs are required for detection. (LSPs in both directions can share one path or use different paths. The following figure shows the latter.)

- When static LSP1 is faulty, BFD for PE2 will perceive the fault and use static LSP2 to inform PE1 of such fault. BFD sessions will be in Down state, and PE1 will delete the Facsimile Transmission Network (FTN) of static LSP1 to avoid selecting LSP1 as a forwarding path.
- When static LSP1 is recovered, the BFD sessions between PE1 and PE2 will return to UP state, and PE1 will reassign an FTN to LSP1 so that LSP1 is available again.

Figure 4- 1



Remarks	Static LSP1 is used to transmit data from PE1 to PE2. Static LSP2 is used to transmit data from PE2 to PE1. BFD-based LSP1 detection is configured on PE1. BFD-based LSP2 detection is configured on PE2.

Deployment

- Run OSPF on the entire network for unicast routing.
- Enable MPLS forwarding on the entire network to forward MPLS packets.
- Configure the static FTN connected to PE2 on PE1, and set P1 as the next hop. Configure the static ILM connected to PE2 on P1, and establish static LSP1.
- Configure the static FTN connected to PE1 on PE2, and set P2 as the next hop. Configure the static ILM connected to PE1 on P2, and establish static LSP2.
- Configure BFD-based static LSP detection on PE1 and PE2 respectively.

4.2.2 Detecting LDP LSPs by Using BFD

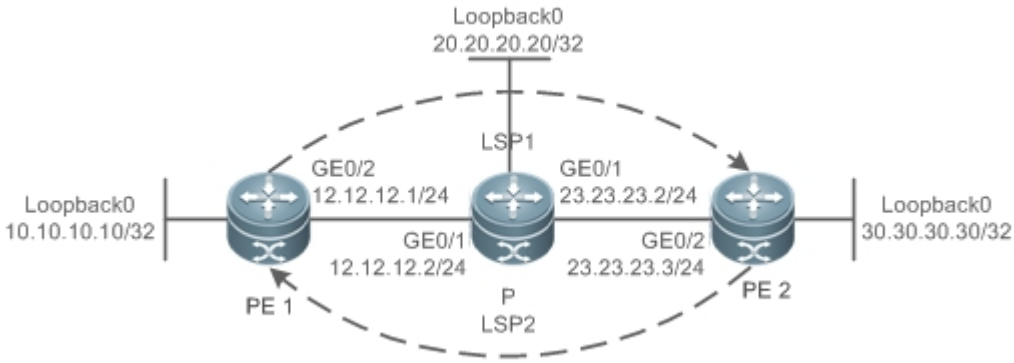
Scenario

If LSPs are established through LDP on anMPLS network, you can detect the connectivity of LDP LSPs by using BFD.

Take PE1 in the following figure as an example (the two LSPs share the same path):

- When LSP1 is faulty, BFD for PE2 will perceive the fault. If LDP LSP2 is available, BFD for PE1 will be informed of such fault; if LSP2 is unavailable, BFD for PE1 will perceive such fault. In both cases, BFD sessions will be in Down state, and PE1 will delete the FTN of LSP1 to avoid selecting LSP1 as a forwarding path.
- When LDP LSP1 is recovered, the BFD sessions between PE1 and PE2 will return to UP state, and PE1 will reassign an FTN to LSP1 so that LSP1 is available again.

Figure 4- 2



Remarks	LDP LSP1 is used to transmit data from PE1 to PE2. LDP LSP2 is used to transmit data from PE2 to PE1. BFD-based LDP LSP detection is configured on PE1. BFD-based LDP LSP detection is configured on PE2.

Deployment

- Run OSPF on the entire network for unicast routing.
- Enable MPLS forwarding on the entire network to forward MPLS packets.
- Enable LDP on the entire network to establish LDP LSPs.
- Configure BFD-based LDP LSP detection on PE1 and PE2 respectively.

4.2.3 Detecting BGP LSPs by Using BFD

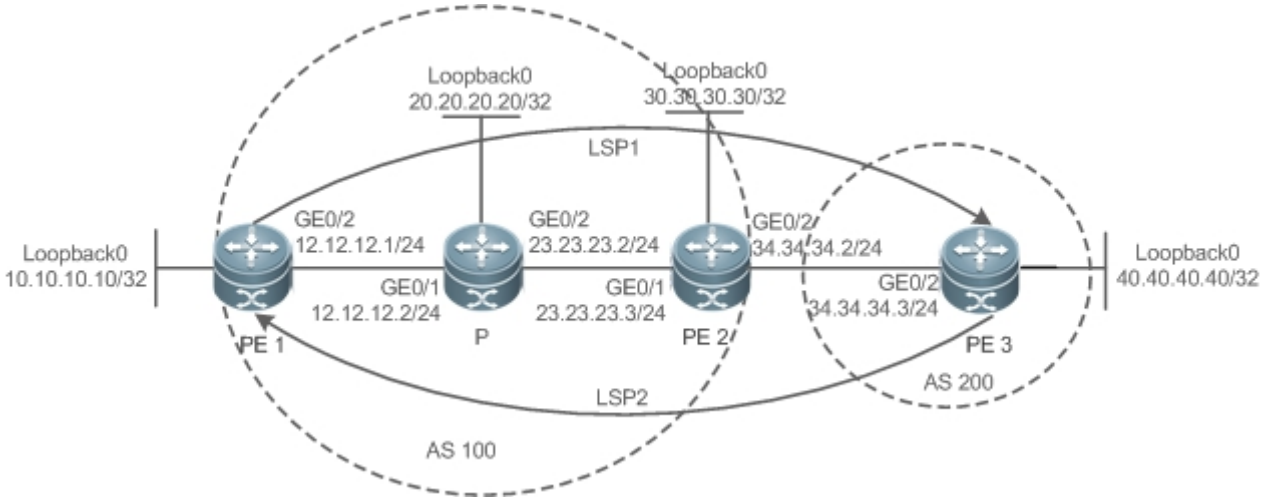
Scenario

If LSPs are established throughBGP on anMPLS network, you can detect the connectivity of BGP LSPs by using BFD.

Take PE1 in the following figure as an example (the two LSPs share the same path):

- When BGP LSP1 is faulty, BFD for PE3 will perceive the fault. If BGP LSP2 is available, BFD for PE1 will be informed of such fault; if LSP2 is unavailable, BFD for PE1 will perceive such fault. In both cases, BFD sessions will be in Down state, and PE1 will delete the FTN of LSP1 to avoid selecting LSP1 as a forwarding path.
- When BGP LSP1 is recovered, the BFD sessions between PE1 and PE3 will return to UP state, and PE1 will reassign an FTN to LSP1 so that LSP1 is available again.

Figure 4- 3



Remarks	BGP LSP1 is used to transmit data from PE1 to PE3. BGP LSP2 is used to transmit data from PE3 to PE1. BFD-based BGP LSP detection is configured on PE1. BFD-based BGP LSP detection is configured on PE3.

Deployment

- Deploy PE1 and PE2 in one Autonomous System (AS), and deploy PE2 and PE3 in the other AS.
- Run OSPF, LDP, and BGP in each AS.
- Run BGP between ASs to assign labels to routes and establish BGP LSPs.
- Enable MPLS forwarding on the entire network to forward MPLS packets.
- Configure BFD-based LSP detection on PE1 and PE3 respectively.

4.3 Features

Basic Concepts

BFD Session

BFD uses local identifier (My Discriminator) and remote identifier (Your Discriminator) to distinguish apart multiple BFD sessions. Identifiers can be configured automatically or manually.

- If identifiers are configured manually, BFD sessions can be established directly.
- If identifiers are automatically configured, before devices establish BFD sessions, local identifiers must be included in LSP Ping echo packets for session negotiation and remote identifiers must be learned.

BFD Mode

Before two devices exchange BFD control packets, they need to establish a BFD session. The control plane and data plane must share the same path. BFD sessions can be operated in two modes:

- Asynchronous mode: Both ends send control packets to each other regularly. If one end fails to receive packets within a specified period of time, it indicates that the session is in Down state.
- Query mode: After a session is established, both ends send control packets to each other only when they need to detect the connectivity. If the recipient does not reply, the sender will consider the session in Down state.

In addition to the preceding modes, BFD defines the echo function. With this function, a device can send customized packets to the peer end. The peer end returns the packets to the sender without processing them. The echo function is applicable to both asynchronous mode and query mode.

Currently, BFD-LSP correlation supports the asynchronous mode but does not support the query mode or the echo function.

Overview

Feature	Description
Establishing BFD Sessions	Establishes BFD sessions between the devices at both ends of the LSP to be detected.
Correlating BFD with LSPs	When any fault is detected by BFD, the LSP needs to be disabled temporarily and restarted after troubleshooting.

4.3.1 Establishing BFD Sessions

During BFD session initialization, the ingress LSR and egress LSR can be active or passive, which depends on the application scenario. However, at least one of them is active. The following describes two scenarios:

Both Active

As LSPs are unidirectional, the following two scenarios must be considered:

- When packets are sent from the ingress LSR to the egress LSR, the LSP is detected by using BFD; when packets are sent from the egress LSR to the ingress LSR, the other LSP is detected by using BFD.

The ingress LSR sends an LSP Ping echo request packet with a local identifier to the egress LSR. The egress LSR obtains a remote identifier from the received packet. In this way, the egress LSR has both local identifier (generated by the egress LSR) and remote identifier. Then the egress LSR sends a BFD control packet to the ingress LSR. The ingress LSR obtains a remote identifier from the received packet. In this way, the ingress LSR has both local identifier (generated by the ingress LSR) and remote identifier. Then the ingress LSR sends a BFD control packet to the egress LSR. A BFD session is then initialized between both ends.

Note: When receiving an echo request packet, the egress LSR can choose to or not to return an echo reply packet. If the egress LSR returns an echo reply packet, a local identifier (generated by the egress LSR) must be included in the echo reply packet. In this way, the ingress LSR can obtain a remote identifier from both the BFD control packet and the echo reply packet.

The egress LSR works in a similar way to the ingress LSR.

- When packets are sent from the ingress LSR to the egress LSR, the LSP is detected by using BFD; when packets are sent from the egress LSR to the ingress LSR, the IP address is detected by using BFD (multiple hops exist).

In this case, BFD session can be established only by manually configuring identifiers on the ingress LSR and the egress LSR. That is, after identifiers are configured manually, BFD sessions can be established without identifier negotiation.

One Active, the Other Passive

The active one sends an LSP Ping echo request packet with a local identifier to the passive one. The passive one obtains a remote identifier from the received packet. In this way, the passive one has both local identifier (generated by the passive one) and remote identifier. Then the passive one sends a BFD control packet to the active one. The active one obtains a remote identifier from the received packet. In this way, the active one has both local identifier (generated by the active one) and remote identifier. Then the active one sends a BFD control packet to the passive one. A BFD session is then initialized between both ends.

Note: When receiving an echo request packet, the passive one can choose to or not to return an echo reply packet. If the passive one returns an echo reply packet, a local identifier (generated by the passive one) must be included in the echo reply packet. In this way, the active one can obtain a remote identifier from both the BFD control packet and the echo reply packet.

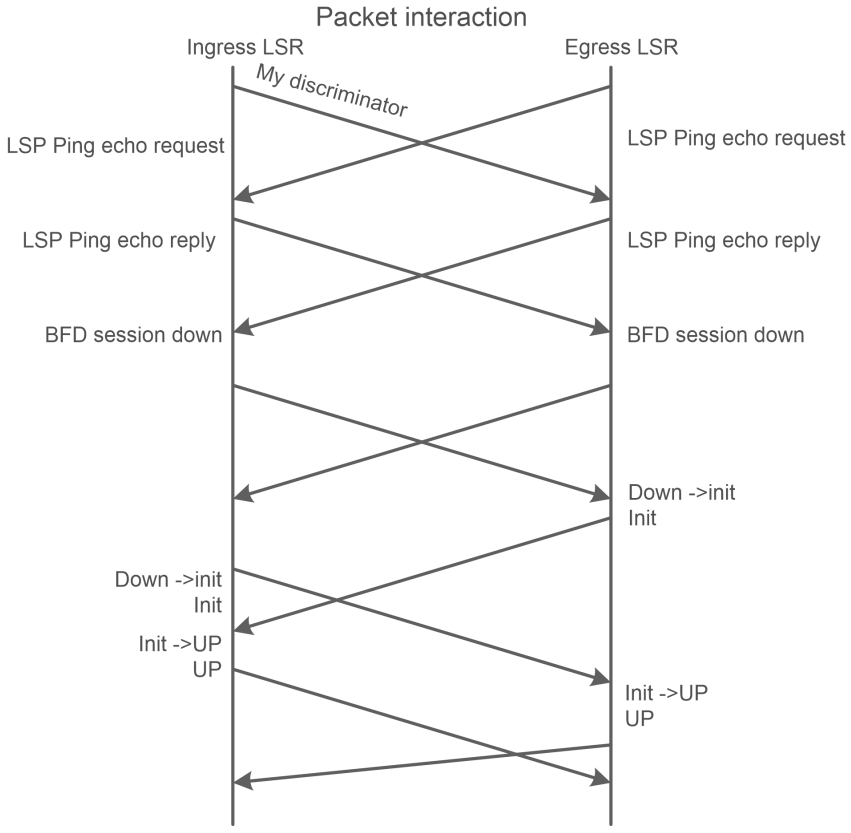
Before receiving an echo request packet from the active one, the passive one will not send a BFD control packet to the active one.

Currently, BFD-LSP correlation supports the "both active" mode but does not support the "one active, the other passive" mode.

Working Principle

The following describes the process of BFD session establishment in the first scenario (when packets are sent from the ingress LSR to the egress LSR, the LSP is detected by using BFD; when packets are sent from the egress LSR to the ingress LSR, the other LSP is detected by using BFD) when both ends are active.

Figure 4- 4



- Before enabling BFD, both ingress LSR and egress LSR must learn remote identifiers and ensure that LSPs are in Up state. As shown in the figure, the ingress LSR sends an LSP Ping echo request packet with a local identifier to the egress LSR. Upon receipt of the echo request packet, the egress LSR returns an echo reply packet with a local identifier (generated by the egress LSR) to the ingress LSR. The case is similar with the egress LSR. Note: When no identifier is specified, both ends need to learn the identifier of the other through an LSP Ping echo packet. If local and remote identifiers are specified by both ends, the devices will skip to the next step. For information on identifier configuration, see the section "Correlating BFD with LSPs".
- The ingress LSR and egress LSR enable BFD. They are in Down state initially. BFD packets are in Down state.
- The egress LSR receives the BFD packets and locally switches BFD to Init state. The BFD packets are in Init state.
- The egress LSR no longer processes BFD packets that are in Down state.

- The case is similar with the ingress LSR.
- The egress LSR receives BFD packets that are in Init state and locally switches BFD to Up state.
- The case is similar with the ingress LSR.
- The Up state indicates that BFD sessions are established successfully.

Related Configuration

Configuring BFD Session Parameters on an Interface

By default, BFD session parameters are not configured on an interface.

By running the **bfd interval** command, you can configure or cancel BFD session parameters on an interface.

The BFD sessions for detecting LSPs can be established only when the **bfd interval** command is configured on the LSP interface, the command for BFD-based LSP detection described in the next section is configured, and the corresponding LSPs exist. The speed of fault detection is determined by the session parameters designated by the**bfd interval**command.

4.3.2 Correlating BFD with LSPs

BFD sessions are distinguished apart through local identifier (My Discriminator) and remote identifier (Your Discriminator). Based on the identifier specification mode, BFD and LSP can be correlated in two modes:

- Manual mode

If identifiers are configured manually, BFD sessions can be established directly.

- Automatic mode

If identifiers are automatically configured, before devices establish BFD sessions, local identifiers must be included in LSP Ping packets for session negotiation and remote identifiers must be learned.

Based on types of LSPs, BFD-LSP correlation includes BFD-based static LSP detection, BFD-based LDP LSP detection, and BFD-based BGP LSP detection.

Working Principle

BFD-based Static LSP Detection

BFD-based static LSP detection can be configured manually or automatically. An LSP is a unidirectional link, while BFD is a bidirectional detection mechanism. Therefore, when detecting unidirectional static LSPs by using BFD, you can select any of the following methods to detect their backward links:

- Detecting backward links with IP
- Detecting backward links with static LSPs

After the detection method is selected, BFD sessions will be established and in UP state. Then, if any fault is detected by BFD, BFD sessions will be in Down state; if the configuration command has the **process-state** parameter, Ingress LSR will delete the FTN of the faulty LSP so that such LSP will not be selected as a forwarding path. If the fault is removed, BFD sessions will return to UP state, and Ingress LSR will re-add the FTN of such LSP so that such LSP will be reused as a forwarding path.

BFD-based LDP LSP Detection

BFD-based LDP LSP detection can be configured manually or automatically. An LSP is a unidirectional link, while BFD is a bidirectional detection mechanism. Therefore, when detecting unidirectional LDP LSPs by using BFD, you can select any of the following methods to detect their backward links:

- Detecting backward links with IP
- Detecting backward links with LDP LSPs

BGP LSPs are used for basic VPN/PW services. Therefore, BFD-based LDP LSP detection detects failures of VPN/PW. By using this function, you can quickly detect various MPLS-based network applications such as VPN Fast Reroute (FRR) and PW FRR, thereby protecting services and improving reliability of MPLS networks. LDP LSPs and static LSPs share the same process of BFD-based fault detection, process of recovery, and relevant processing.

BFD-based BGP LSP Detection

The function can be configured both manually and automatically. An LSP is a unidirectional link, while BFD is a bidirectional detection mechanism. Therefore, when detecting unidirectional BGP LSPs by using BFD, you can select any of the following methods to detect their backward links:

- Detecting backward links with IP
- Detecting backward links with BGP LSPs

BGP LSPs are used for basic VPN/PW services. Therefore, BFD-based BGP LSP detection detects faults of VPN/PW. By using this function, you can quickly detect various MPLS-based network applications such as VPN Fast Reroute (FRR) and PW FRR, thereby protecting services and improving reliability of MPLS networks. LDP LSPs, static LSPs and LDP LSPs share the same process of BFD-based fault detection, process of recovery, and relevant processing.

Related Configuration

Configuring BFD-based Static LSP Detection

By default, BFD-based static LSP detection is disabled.

Run the **bfd bind static-lsp** command to enable BFD-based static LSP detection.

The number of BFD sessions is limited. If the number of BFD sessions submitted by a user exceeds the maximum allowed, the system will generate log information to prompt the user.

Configuring BFD-based LDP LSP Detection

By default, BFD-based LDP LSP detection is disabled.

Run the **bfd bind ldp-lsp** command to enable BFD-based LDP LSP detection.

The number of BFD sessions is limited. If the number of BFD sessions submitted by a user exceeds the maximum allowed, the system will generate log information to prompt the user.

Configuring BFD-based BGP LSP Detection

By default, BFD-based BGP LSP detection is disabled.

Run the **bfd bind bgp-lsp** command to enable BFD-based BGP LSP detection.

The number of BFD sessions is limited. If the number of BFD sessions submitted by a user exceeds the maximum allowed, the system will generate log information to prompt the user.

4.4 Configuration

Configuration items	Description and Command	
Configuring BFD-based Static LSP Detection	(Mandatory) It is used to detect the connectivity of static LSPs by using BFD.	
	bfd bind static-lsp	Configures BFD-based static LSP detection.
	bfd bind backward-lsp-with-ip	Configures BFD-based detection of backward LSPs with IP.
Configuring BFD-based LDP LSP Detection	(Mandatory) It is used to detect the connectivity of LDP LSPs by using BFD.	
	bfd bind ldp-lsp	Configures BFD-based LDP LSP detection.
	bfd bind backward-lsp-with-ip	Configures BFD-based detection of backward LSPs with IP.
Detecting BGP LSPs by Using BFD	(Mandatory) It is used to detect the connectivity of BGP LSPs by using BFD.	
	bfd bind bgp-lsp	Configures BFD-based BGP LSP detection.
	bfd bind backward-lsp-with-ip	Configures BFD-based detection of backward LSPs with IP.

4.4.1 Configuring BFD-based Static LSP Detection

Configuration Effect

The connectivity of static LSPs can be detected by using BFD. Correlating static LSPs with BFD can prevent a private route from using a faulty static LSP as a forwarding path.

Notes

Only host-routed static LSPs can be detected by using BFD.

In application scenarios where failures need to be detected by using BFD, for example, the BFD-LSP correlation scenario, you need to configure **process-state**.

Before configuring BFD-based static LSP detection, you need to:

- Enable MPLS.
- Configure static LSPs.

Before configuring BFD-based static LSP detection, you need to configure the following:

- Local identifier (My Discriminator) and remote identifier (Your Discriminator) of a BFD session
- Backward link detection
- BFD session parameters: interval for sending BFD control packets, interval for receiving BFD control packets, and BFD control packet detection multiplier

Configuration Steps

Configuring BFD for Ingress LSR

- Mandatory.

- Configure BFD sessions on the outbound interface of an LSP.
- Configure BFD-based static LSP detection and process the BFD session state. If BFD-based static LSP detection is configured manually on the ingress LSR, it must be configured manually on the egress LSR. That is, BFD-based static LSP detection must be configured in the same way at both ends.

If identifiers are configured manually, local and remote identifiers configured on the ingress LSR must match those configured on the egress LSR.

Configuring BFD for Egress LSR

- Mandatory.
- Configure BFD sessions on the outbound interface of the backward LSP (route if IP is used).
- Configure BFD-based static LSP detection. Do not process the BFD session state. If BFD-based static LSP detection is configured manually on the ingress LSR, it must be configured manually on the egress LSR. That is, BFD-based static LSP detection must be configured in the same way at both ends.

Or:

- Configure BFD-based detection of backward LSPs with IP. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs.

If identifiers are configured manually, local and remote identifiers configured on the egress LSR must match those configured on the ingress LSR.

Verification

Run the **show bfd neighbors** command to check whether BFD sessions are in Up state.

Related Commands

Configuring BFD-based Static LSP Detection

Command	**bfd bind static-lsp peer-ip***ip-addresss***ource-ip** *ip-address* [**local-discriminator** *discr-value***remote-discriminato**r *discr-value*] [**process-state**]
Parameter Description	**peer-ip***ip-address:* Indicates the peer IP address of the static LSP bound with BFD sessions. **source-ip***ip-address:* Indicates the source IP address of BFD packets. **local-discriminator***discr-value:* Configures the local identifier of a BFD session. The value ranges from 1 to 8191. **remote-discriminator***discr-value:* Configures the remote identifier of the BFD session. The value ranges from 1 to 8191. **process-state:** Processes the BFD session state. In application scenarios where failures need to be detected by using BFD, for example, the BFD-LSP correlation scenario, you need to configure **process-state**.
Command Mode	Global configuration mode
Usage Guide	● This command can be run on the ingress node of the LSP. ● If BFD has been configured, BFD configuration items cannot be established. After BFD is configured, if a static LSP exists, BFD sessions will be immediately created; if no static LSP exists, BFD sessions will not be created until astatic LSP exists. ● If a static LSP is deleted, the BFD sessions bound with the LSP will be deleted. However, configuration items of the BFD sessions will be retained. When a static LSP exists, BFD sessions will be re-created.

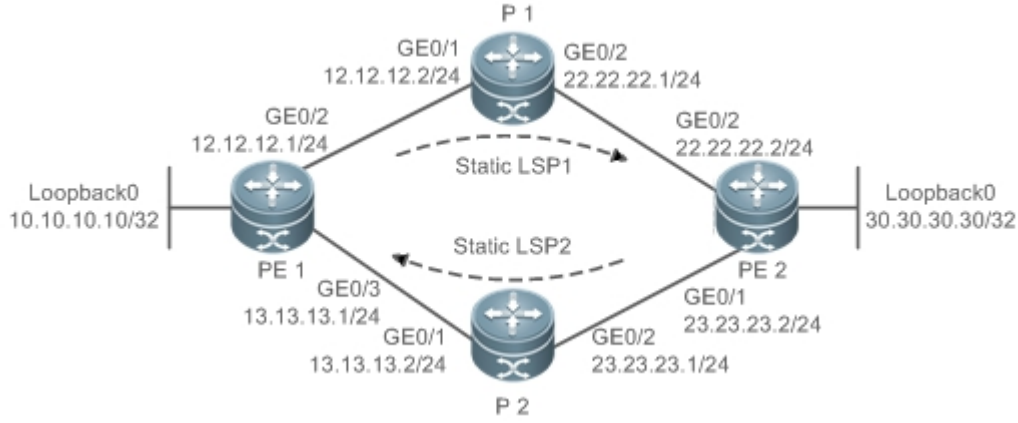
	● You can configure local and remote identifiers for BFD sessions. If no local identifier is configured, the system will automatically elect a local identifier. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs. ● After BFD sessions are established, identifiers cannot be modified. ● The system regularly queries submitted BFD configuration items to establish BFD sessions. ● The number of BFD sessions is limited. If the number of requested BFD sessions exceeds the upper limit, the system will generate a log to prompt you.

Configuring BFD-based Detection of Backward LSPs with IP

Command	**bfd bind backward-lsp-with-ip peer-ip***ip-address*[**vrf***vrf-name*] **interface** *interface-typeinterface-number*[**source-ip** *ip-address*] {**local-discriminator***discr-value* **remote-discriminator** *discr-value*}
Parameter Description	**peer-ip***ip-address:* Indicates the peer IP address of the LSP bound with BFD sessions. **vrf***vrf-name:* Indicates the name of VRF bound with BFD sessions. **interface** *interface-type interface-number:* Configures the interface type and number. **source-ip***ip-address:* Indicates the source IP address of BFD packets. **local-discriminator***discr-value:* Configures the local identifier of a BFD session. The value ranges from 1 to 8191. **remote-discriminator***discr-value:* Configures the remote identifier of the BFD session. The value ranges from 1 to 8191.
Command Mode	Global configuration mode
Usage Guide	Configures BFD-based detection of backward LSPs with IP. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs.

Configuration Example

Detecting the Connectivity of Static LSPs by Using BFD

Scenario **Figure 4- 5**	
Configuration Steps	● Configure the IP address and OSPF on each node. ● On each node, enable global MPLS forwarding, and enable MPLS labeled packet forwarding on an interface. ● On each node, configure static MPLS routes to enable MPLS traffic forwarding. ● Configure BFD-based static LSP1 detection on PE1. ● Configure BFD-based static LSP2 detection on PE2 (detect backward links with LSPs).
PE1	FS#configure terminal

	Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 12.12.12.1 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface gigabitEthernet 0/3 FS(config-if-GigabitEthernet 0/3)#ip address 13.13.13.1 255.255.255.0 FS(config-if-GigabitEthernet 0/3)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 10.10.10.10 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network12.12.12.1 255.255.255.0 area 0 Router(config-router)#network13.13.13.1 255.255.255.0 area 0 Router(config-router)#network10.10.10.10 255.255.255.255 area 0 Router(config-router)#exit FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#interface gigabitEthernet 0/3 FS(config-if-GigabitEthernet 0/3)#label-switching FS(config-if-GigabitEthernet 0/3)#mpls ip Router(config-if-GigabitEthernet 0/3)#exit FS(config)#mpls static ftn 30.30.30.30/32 out-label 16 nexthop gigabitEthernet 0/2 12.12.12.2 FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#bfd interval 50 min_rx 50 multiplier 3 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#bfd bind static-lsp peer-ip 30.30.30.30source-ip 10.10.10.10 local-discriminator 1 remote-discriminator 2 process-state FS(config)#exit
PE2	FS#configure terminal

	Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 22.22.22.2 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#ip address 23.23.23.2 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 30.30.30.30 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network 22.22.22.2 255.255.255.0 area 0 Router(config-router)#network 23.23.23.2 255.255.255.0 area 0 Router(config-router)#network 30.30.30.30 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip Router(config-if-GigabitEthernet 0/1)#exit FS(config)#mpls static ftn 10.10.10.10/32 out-label 16 nexthop gigabitEthernet 0/1 23.23.23.1 FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#bfd interval 50 min_rx 50 multiplier 3 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#bfd bind static-lsp peer-ip 10.10.10.10source-ip 30.30.30.30 local-discriminator 2 remote-discriminator 1 FS(config)#exit
P1	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#ip address 12.12.12.2 255.255.255.0

	FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 22.22.22.1 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 20.20.20.20 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network12.12.12.2 255.255.255.0 area 0 Router(config-router)#network 22.22.22.1 255.255.255.0 area 0 Router(config-router)#network 20.20.20.20 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip Router(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#mpls static ilm in-label 16 forward-action swap-label 3nexthop gigabitEthernet 0/2 22.22.22.2 fec 30.30.30.30/32
P2	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#ip address 13.13.13.2 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 23.23.23.1 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 40.40.40.40 255.255.255.255 FS(config-Loopback 0)#exit

	FS(config)#router ospf 1 Router(config-router)#network 13.13.13.2 255.255.255.0 area 0 Router(config-router)#network 23.23.23.1 255.255.255.0 area 0 Router(config-router)#network 40.40.40.40 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip Router(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#mpls static ilm in-label 16 forward-action swap-label 3nexthop gigabitEthernet 0/113.13.13.1 fec 10.10.10.10/32
Verification	BFD sessions are in Up state on PE1 and PE2.
PE1	FS# show bfd neighbors details OurAddr NeighAddr LD/RD RH Holdown(mult) State Int 10.10.10.10 30.30.30.30 1/2 1 532 (3) Up Ge2/1 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 200000, MinRxInt: 200000, Multiplier: 3 Received MinRxInt: 50000, Received Multiplier: 3 Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332 Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196 Registered protocols: static-lsp Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24 My Discr.: 2 - Your Discr.: 1

	Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0
PE2	FS# show bfd neighbors details OurAddr NeighAddr LD/RD RH Holdown(mult) State Int 30.30.30.30 10.10.10.10 1/2 1 532 (3) Up Ge2/1 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 200000, MinRxInt: 200000, Multiplier: 3 Received MinRxInt: 50000, Received Multiplier: 3 Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332 Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196 Registered protocols: static-lsp Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24 My Discr.: 2 - Your Discr.: 1 Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0

Common Errors

- BFD sessions are not configured on the outbound interface of the LSP.
- The interface on which BFD sessions are configured is not the outbound interface of the LSP.
- Static FTN/ILM entries do not take effect.
- The LSP is incomplete.
- The global **mpls ip** command is not run.
- The **label-switching** command for enabling MPLS forwarding on the interface is not run.
- The source address and destination address of the BFD-detected devices at both ends of an LSP are not matched.
- The identifiers configured (manually) for the BFD-detected devices at both ends of an LSP are not matched.
- Forward Error Correction (FEC) corresponding to the detected LSP is not routed by the host.

4.4.2 Configuring BFD-based LDP LSP Detection

Configuration Effect

- The connectivity of LDP LSPs can be detected by using BFD. Correlating an LDP LSP with BFD can prevent a private route from using the LSP as a forwarding path when it is in Down state.

Notes

- Only host-routed LDP LSPs can be detected by using BFD.
- One LSP can be bound with only one BFD session.
- An LDP LSP must be bound with BFD at its ingress.
- In application scenarios where failures need to be detected by using BFD, for example, the BFD-LSP correlation scenario, you need to configure **process-state.**
- Before configuring BFD-based LDP LSP detection, you need to:

23. Enable MPLS.
24. Enable LDP.

- Before configuring BFD-based LDP LSP detection, you need to configure the following:

25. Local identifier (My Discriminator) and remote identifier (Your Discriminator) of a BFD session
26. Backward link detection
27. BFD session parameters: interval for sending BFD control packets, interval for receiving BFD control packets, and BFD control packet detection multiplier

Configuration Steps

Configuring BFD for the Ingress LSR

- Mandatory.
- Configure BFD sessions on the outbound interface of an LSP.
- Configure BFD-based LDP LSP detection and process the BFD session state. If BFD-based LDP LSP detection is configured manually on the ingress LSR, it must be configured manually on the egress LSR. That is, BFD-based LDP LSP detection must be configured in the same way at both ends.

If identifiers are configured manually, local and remote identifiers configured on the ingress LSR must match those configured on the egress LSR.

Configuring BFD for Egress LSR

- Mandatory.
- Configure BFD sessions on the outbound interface of an LSP.
- Configure BFD-based LDP LSP detection. Do not process the BFD session state. If BFD-based LDP LSP detection is configured manually on the ingress LSR, it must be configured manually on the egress LSR. That is, BFD-based LDP LSP detection must be configured in the same way at both ends.

Or:

- Configures BFD-based detection of backward LSPs with IP. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs.

If identifiers are configured manually, local and remote identifiers configured on the egress LSR must match those configured on the ingress LSR.

Only host-routed LDP LSPs can be detected by using BFD.

One LSP can be bound with only one BFD session.

An LDP LSP must be bound with BFD at its ingress.

Verification

Run the **show bfd neighbors** command to check whether BFD sessions are in Up state.

Related Commands

Configuring BFD-based LDP LSP Detection

Command	**bfd bind ldp-lsp peer-ip***ip-address*[**vrf***vrf-name*] **nexthop***ip-address*[**interface***interface-typeinterface-number*] **source-ip***ip-address*[**local-discriminator***discr-value***remote-discriminator** *discr-value*] [**process-state**]
Parameter Description	**peer-ip***ip-address:* Indicates the peer IP address of the static LSP bound with BFD sessions. **vrf***vrf-name:* Indicates VRF corresponding to LDP LSPs. **nexthop***ip-address:* Specifies the next-hop IP address of LDP LSPs. **interface** *interface-type interface-number:* Configures the interface type and number. **source-ip***ip-address:* Indicates the source IP address of BFD packets. **local-discriminator***discr-value:* Configures the local identifier of a BFD session. The value ranges from 1 to 8191. **remote-discriminator***discr-value:* Configures the remote identifier of the BFD session. The value ranges from 1 to 8191. **process-state:** Processes the BFD session state. In application scenarios where failures need to be detected by using BFD, for example, the BFD-LSP correlation scenario, you need to configure **process-state**.
Command Mode	Global configuration mode
Usage Guide	● This command can be run on the ingress node of the LSP. ● If BFD has been configured, BFD configuration items cannot be established. After BFD is configured, if an LDP LSP exists, BFD sessions will be immediately created; if no LDP LSP exists, BFD sessions will not be created until an LDP LSP exists. ● If an LDP LSP is deleted, the BFD sessions bound with the LSP will be deleted. However, configuration items of the BFD sessions will be retained. When an LDP LSP exists, BFD sessions will be re-created. ● You can configure local and remote identifiers for BFD sessions. If no local identifier is configured, the system will automatically elect a local identifier. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs. ● If the outbound interface of the detected LSP is lent or borrowed, you need to specify an outbound interface. ● After BFD sessions are established, identifiers cannot be modified. ● The system regularly queries submitted BFD configuration items to establish BFD sessions. ● The number of BFD sessions is limited. If the number of requested BFD sessions exceeds the upper limit, the system will generate a log to prompt you.

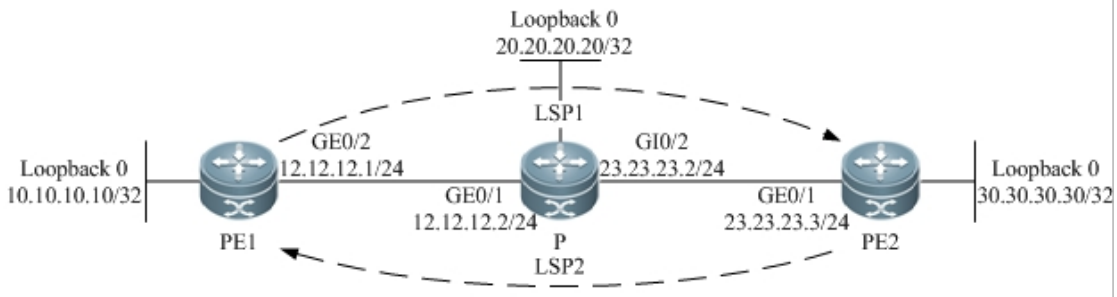
Configuring BFD-based Detection of Backward LSPs with IP

Command	**bfd bind backward-lsp-with-ip peer-ip***ip-address*[**vrf** *vrf-name*] **interface***interface-typeinterface-number*[**source-ip**

	ip-address] l**ocal-discriminator** *discr-value***remote-discriminato**r *discr-value*
Parameter Description	**peer-ip***ip-address:* Indicates the peer IP address of the LSP bound with BFD sessions. **vrf***vrf-name:* Indicates the name of VRF bound with BFD sessions. **interface** *interface-type interface-number:* Configures the interface type and number. **source-ip***ip-address:* Indicates the source IP address of BFD packets. **local-discriminator***discr-value:* Configures the local identifier of a BFD session. The value ranges from 1 to 8191. **remote-discriminator***discr-value:* Configures the remote identifier of the BFD session. The value ranges from 1 to 8191.
Command Mode	Global configuration mode
Usage Guide	Configures BFD-based detection of backward LSPs with IP. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs.

Configuration Example

Detecting the Connectivity of LDP LSPs by Using BFD

Scenario **Figure 4- 6**	
Configuration Steps	● Configure the IP address and OSPF on each node. ● On each node, enable global MPLS forwarding, and enable MPLS labeled packet forwarding and LDP on an interface. ● Configure LDP to enable MPLS traffic forwarding. ● Configure BFD-based LDP LSP1 detection on PE1. ● Configure BFD-based LDP LSP2 detection on PE2 (detect backward links with LSPs).
PE1	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 12.12.12.1 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 10.10.10.10 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1

	Router(config-router)#network12.12.12.1 255.255.255.0 area 0 Router(config-router)#network10.10.10.10 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#bfd interval 50 min_rx 50 multiplier 3 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#bfd bind ldp-lsp peer-ip 30.30.30.30nexthop 12.12.12.2 interface gigabitEthernet 0/2 source-ip 10.10.10.10 local-discriminator 1 remote-discriminator 2 process-state
PE2	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#ip address 23.23.23.3 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 30.30.30.30 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network 23.23.23.3 255.255.255.0 area 0 Router(config-router)#network 30.30.30.30 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip FS(config-if-GigabitEthernet 0/1)#bfd interval 50 min_rx 50 multiplier 3 FS(config-if-GigabitEthernet 0/1)#exit

	FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#exit FS(config)#bfd bind ldp-lsp peer-ip 10.10.10.10nexthop 23.23.23.2 interface gigabitEthernet 0/1 source-ip 30.30.30.30 local-discriminator 2 remote-discriminator 1
P1	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#ip address 12.12.12.2 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 23.23.23.2 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 20.20.20.20 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network12.12.12.2 255.255.255.0 area 0 Router(config-router)#network 23.23.23.2 255.255.255.0 area 0 Router(config-router)#network 20.20.20.20 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip Router(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force
Verification	● BFD sessions are in Up state on PE1 and PE2.

PE1	FS# show bfd neighbors details OurAddr NeighAddr LD/RD RH Holdown(mult) State Int 10.10.10.10 30.30.30.30 1/2 1 532 (3) Up Ge0/2 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 200000, MinRxInt: 200000, Multiplier: 3 Received MinRxInt: 50000, Received Multiplier: 3 Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332 Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196 Registered protocols: ldp-lsp Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24 My Discr.: 2 - Your Discr.: 1 Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0
PE2	FS# show bfd neighbors details OurAddr NeighAddr LD/RD RH Holdown(mult) State Int 30.30.30.30 10.10.10.10 2/1 1 532 (3) Up Ge0/1 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 200000, MinRxInt: 200000, Multiplier: 3 Received MinRxInt: 50000, Received Multiplier: 3 Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332 Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196 Registered protocols: ldp-lsp Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24

	My Discr.: 2 - Your Discr.: 1 Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0

Common Errors

- BFD sessions are not configured on the outbound interface of an LSP.
- The interface on which BFD sessions are configured is not the outbound interface of an LSP.
- FTN/ILM entries are not assigned through LDP.
- An LSP is incomplete.
- The global **mpls ip** command is not run.
- The **label-switching** command for enabling MPLS forwarding on an interface is not run.
- The source address and destination address of the BFD-detected devices at both ends of an LSP are not matched.
- The identifiers configured (manually) for the BFD-detected devices at both ends of an LSP are not matched.
- The forwarding equivalence class (FEC) corresponding to the detected LSP is not routed by the host.

4.4.3 Configuring BFD-based BGP LSP Detection

Configuration Effect

- The connectivity of BGP LSPs can be detected by using BFD. By correlating a BGP LSP with BFD, you can prevent the private route from using the LSP as a forwarding path when it is in Down state.

Notes

- Only host-routed BGP LSPs can be detected by using BFD.
- One LSP can be bound with only one BFD session.
- A BGP LSP must be bound with BFD at its ingress.
- Before configuring BFD-based BGP LSP detection, you need to:

28. Enable MPLS.
29. Assign labels via BGP.

- Before configuring BFD-based BGP LSP detection, you need to configure the following:

1. Local identifier (My Discriminator) and remote identifier (Your Discriminator) of a BFD session
2. Backward link detection
3. BFD session parameters: interval for sending BFD control packets, interval for receiving BFD control packets, and BFD control packet detection multiplier

Configuration Steps

↘ **Configuring BFD for the Ingress LSR**

- Mandatory.

- Configure BFD sessions on the outbound interface of an LSP.
- Configure BFD-based BGP LSP detection and process the status of BFD sessions. If BFD is configured manually for Ingress LSR, it must be configured manually for Egress LSR.That is to say, configurations at both ends must be consistent.

If identifiers are configured manually, local and remote identifiers configured on the ingress LSR must match those configured on the egress LSR.

Configuring BFD for the Egress LSR

- Mandatory.
- Configure BFD sessions on the outbound interface of an LSP.
- Configure BFD-based BGP LSP detection. Do not process the BFD session state. If BFD-based BGP LSP detection is configured manually on the ingress LSR, it must be configured manually on the egress LSR. That is, BFD-based BGP LSP detection must be configured in the same way at both ends.

Or:

- Configures BFD-based detection of backward LSPs with IP. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs.

If identifiers are configured manually, local and remote identifiers configured on the egress LSR must match those configured on the ingress LSR.

Only host-routed BGP LSPs can be detected by using BFD.

One LSP can be bound with only one BFD session.

A BGP LSP must be bound with BFD at its ingress.

Verification

Run the **show bfd neighbors** command to check whether BFD sessions are in Up state.

Related Commands

Detecting the Connectivity of BGP LSPs by Using BFD

Command	**bfd bind bgp-lsppeer-ip***ip-address***source-ip** *ip-address* [**local-discriminator***discr-value***remote-discriminator***discr-value*]
Parameter Description	**peer-ip***ip-address:* Indicates the sink IP address of the BGP LSP bound with BFD sessions. **source-ip***ip-address:* Indicates the source IP address of BFD packets. **local-discriminator***discr-value:* Configures the local identifier of a BFD session. The value ranges from 1 to 8191. **remote-discriminator***discr-value:* Configures the remote identifier of the BFD session. The value ranges from 1 to 8191.
Command Mode	Global configuration mode
Usage Guide	● Enable global MPLS forwarding on each node, and enable the forwarding of MPLS label packets on the interface.Configure IBGP and assign labelsvia BGP between PE1 and PE2. Configure EBGP and assign labelsvia BGP between PE2 and PE3.Configure BFD-based BGP LSP1 detection on PE1.This command can be run on the ingress node of the LSP. ● If BFD has been configured, BFD configuration items cannot be established. After BFD is configured, if a BGP LSP

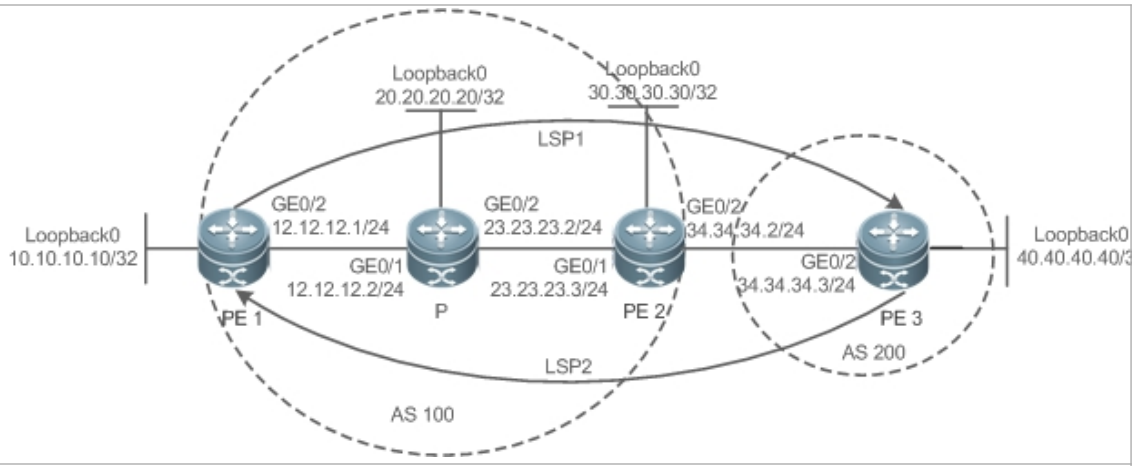
	exists, BFD sessions will be immediately created; if no BGP LSP exists, BFD sessions will not be created until an LSP exists. ● If a BGP LSP is deleted, the BFD sessions bound with the LSP will be deleted. However, configuration items of the BFD sessions will be retained. When a BGP LSP exists, BFD sessions will be re-created. ● You can configure local and remote identifiers for BFD sessions. If no local identifier is configured, the system will automatically elect a local identifier. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs. ● After BFD sessions are established, identifiers cannot be modified. ● The system regularly queries submitted BFD configuration items to establish BFD sessions. ● The number of BFD sessions is limited. If the number of requested BFD sessions exceeds the upper limit, the system will generate a log to prompt you.

Configuring BFD-based Detection of Backward LSPs with IP

Command	**bfd bind backward-lsp-with-ip peer-ip***ip-address*[**vrf** *vrf-name*] **interface***interface-typeinterface-number*[**source-ip** *ip-address*] l**ocal-discriminator** *discr-value***remote-discriminato**r *discr-value*
Parameter Description	**peer-ip***ip-address:* Indicates the peer IP address of the LSP bound with BFD sessions. **vrf***vrf-name:* Indicates the name of VRF bound with BFD sessions. **interface** *interface-type interface-number:* Configures the interface type and number. **source-ip***ip-address:* Indicates the source IP address of BFD packets. **local-discriminator***discr-value:* Configures the local identifier of a BFD session. The value ranges from 1 to 8191. **remote-discriminator***discr-value:* Configures the remote identifier of the BFD session. The value ranges from 1 to 8191.
Command Mode	Global configuration mode
Usage Guide	Configures BFD-based detection of backward LSPs with IP. If backward LSPs with IP are detected, local and remote identifiers must be configured manually for forward LSPs.

Configuration Example

Detecting the Connectivity of BGP LSPs by Using BFD

Scenario **Figure 4- 7**	
Configuration Steps	● Configure the IP address and loopback address on each node. ● Configure OSPF and LDP between PE1 and PE2.

	● On each node, enable global MPLS forwarding, and enable MPLS labeled packet forwarding on an interface. ● Configure IBGP and assign labels via BGP between PE1 and PE2. Configure EBGP and assign labels via BGP between PE2 and PE3. ● Configure BFD-based BGP LSP1 detection on PE1. ● Configure BFD-based BGP LSP2 detection on PE3 (detect backward links with LSPs).
PE1	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 12.12.12.1 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 10.10.10.10 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#interface loopback 1 FS(config-Loopback 0)#ip address 100.100.100.100 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network12.12.12.1 255.255.255.0 area 0 Router(config-router)#network10.10.10.10 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#exit FS(config)#router bgp 100 Router(config-router)#neighbor 30.30.30.30 remote-as 100 Router(config-router)#neighbor 30.30.30.30 update-source loopback 0 Router(config-router)#neighbor 30.30.30.30 send-label Router(config-router)#network 100.100.100.100 mask 255.255.255.255 Router(config-router)#exit FS(config)#interface gigabitEthernet 0/2

	FS(config-if-GigabitEthernet 0/2)#bfd interval 50 min_rx 50 multiplier 3 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#bfd bind bgp-lsp peer-ip 200.200.200.200source-ip 100.100.100.100 local-discriminator 1 remote-discriminator 2
P	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#ip address 12.12.12.2 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#ip address 23.23.23.2 255.255.255.0 FS(config-if-GigabitEthernet 0/2)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 20.20.20.20 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network12.12.12.2 255.255.255.0 area 0 Router(config-router)#network 23.23.23.2 255.255.255.0 area 0 Router(config-router)#network 20.20.20.20 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip Router(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)#label-switching FS(config-if-GigabitEthernet 0/2)#mpls ip Router(config-if-GigabitEthernet 0/2)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force
PE2	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/1

	FS(config-if-GigabitEthernet 0/1)#ip address 23.23.23.3 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/1)#ip address 34.34.34.2 255.255.255.0 FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface loopback 0 FS(config-Loopback 0)#ip address 30.30.30.30 255.255.255.255 FS(config-Loopback 0)#exit FS(config)#router ospf 1 Router(config-router)#network 23.23.23.3 255.255.255.0 area 0 Router(config-router)#network 30.30.30.30 255.255.255.255 area 0 Router(config-router)#exit FS(config)#mpls ip FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#mpls ip FS(config-if-GigabitEthernet 0/1)#exit FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/1)#label-switching FS(config-if-GigabitEthernet 0/1)#exit FS(config)#mpls router ldp FS(config-mpls-router)#ldp router-id interface loopback 0 force FS(config-mpls-router)#exit FS(config)#router bgp 100 Router(config-router)#neighbor 10.10.10.10 remote-as 100 Router(config-router)#neighbor 10.10.10.10 update-source loopback 0 Router(config-router)#neighbor 10.10.10.10 send-label Router(config-router)#neighbor 34.34.34.3 remote-as 200 Router(config-router)#neighbor 34.34.34.3 send-label Router(config-router)#exit
PE3	FS#configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)#interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/1)#ip address 34.34.34.3 255.255.255.0

```
FS(config-if-GigabitEthernet 0/1)#exit
FS(config)#interface loopback 0
FS(config-Loopback 0)#ip address 40.40.40.40 255.255.255.255
FS(config-Loopback 0)#exit
FS(config)#interface loopback 1
FS(config-Loopback 0)#ip address 200.200.200.200 255.255.255.255
FS(config-Loopback 0)#exit
FS(config)#mpls ip
FS(config)#interface gigabitEthernet 0/2
FS(config-if-GigabitEthernet 0/2)#label-switching
Router(config-if-GigabitEthernet 0/2)#exit
FS(config)#router bgp 200
Router(config-router)#neighbor 34.34.34.2 remote-as 100
Router(config-router)#neighbor 34.34.34.2 send-label
Router(config-router)# network 200.200.200.200 mask 255.255.255.255
Router(config-router)#exit
FS(config)#interface gigabitEthernet 0/2
FS(config-if-GigabitEthernet 0/2)#bfd interval 50 min_rx 50 multiplier 3
FS(config-if-GigabitEthernet 0/2)#exit
FS(config)#bfd bind bgp-lsp peer-ip 100.100.100.100source-ip 200.200.200.200 local-discriminator 2
remote-discriminator 1
```

Verification

- BFD sessions are in Up state on PE1 and PE2.

PE1

```
FS# show bfd neighbors details
OurAddr          NeighAddr          LD/RD    RH    Holdown(mult)   State    Int
100.100.100.100  200.200.200.200 1/2      1       532 (3 )        Up       Ge0/2
Local Diag: 0, Demand mode: 0, Poll bit: 0
MinTxInt: 200000, MinRxInt: 200000, Multiplier: 3
Received MinRxInt: 50000, Received Multiplier: 3
Holdown (hits): 600(22), Hello (hits): 200(84453)
Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332
Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196
Registered protocols: bgp-lsp
Uptime: 02:18:49
```

	Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24 My Discr.: 2 - Your Discr.: 1 Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0
PE3	FS# show bfd neighbors details OurAddr NeighAddr LD/RD RH Holdown(mult) State Int 200.200.200.200 100.100.100.100 2/1 1 532 (3) Up Ge0/2 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 200000, MinRxInt: 200000, Multiplier: 3 Received MinRxInt: 50000, Received Multiplier: 3 Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332 Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196 Registered protocols: bgp-lsp Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24 My Discr.: 2 - Your Discr.: 1 Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0

Common Errors

- BFD sessions are not configured on the outbound interface of an LSP.
- The interface on which BFD sessions are configured is not the outbound interface of an LSP.
- FTN/ILM entries are not assigned through LDP.
- An LSP is incomplete.
- The global **mpls ip** command is not run.
- The **label-switching** command for enabling MPLS forwarding on an interface is not run.
- The source address and destination address of the BFD-detected devices at both ends of an LSP are not matched.

- The identifiers configured (manually) for the BFD-detected devices at both ends of an LSP are not matched.
- The FEC corresponding to the detected LSP is not routed by the host.

4.5 Monitoring and Maintenance

Displaying

Description	Command
Displays the BFD session state.	**show bfd neighbors**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs MPLS BFD.	**debug mpls bfd**

Chapter6 Multicast Configuration

1. Configuring IP Multicast
2. Configuring IGMP Snooping
3. Configuring IGMP
4. Configuring MLD
5. Configuring PIM-DM
6. Configuring PIM-SM
7. Configuring PIM-SMv6
8. Configuring IGMP Snooping
9. Configuring MLD Snooping
10. Configuring MSDP

1 Configuring IP Multicast

1.1 Overview

IP multicast is abstracted hardware multicasting and an extended multicast routing protocol on the standard IP network layer.

In traditional IP transmission, only one host can send packets to a single host (unicast communication) or all hosts (broadcast communication). However, the multicast technology provides the third choice: a host can send packets to certain specified hosts.

IP multicast is applicable to one-to-many multimedia applications.

1.2 Features

Basic Concepts

PIM Routers and PIM Interfaces

Routers enabled with PIM are called PIM routers. Interfaces enabled with PIM protocol are called PIM interfaces.

Multicast packets are forwarded on PIM routers. The PIM interfaces for receiving multicast packets are called upstream interfaces, and the PIM interfaces for sending multicast packets are called downstream interfaces.

The network segments where upstream interfaces are located are called upstream network segments. The network segments where downstream interfaces are located are called downstream network segments.

PIM Network and PIM Domain

PIM routers are connected through PIM interfaces and form a PIM network.

On certain PIM interfaces, borders are configured to divide a large PIM network into multiple PIM domains. Borders may reject specified multicast packets or limit transmission of PIM messages.

Multicast Distribution Tree, DR and RP

Multicast packets are transmitted from one point to multiple points. The forwarding path is in a tree structure. This forwarding path is called a multicast distribution tree (MDT) and has the following types:

- Rendezvous Point Tree (RPT): The RP is regarded as the root and the designated router (DR) that connects group members is regarded as a leaf.
- Shortest Path Tree (SPT): The DR that connects multicast sources is regarded as the root, and RP or DR that connects group members is regarded as a leaf.

The DR and RP are functional roles for a PIM router.

- The RP collects multicast sources and group member information on the network.
- The DR that connects multicast sources reports multicast source information to the RP. The DR that connects group members reports group member information to the RP.

(*,G) and (S,G)

- (*,G): Packets sent from any source to group G, routing entries corresponding to the packets, and forwarding path (RPT) corresponding to the packets.
- (S,G): Packets sent from source S to group G, routing entries corresponding to the packets, and forwarding path (SPT) corresponding to the packets.

ASM and SSM

PIM-SM supports the following multicast models that are applicable to different multicast address segments:

- Any-Source Multicast (ASM): In the ASM model, user hosts cannot select multicast sources. User hosts join a group and receive packets sent from all sources to the group.
- Source-Specific Multicast (SSM): In the SSM model, user hosts can select multicast sources. User hosts specify source addresses when joining a group and receive only packets sent from specified sources to the group.

SSM model requirements: User hosts must know the multicast source address in advance using other network services so that the hosts can select multicast sources.

Overview

Feature	Description
Configuring Layer-2 Direction Control for Multicast Streams	Allows a specified multicast stream to be configured with multiple commands, that is, to be configured with multiple ports that can forward the stream. Once direction control is configured for a multicast stream, the stream can be forwarded only by these configured interfaces. Other interfaces are not permitted to forward the stream.
Configuring Multicast Non-Stop Forwarding Parameters	During normal running, SSP synchronizes the hardware multicast forwarding table to the management board in real time. After the management board is switched, the command for configuring the multicast control plane of the original slave management board is loaded, and the multicast protocol (such as PIM-SM or IGMP Snooping) re-converges. The multicast non-stop forwarding function ensures continuous forwarding of multicast data streams during re-convergence of the multicast protocol.
Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries	Deletes the earliest hardware entries and adds new entries if the hardware forwarding table overflows when you create multicast forwarding entries.

1.2.1 Configuring Layer-2 Direction Control for Multicast Streams

Configure layer-2 direction control for multicast streams to control the forwarding of multicast streams on an interface.

Working Principle

Configure layer-2 direction control for multicast streams and a forwarding interface so that multicast streams can be forwarded only through configured interfaces. In this case, layer-2 forwarding of multicast streams can be controlled.

Related Configuration

Configuring Layer-2 Direction Control for Multicast Streams

By default, layer-2 direction control for multicast streams is disabled.

Run **ip multicast static** *source-address group-address interface-type interface-number* to configure layer-2 direction control for multicast streams.

1.2.2 Configuring Multicast Non-Stop Forwarding Parameters

The non-stop forwarding function ensures continuous forwarding of multicast data streams during the re-convergence of multicast protocols.

Working Principle

During normal running, SSP synchronizes the hardware multicast forwarding table to the management board in real time. After the management board is switched, the command for configuring the multicast control plane of the original slave management board is loaded, and the multicast protocol (such as PIM-SM or IGMP Snooping) re-converges. The multicast non-stop forwarding function ensures continuous forwarding of multicast data streams during re-convergence of multicast protocols.

After the configured protocol convergence period times out, all multicast forwarding table entries that are not updated during the convergence period are deleted.

Related Configuration

Configuring the Maximum Period for Multicast Protocol Convergence

By default, the maximum period for multicast protocol convergence is 20s.

Run **msf nsf convergence-time** *time* to configure the maximum period for multicast protocol convergence. The value ranges from 0 to 3600s.

A larger value of *time* means a longer maximum period for multicast protocol convergence.

Configuring the Multicast Packet Leakage Period

By default, the multicast packet leakage period is 30s.

Run **msf nsf leak** *interval* to configure the multicast packet leakage period. The value ranges from 0 to 3600s.

A larger value of *interval* means a longer leakage period.

1.2.3 Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries

Delete the earliest hardware entries and adds new entries if the hardware forwarding table overflows when you create multicast forwarding entries.

Working Principle

Delete the earliest hardware entries and adds new entries if the hardware forwarding table overflows when you create multicast forwarding entries .

Related Configuration

Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries

By default, the overwriting mechanism upon the overflow of multicast hardware forwarding entries is disabled.

Run **msf ipmc-overflow override** to configure the overwriting mechanism upon overflow of multicast hardware forwarding entries.

1.3 Configuration

Configuration	Description and Command	
Configuring Basic Functions of IP Multicast	(Mandatory) It is used to configure the multicast service.	
	ip multicast-routing	Enables the IPv4 multicast routing function.
Configuring a TTL Threshold	Optional.	
	ip multicast ttl-threshold *ttl-value*	Configures a TTL threshold for an interface.
Configuring the Number of Entries That Can Be Added to the Multicast Routing Table	**ip multicast route-limit** *limit* [*threshold*]	Limits the number of entries that can be added to the multicast routing table.
Configuring an IP Multicasting Border	**ip multicast boundary** *access-list* [**in** | **out**]	Configures an interface as a multicast border for a specified group.
Configuring an IP Multicasting Static Route	**ip mroute** *source-address mask* { [**bgp** | **isis** | **ospf** | **rip** | **static**] { *v4rpf-addrress* | *interface-type interface-number* } } [*distance*]	Configures an IP multicasting static route.
Configuring Layer-2 Direction Control for Multicast Streams	**ip multicast static** *source-address group-address interface-type interface-number*	Controls the direction of data streams on layer-2 interfaces.
Configuring RPF Route Selection Based on the Longest Match Rule	**ip multicast rpf longest-match**	Configures RPF route selection based on the longest match rule.
Configuring Multicast Non-Stop Forwarding Parameters	**msf nsf convergence-time** *time*	Configures the maximum period for multicast protocol convergence.
	msf nsf leak *time*	Configures the multicast packet leakage period.

Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries	**msf ipmc-overflow override**	Configures the overwriting mechanism upon overflow of multicast hardware forwarding entries.

1.3.1 Configuring Layer-2 Direction Control for Multicast Streams

Configuration Effect

Configure layer-2 direction control for multicast streams to control the forwarding of multicast streams on an interface.

Notes

- The basic functions of IP multicast must be configured.

Configuration Steps

- Layer-2 direction control for multicast streams can be configured on layer-2 devices unless otherwise specified.

Verification

Send multicast packets on the network containing layer-2 device A, connect multiple user hosts to VLAN 1 of layer-2 device A to receive the group, configure layer-2 direction control for multicast streams on device A, and check whether multicast packets are sent to the configured layer-2 interface.

Related Commands

Configuring Layer-2 Direction Control for Multicast Streams

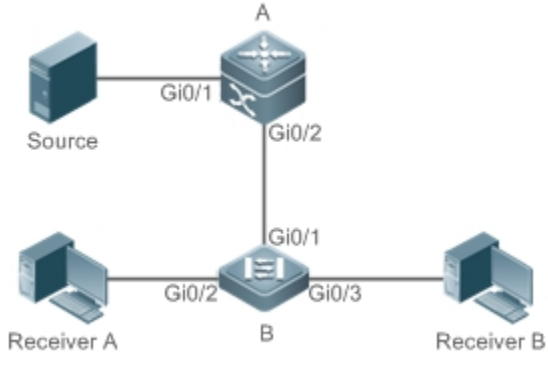
Command	**ip multicast static** *source-address group-address interface-type interface-number*
Parameter Description	*source -address*: Specifies the multicast source address. *group-address*: Specifies the multicast group address. *interface-type interface-number*: Specifies a layer-2 interface that is allowed to forward the multicast flow.
Command Mode	Global configuration mode
Usage Guide	Allow a specified multicast flow to be configured with multiple commands, that is, to be configured with multiple interfaces. Once direction control is configured for a multicast stream, the stream can be forwarded only by these configured interfaces. Other interfaces are not permitted to forward the stream. This command controls only the forwarding of multicast streams on the interface, but does not directly affect the processing of multicast protocols on the protocol packets. However, since certain features of the multicast protocol are driven by multicast data streams, behaviors of the multicast routing protocols may also be affected.

Configuration Example

Creating the IP Multicast Service on the IPv4 Network and Configuring Layer-2 Direction Control for Multicast Streams

Scenario

Figure 1- 1



Configuration Steps	● Configure the basic functions of IP multicast. (Omitted) ● Configure layer-2 direction control for multicast streams on device B so that the streams are sent only to the Gi 0/2 interface.
B	A# configure terminal A(config)# ip multicast static 192.168.1.100 233.3.3.3 gigabitEthernet0/2
Verification	Enable the multicast source (192.168.1.100) to send packets to G (233.3.3.1). Enable receivers A and B to join G. ● Check multicast packets received by receiver A. Receiver B should not be able to receive multicast packets from G.

Common Errors

- An IPv4 unicast route is incorrectly configured.

1.3.2 Configuring Multicast Non-Stop Forwarding Parameters

Configuration Effect

- The non-stop forwarding function ensures continuous forwarding of multicast data streams during re-convergence of multicast protocols.

Notes

- The basic functions of IP multicast must be configured.

Configuration Steps

Configuring the Maximum Period for Multicast Protocol Convergence

- The maximum period for multicast protocol convergence can be specified on each device unless otherwise specified.

Configuring the Multicast Packet Leakage Period

- The multicast leakage period can be configured on each device unless otherwise specified.

Verification

Run **show msf nsf** to check the configured multicast non-stop forwarding parameters.

Related Commands

Configuring the Maximum Period for Multicast Protocol Convergence

Command	**msf nsf convergence-time** *time*
Parameter Description	**convergence-time** *time*: Specifies the maximum period for multicast protocol convergence. The value ranges from 0 to 3600s. The default value is 20s.
Command Mode	Global configuration mode
Usage Guide	-

Configuring the Multicast Packet Leakage Period

Command	**msf nsf leak** *interval*
Parameter Description	**leak** *intervald*: Specifies the multicast packet leakage period. The value ranges from 0 to 3600s. The default value is 30s.
Command Mode	Global configuration mode
Usage Guide	-

Displaying Multicast Non-Stop Forwarding Configurations

Command	**show msf nsf**
Parameter Description	**-**
Command Mode	Privilege, global and interface configuration modes
Usage Guide	-

Configuration Example

Creating the IP Multicast Service on the IPv4 Network and Supporting PIM-DM

Scenario	Basic environment of the IP multicast service (Omitted)
Configuration Steps	● Configure the basic functions of IP multicast. (Omitted) ● Configure the maximum period for multicast protocol convergence. ● Configure the multicast packet leakage period.
A	A# configure terminal A(config)# msf nsf convergence-time 200 A(config)# msf nsf leak 300
Verification	Run **show msf nsf** to display multicast non-stop forwarding configurations.
A	A# show msf nsf Multicast HA Parameters ----------+------+ protocol convergence timeout 200 secs flow leak interval 300 secs

1.3.3 Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries

Configuration Effect

- Delete the earliest hardware entries and adds new entries if the hardware forwarding table overflows when you create multicast forwarding entries.

Notes

- The basic functions of IP multicast must be configured.

Configuration Steps

- The overwriting mechanism upon overflow of multicast hardware forwarding entries can be configured on each device unless otherwise specified.

Verification

Run **show running-config** to check whether the overwriting mechanism upon overflow of multicast hardware forwarding entries is configured.

Related Commands

Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries

Command	**msf ipmc-overflow override**
Parameter Description	**-**

Command Mode	Global configuration mode
Usage Guide	-

Configuration Example

Creating the IP Multicast Service on the IPv4 Network and Configuring an Overwriting Mechanism Upon Overflow of Multicast Hardware Forwarding Entries

Scenario	Basic environment of the IP multicast service (Omitted)
Configuration Steps	● Configure the basic functions of IP multicast. (Omitted) ● Configure the overwriting mechanism upon overflow of multicast hardware forwarding entries.
A	A# configure terminal A(config)#msf ipmc-overflow override
Verification	Run **show running-config** to check whether the overwriting mechanism upon overflow of multicast hardware forwarding entries is configured.
A	A# show running-config … msf ipmc-overflow override …

1.4 Monitoring

Displaying

Description	Command
Displays the IPv4 multi-layer multicast forwarding table.	**show msf msc**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the processing of IPv4 multi-layer multicast packet forwarding.	**debug msf forwarding**
Debugs the operation on multi-layer multicast forwarding entries on an IPv4 network.	**debug msf mfc**
Debugs the bottom-layer hardware processing of IPv4 multi-layer multicast packet forwarding.	**debug msf ssp**
Debugs the invocation of API interfaces provided by IPv4 multi-layer multicast forwarding.	**debug msf api**
Debugs the processing of multi-layer multicast forwarding events on an IPv4	**debug msf event**

Description	Command
network.	

2 Configuring IPv6 Multicast

2.1 Overview

IPv6 multicast is enrichment and enhancement of IPv4 multicast. In comparison with IPv4 multicast, the IPv6 multicast address mechanism is greatly enriched.

In traditional IP transmission, a host is allowed to send packets only to a single host (unicast communication) or all hosts (broadcast communication). The multicast technology provides a third choice: A host is allowed to send packets to certain hosts.

The IP multicast technology is applicable to one-to-many multimedia applications.

Protocols and Standards

IPv6 multicast covers the following protocols:

- Multicast Listener Discovery (MLD): Runs between a multicast device and a host, and tracks and learns relationships of group members.
- Protocol Independent Multicast – Spare Mode for IPv6 (PIM-SMv6): Runs between devices and implements multicast packet forwarding by establishing a multicast routing table.

2.2 Applications

Application	Description
Typical Application of PIM-SMv6	The PIM-SMv6 multicast service is provided in the same network.

2.2.1 Typical Application of PIM-SMv6

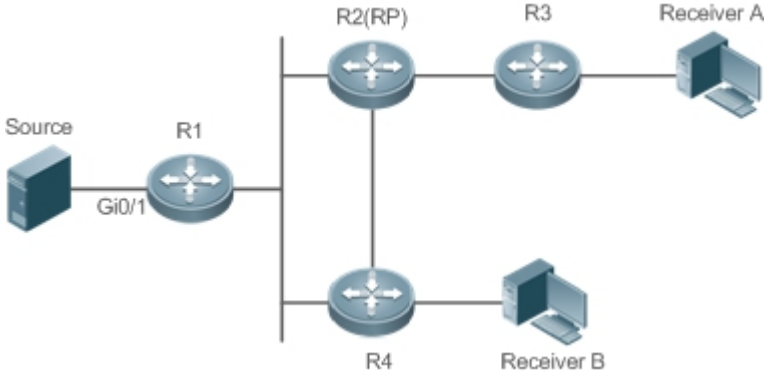
Scenario

The PIM-SMv6 multicast service is provided in the same network.

As shown in the following figure:

- R1 and the multicast source are in the same network, R2 is configured as a rendezvous point (RP), R3 is in the same network as Receiver A, and R4 is in the same network as Receiver B. Assume that devices and hosts are correctly connected, IPv6 is enabled on each interface, and IPv6 unicast is enabled on each device.

Figure 2- 1



Remarks	R1, R2, R3, and R4 are Layer-3 devices and R2 functions as an RP. The multicast source is directly connected to R1, Receiver A is directly connected to R3, and Receiver B is directly connected to R4.

Deployment

- Run the Open Shortest Path First for IPv6 (OSPFv6) protocol in the same network to implement unicast routing.
- Run the PIM-SMv6 protocol in the same network to implement multicast routing.

2.3 Features

Basic Concepts

PIM Router and PIM Interface

Routers where the PIM protocol is enabled are called PIM routers. Interfaces where the PIM protocol is enabled are called PIM interfaces.

Multicast packets are forwarded by PIM routers. The PIM interfaces for receiving multicast packets are called upstream interfaces, and the PIM interfaces for transmitting multicast packets are called downstream interfaces.

Network segments where upstream interfaces are located are called upstream network segments. Network segments where downstream interfaces are located are called downstream network segments.

PIM Network and PIM Domain

PIM routers are connected through PIM interfaces and form a PIM network.

On some PIM interfaces, borders are set to divide a large PIM network into multiple PIM domains. The borders may reject specific multicast packets or limit transmission of PIM messages.

Multicast Distribution Tree, DR, RP

Multicast packets are transmitted from one point to multiple points. The forwarding path presents a tree structure. This forwarding path is called a multicast distribution tree (MDT). MDTs are classified into two types:

- Rendezvous point tree (RPT): Uses the rendezvous point (RP) as the root and designated routers (DRs) connected to group members as leaves.
- Shortest path tree (SPT): Use the DR connected to a multicast source as the root and the RPs or DRs connected to group members as leaves.

DRs and RPs are function roles of PIM routers.

- RPs collect information about multicast sources and group members in the network.
- The DR connected to a multicast source reports multicast source information to the RP and the DRs connected to group members report the group member information to the RP.

(*,G), (S,G)

- (*,G): Indicates the packets transmitted from any source to Group G, routing entries corresponding to the packets, and forwarding path (RPT) corresponding to the packets.

- (S,G): Indicates the packets transmitted from Source S to Group G, routing entries corresponding to the packets, and forwarding path (SPT) corresponding to the packets.

ASM, SSM

PIM-SM supports two multicast service models: any-source multicast (ASM) and source-specific multicast (SSM), which are applicable to different multicast address segments.

- ASM: In the ASM model, a user host cannot select a multicast source. The user host joins a multicast group and receives all packets sent from all sources to the multicast group.
- SSM: In the SSM model, a user host can select a multicast source. The user host specifies the source address when joining a multicast group, and then receives packets only from the specified source to the multicast group.

SSM model requirement: Other network services must be used to enable a user host to know the position of a multicast source in advance so that the user host selects the multicast source.

Overview

Feature	Description
Configuring IPv6 Multicast Basic Functions	Creates a PIM network to provide the IPv6 multicast service for data sources and user terminals in the network.
Configuring the Number of Entries That Can Be Added to the IPv6 Multicast Routing Table	Restricts the number of entries that can be added to the multicast routing table.
Configuring the IPv6 Multicast Border	Sets an interface as the multicast border of a specific group range.
Configuring IPv6 Multicast Static Routing	Configures multicast static routing to adopt multicast forwarding paths different from unicast forwarding paths.
Configuring Layer-2 Flow Direction Control for Multicast Streams	Multiple commands can be configured for a multicast stream, that is, multiple ports can be allowed to forward the multicast stream. If flow direction control is configured for a multicast stream, the multicast stream can be forwarded only by the configured ports. Other ports are not allowed to forward the multicast stream.
Configuring RPF Route Selection According to the Longest Matching Principle	One optimal route is selected from each of the multicast static routing table, MBGP routing table, and unicast routing table according to RPF rules. Among the three optimal routes, the route with the longest subnet mask matching is selected as the RPF route.

2.3.1 Configuring IPv6 Multicast Basic Functions

Create a PIM network to provide the IPv6 multicast service for data sources and user terminals in the network.

Working Principle

A device maintains the routing table used for multicast packet forwarding over an IPv6 multicast routing protocol (such as PIM-SMv6), and learns information about the status of group members in the directly-connected network segments over the MLDv1/v2 protocol. A host joins a specific IPv6 multicast group by transmitting the MLD REPORT message.

Related Configuration

Enabling the IPv6 Multicast Routing Function

The IPv6 multicast routing function is disabled by default.

Run the **ipv6 multicast-routing** command to enable the IPv6 multicast routing function.

Configuring an IP Multicast Protocol on an Interface

The IPv6 multicast protocol is disabled on an interface by default.

Run the **ipv6 pim dense-mode** command to enable the IPv6 multicast protocol on an interface.

2.3.2 Configuring the Number of Entries That Can Be Added to the IPv6 Multicast Routing Table

Every multicast data packet received by the device is used to maintain relevant IPv6 multicast routing entries. Excessive multicast routing entries, however, may deplete the device memory and degrade the device performance. Users can restrict the number of entries in the IPv6 multicast routing table based on the actual networking conditions and service performance requirements.

Working Principle

Restrict the number of entries in the IPv6 multicast routing table based on the actual networking conditions and service performance requirements, so as to sustain the device performance.

Related Configuration

Configuring the Number of Entries That Can Be Added to the IPv6 Multicast Routing Table

By default, 1,024 entries can be added to the IP multicast routing table.

Run the **ipv6 multicast route-limit** *limit* [*threshold*] command to adjust the number of entries that can be added to the IPv6 multicast routing table. The value ranges from 1 to 65,536.

A larger value of *limit* means that more entries can be added to the IPv6 multicast routing table, and a smaller value of *limit* means that fewer entries can be added to the IPv6 multicast routing table.

2.3.3 Configuring the IPv6 Multicast Border

Configure the IPv6 multicast border to restrict the transmission scope of multicast packets.

Working Principle

Configure the multicast border to specify the transmission scope of multicast packets. When the multicast forwarding border is configured on an interface, multicast packets including multicast packets sent by the local device cannot be forwarded or received by this interface.

Related Configuration

Configuring the IPv6 Multicast Border

No multicast border is configured by default.

Run the **ipv6 multicast boundary** *access-list-name* [**in** | **out**] command to configure the multicast border.

2.3.4 Configuring IPv6 Multicast Static Routing

Configure IPv6 multicast static routing to specify a reverse path forwarding (RPF) interface or RPF neighbor for multicast packets from a specific multicast source.

Working Principle

The RPF check is conducted during forwarding of multicast packets. IPv6 multicast static routing can be configured to specify an RPF interface or RPF neighbor for multicast packets from a specific multicast source.

Related Configuration

Configuring IPv6 Multicast Static Routing

No multicast static routing is configured by default.

Run the **ipv6 mroute** *ipv6-prefix/prefix-length* [**bgp** | **isis** | **ospfv3** | **ripng** | **static**] { *ipv6-prefix* | *interface-type interface-number* } [*distance*] command to configure IPv6 multicast static routing.

2.3.5 Configuring Forced Forwarding of Multicast Packets by Software

IPv6 multicast data packets destined for the CPU are forcedly forwarded by software.

Working Principle

After configuring this function, all IPv6 multicast data packets destined for the CPU are forcedly forwarded by software.

Related Configuration

Configuring Forced Forwarding of CPU-destined IPv6 Multicast Data Packets by Software

This function is disabled by default.

Run **msf force-forwarding** to enable IPv6 multicast data packets destined for the CPU to be forcedly forwarded by software.

2.3.6 Configuring Layer-2 Flow Direction Control for Multicast Streams

Configure Layer-2 flow direction control for multicast streams to control the forwarding behavior of multicast streams on ports.

Working Principle

Configure Layer-2 flow direction control for multicast streams to configure the ports that are allowed to forward multicast streams. Then, multicast streams are forwarded only by the configured ports, thereby controlling Layer-2 forwarding of multicast streams.

Related Configuration

Configuring Layer-2 Flow Direction Control for Multicast Streams

Layer-2 flow direction control is disabled for multicast streams by default.

Run the **ipv6 multicast static** *source-address group-address interface-type interface-number* command to configure the Layer-2 flow direction control for multicast streams.

2.3.7 Configuring RPF Route Selection According to the Longest Matching Principle

Among the three optimal routes selected from the multicast static routing table, Multiprotocol Border Gateway Protocol (MBGP) routing table, and unicast routing table, select the optimal route with the longest subnet mask matching as the RPF route.

Working Principle

According to RPF rules, select a multicast static route, MBGP route, and unicast route used for the RPF check respectively from the multicast static routing table, MBGP routing table, and unicast routing table.

- If route selection according to the longest matching principle is configured, the route with the longest subnet mask matching is selected out of the three routes as the RPF route. If the three routes share the same subnet mask, the route with the highest priority is selected. If the three routes have the same priority, the RPF route is selected according to the sequence of multicast static route, MBGP route, and unicast route.

- If route selection according to the longest matching principle is not configured, the route with the highest priority is selected. If the three routes have the same priority, the RPF route is selected according to the sequence of multicast static route, MBGP route, and unicast route.

Related Configuration

Configuring RPF Route Selection According to the Longest Matching Principle

A route with the highest priority is selected as the RPF route by default. If the routes have the same priority, the RPF route is selected according to the sequence of multicast static route, MBGP route, and unicast route.

Run the **ipv6 multicast rpf longest-match** command to configure RPF route selection according to the longest matching principle.

2.4 Configuration

Configuration	Description and Command	
Configuring IPv6 Multicast Basic Functions	(Mandatory) It is used to create a multicast service.	
	ipv6 multicast-routing	Enables the IPv6 multicast routing function.
Configuring the Number of Entries That Can Be Added to the IPv6 Multicast Routing Table	Optional.	
	ipv6 multicast route-limit *limit* [*threshold*]	Restricts the number of entries that can be added to the multicast routing table.
Configuring the IPv6 Multicast Border	**ipv6 multicast boundary** *access-list-name* [**in** | **out**]	Sets an interface as the multicast border of a specific group range.
Configuring IPv6 Multicast Static Routing	**ipv6 mroute** *ipv6-prefix/prefix-length* [*protocol*] { *v6rpf-addrress* | *interface-type interface-number* } [*distance*]	Configures IPv6 multicast static routing.
Configuring Forced Forwarding of Multicast Packets by Software	**msf6 force-forwarding**	Configures forced forwarding of multicast packets by software.
Configuring Layer-2 Flow Direction Control for Multicast Streams	**ipv6 multicast static** *source-address group-address interface-type interface-number*	Controls the flow direction of data streams on Layer-2 ports.
Configuring RPF Route Selection According to the Longest Matching Principle	**ipv6 multicast rpf longest-match**	Configures RPF route selection according to the longest matching principle.

2.4.1 Configuring IPv6 Multicast Basic Functions

Configuration Effect

- Create a PIM network to provide the IPv6 multicast service for data sources and user terminals in the network.

Notes

- The PIM network needs to use existing unicast routing in the network. Therefore, IPv6 unicast routing must be configured in the network.

Configuration Steps

Enabling the IPv6 Multicast Routing Function

- Mandatory.
- Enable the IPv6 multicast routing function on each router unless otherwise specified.

Enabling an IP Multicast Protocol on Interfaces

- Mandatory.
- Enable the IPv6 multicast protocol function on interfaces unless otherwise specified.

Verification

Make multicast sources in the network send multicast packets and make a user host join the groups.

- Check whether the user host can successfully receive packets from each group.

Related Commands

Enabling the IPv6 Multicast Routing Function

Command	**ipv6 multicast-routing**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	The IPv6 multicast routing function must be enabled before various IPv6 multicast protocols are enabled. The IPv6 multicast routing function and the MLD snooping function are mutually exclusive.

Configuring IPv6 Multicast Protocols

For details about the MLD configuration method, see the *Configuring MLD*.

For details about the PIM-SMv6 configuration method, see the *Configuring PIM-SMv6*.

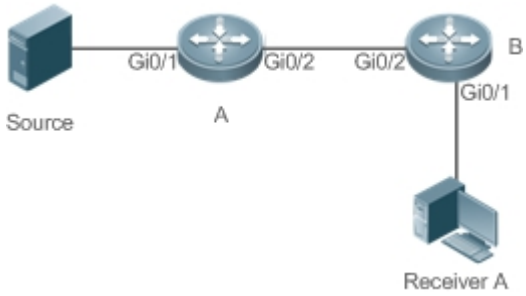
After the Layer-3 multicast function is enabled on a private VLAN and Super VLAN, if there is a multicast source in the sub-VLAN, an entry needs to be additionally copied, with the inlet of the sub-VLAN where multicast streams enter because the validity check needs to be conducted at the inlet during multicast packet forwarding. As a result, one more multicast hardware entry is occupied, and the multicast capacity needs to be decreased by one.

Displaying Multicast Forwarding Table Information

Command	**show ipv6 mroute** [*group-or-source-address* [*group-or-source-address*]] [**sparse**] [**summary** | **count**]
Parameter Description	*group-or-source-address*: Indicates the group address or source address. *group-or-source-address*: Indicates the group address or source address. **sparse**: Displays the core entry of the PIM-SMv6 multicast routing table. **summary**: Displays the summary of IPv6 multicast routing entries. **count**: Displays the count information about IPv6 multicast routing entries.
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	N/A

Configuration Example

Creating the IPv6 Multicast Service on an IPv6 Network to Support PIMv6-SM

Scenario **Figure 2- 2**	
Configuration Steps	● Configure an IPv6 unicast routing protocol (for example, OSPFv3) on routers. ● Enable the IPv6 multicast routing function on all routers. ● Enable the PIMv6-SM function on device interconnection interfaces, interface for connecting to the user host, and interface for connecting to the multicast source.
A	A# configure terminal A(config)# ipv6 multicast-routing A(config)# interface gigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ipv6 pim sparse-mode A(config-if)# exit A(config)# interface gigabitEthernet 0/2 A(config-if-GigabitEthernet 0/2)#ipv6 pim sparse-mode A(config-if)# exit
B	B# configure terminal B(config)# ipv6 multicast-routing B(config)# interface gigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)#ipv6 pim sparse-mode B(config-if)# exit B(config)# interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)#ipv6 pim sparse-mode B(config-if)# exit
Verification	Make Multicast Source (2001::1) send packets to G(ff16::16) and make Receiver A join G. ● Check multicast packets received by Receiver A. Receiver A should be able to receive multicast packets from G. ● Check the multicast forwarding table on Receiver A and Device B.
A	A# show ipv6 mroute

	IPv6 Multicast Routing Table Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed, R - RPT, S - SPT, s - SSM Group Timers: Uptime/Stat Expiry Interface State: Interface (2001::1, ff16::16), uptime 00:03:12, stat expires 00:02:03 Owner PIMSMV6, Flags: TFS Incoming interface: GigabitEthernet 0/1 Outgoing interface list: GigabitEthernet 0/2
B	B# show ipv6 mroute IPv6 Multicast Routing Table Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed, R - RPT, S - SPT, s - SSM Group Timers: Uptime/Stat Expiry Interface State: Interface (2001::1, ff16::16), uptime 00:00:23, stat expires 00:03:07 Owner PIMSMV6, Flags: TFR Incoming interface: GigabitEthernet 0/2 Outgoing interface list: GigabitEthernet 0/1

Common Errors

- IPv6 unicast routing is incorrectly configured.
- IPv6 multicast routing is not enabled on a router.
- No IPv6 multicast protocol is enabled on an interface.

2.4.2 Configuring the Number of Entries That Can Be Added to the IPv6 Multicast Routing Table

Configuration Effect

- Every multicast data packet received by the device is used to maintain relevant IPv6 multicast routing entries. Excessive multicast routing entries, however, may deplete the device memory and degrade the device performance. Users can restrict the number of entries in the IPv6 multicast routing table based on the actual networking conditions and service performance requirements.

Notes

- The IPv6 multicast basic functions must be configured.

Configuration Steps

- Restrict the number of entries in the IPv6 multicast routing table based on the actual networking conditions and service performance requirements.

Verification

Make multicast sources in the network send multicast packets to N different multicast groups and make a user host join these groups. Set the number of entries that can be added to the IPv6 multicast routing table to N-1 on the device and check that multicast packets received by the user host are from N-1 groups.

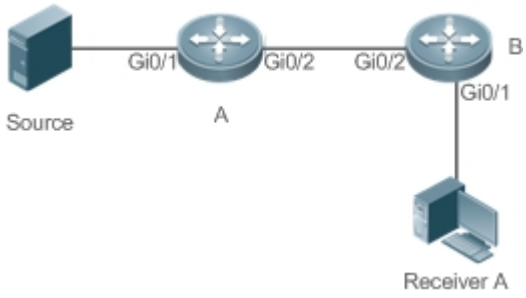
Related Commands

Configuring the Number of Entries That Can Be Added to the IP Multicast Routing Table

Command	**ipv6 multicast route-limit** *limit* [*threshold*]
Parameter Description	*limit*: Indicates the number of multicast routing entries. The value ranges from 1 to 65,536 and the default value is 1,024. *threshold*: Indicates the multicast routing entry quantity for triggering an alarm. The default value is 65,536.
Command Mode	Global configuration mode
Usage Guide	Routing entries that are beyond the allowable range of hardware can be forwarded only by software due to hardware resource restrictions, making the performance deteriorate.

Configuration Example

Creating the IPv6 Multicast Service on an IPv6 Network and Configuring the Number of Entries That Can Be Added to the IPv6 Multicast Routing Table

Scenario **Figure 2- 3**	
Configuration Steps	- Configure IP multicast basic functions (omitted). - Set the number of entries that can be added to the IP multicast routing table to 2 on Device B.
B	B# configure terminal B(config)# ipv6 multicast route-limit 2

Verification	Make Multicast Source (2001: : 1) send packets to G1(ff16::16), G2(ff16::17), and G3(ff16::18) and make Receiver A join G1, G2, and G3. ● Check multicast packets received by Receiver A. Receiver A should be able to receive multicast packets from two groups of G1, G2, and G3. ● Check multicast routing entries on Receiver A and Device B. ● A prompt is displayed when the number of entries in the multicast routing table reaches the upper limit.
A	A# show ipv6 mroute IPv6 Multicast Routing Table Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed, R - RPT, S - SPT, s - SSM Group Timers: Uptime/Stat Expiry Interface State: Interface (2001::1, ff16::16), uptime 00:01:01, stat expires 00:02:29 Owner PIMSMV6, Flags: TFS Incoming interface: GigabitEthernet 0/1 Outgoing interface list: GigabitEthernet 0/2 (2001::1, ff16::17), uptime 00:01:01, stat expires 00:02:29 Owner PIMSMV6, Flags: TFS Incoming interface: GigabitEthernet 0/1 Outgoing interface list: GigabitEthernet 0/2 (2001::1, ff16::18), uptime 00:00:57, stat expires 00:02:33 Owner PIMSMV6, Flags: TFS Incoming interface: GigabitEthernet 0/1 Outgoing interface list: GigabitEthernet 0/2
B	B# show ipv6 mroute

	IPv6 Multicast Routing Table Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed, R - RPT, S - SPT, s - SSM Group Timers: Uptime/Stat Expiry Interface State: Interface (2001::1, ff16::16), uptime 00:00:29, stat expires 00:03:01 Owner PIMSMV6, Flags: TFR Incoming interface: GigabitEthernet 0/2 Outgoing interface list: GigabitEthernet 0/1 (2001::1, ff16::17), uptime 00:00:29, stat expires 00:03:01 Owner PIMSMV6, Flags: TFR Incoming interface: GigabitEthernet 0/2 Outgoing interface list: GigabitEthernet 0/1
	A prompt is displayed when the number of entries in the multicast routing table reaches the upper limit. B#* Jan 3 21:40:07: %MROUTE-4-ROUTELIMIT: IPv6 Multicast route limit 2 exceeded..

Common Errors

- IPv6 unicast routing is incorrectly configured.

2.4.3 Configuring the IPv6 Multicast Border

Configuration Effect

- Configure the IPv6 multicast border to restrict the transmission scope of multicast packets.

Notes

- The IPv6 multicast basic functions must be configured.

Configuration Steps

- Configure the IPv6 multicast border on each PIM router interface unless otherwise specified.

Verification

Make multicast sources send multicast packets to multicast groups and make a user host join these multicast groups. Configure the IPv6 multicast border on the PIM router interface connected to the user host and check whether the user host can receive multicast packets.

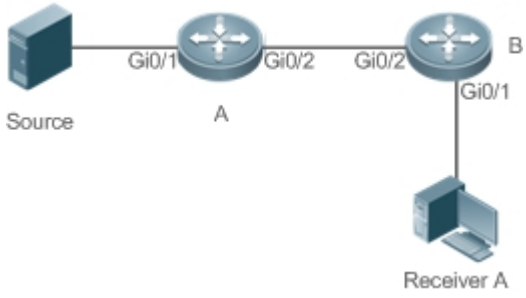
Related Commands

Enabling the IPv6 Multicast Routing Function

Command	**ipv6 multicast boundary** *access-list-name* [**in** | **out**]
Parameter Description	*access-list-name*: Uses the group address range defined by an access control list (ACL). **in**: Indicates that the multicast border takes effect in the incoming direction of multicast streams. **out**: Indicates that the multicast border takes effect in the outgoing direction of multicast streams.
Command Mode	Interface configuration mode
Usage Guide	The ACL referenced in this command can be a standard ACL or an extended ACL. If an extended ACL is used, only destination addresses need to be matched. This command can be used to filter MLD and PIM-SMv6 protocol packets relevant to the IPv6 multicast group range. Multicast data streams are not transmitted or received by multicast border interfaces.

Configuration Example

Creating the IPv6 Multicast Service on an IPv6 Network and Configuring the IPv6 Multicast Border

Scenario **Figure 2-4**	
Configuration Steps	● Configure IP multicast basic functions (omitted). ● Configure an ACL on Device A. ● Configure the IP multicast border on Interface Gi0/1 of Device A.
A	A# configure terminal A(config)# ipv6 access-list ip_multicast A(config-ipv6-acl)#deny udp any any A(config-ipv6-acl)#exit A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ip multicast boundary ip_multicast A A(config-if-GigabitEthernet 0/1)# exit
Verification	Make Multicast Source (192.168.1.100) send packets to G (233.3.3.3) and make Receiver A join G. ● Run the **debug ipv6 pim sparse-mode events** command to debug multicast events in SM mode.

A	A# debug ipv6 pim sparse-mode events Dec 28 11:54:07: %7: No cache message: src 2001::1 for ff16::16 vif 1 *Dec 28 11:54:07: %7: Ignore No cache message: src 2001::1 for ff16::16 vif 1 in PIM6_BOUNDARY_FLT_BOTH range

Common Errors

- IPv6 unicast routing is incorrectly configured.

2.4.4 Configuring IPv6 Multicast Static Routing

Configuration Effect

- Configure IPv6 multicast static routing to specify an RPF interface or RPF neighbor for multicast packets from a specific multicast source.

Notes

- The IPv6 multicast basic functions must be configured.

Configuration Steps

- Configure IPv6 multicast static routing on each device unless otherwise specified.

Verification

Configure IPv6 multicast static routing and then run the **show ipv6 rpf** *v6source-address* command to check RPF information about a specific multicast source.

Related Commands

Configuring IPv6 Multicast Static Routing

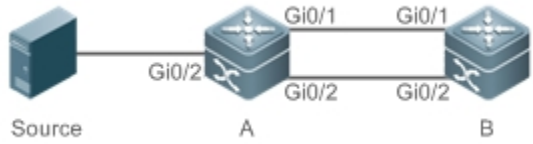
Command	**ipv6 mroute** *ipv6-prefix/prefix-length* [*protocol*] { *v6rpf-addrress* | *interface-type interface-number* } [*distance*]
Parameter Description	*ipv6-prefix*: Indicates the IPv6 address of a multicast source. *prefix-length*: Indicates the subnet mask of the IPv6 address of the multicast source. **fallback-lookup** { **global** | **vrf** *vrf-name* }: Specifies the VRF used for RPF search. *protocol*: Indicates the unicast routing protocol that is being used currently. *v6rpf-addrress*: Indicates the IPv6 address of the RPF neighbor (next hop to the multicast source). *interface-type interface-number*: Indicates the RPF interface (outbound interface to the multicast source). *distance*: Indicates the route management distance. The value ranges from 0 to 255 and the default value is 0.
Command Mode	Global configuration mode
Usage Guide	IPv6 multicast static routing is used only for the RPF check. To specify the outbound interface rather than the next-hop IP address of IPv6 static multicast routing, the outbound interface must be of the point-to-point type.

Displaying RPF Information About a Specific Source Address

Command	**show ipv6 rpf** *v6source-address*
Parameter Description	*v6source-address*: Indicates the IPv6 source address.
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	N/A

Configuration Example

Creating the IPv6 Multicast Service on an IPv6 Network and Configuring IPv6 Multicast Static Routing

Scenario Figure 2- 5	
Configuration Steps	● Configure IPv6 multicast basic functions (omitted). ● Configure a static route to the receiver on Device B.
A	B# configure terminal B(config)# ipv6 mroute 2005::/64 ospfv3 2002::2
Verification	Run the **show ipv6 rpf** command to display the RPF information received by the receiver before and after configuration.
Before Configuration	B# show ipv6 rpf 2005::1 RPF information for 2005::1 RPF interface: GigabitEthernet 0/1 RPF neighbor: fe80::2d0:f8ff:fe22:341b RPF route: 2005::1/128 RPF type: unicast (ospf) RPF recursion count: 0 Doing distance-preferred lookups across tables Distance: 110 Metric: 1
After Configuration	B# show ipv6 rpf 2005::1 RPF information for 2005::1 RPF interface: GigabitEthernet 0/2 RPF neighbor: 2002::2

	RPF route: 2005::/64 RPF type: unicast (ospf) RPF recursion count: 0 Doing distance-preferred lookups across tables Distance: 110 Metric: 1

Common Errors

- IPv6 unicast routing is incorrectly configured.
- IPv6 multicast routing is not enabled on a router.

2.4.5 Configuring Forced Forwarding of Multicast Packets by Software

Configuration Effect

- After configuring this function, all IPv6 multicast data packets destined for the CPU are forcedly forwarded by software.

Notes

- The basic functions of IPv6 multicasting must be configured.

Configuration Steps

- Configure forced forwarding of multicast packets by software on each device unless otherwise specified.

Verification

Run **show running-config** to check whether forced forwarding of multicast packets by software is configured.

Related Commands

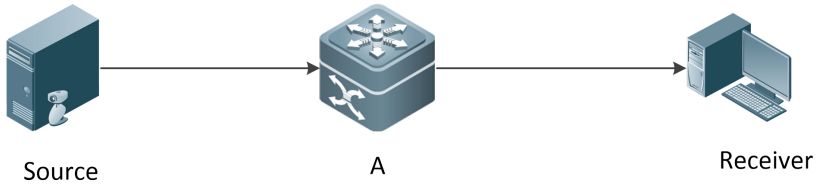
Configuring Forced Forwarding of Multicast Packets by Software

Command	**msf6 force-forwarding**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Only configuration related to IPv6 multicasting is described.

Creating the IPv6 Multicast Service on the IPv6 Network and Configuring Forced Forwarding of Multicast Packets by Software

Scenario	Basic environment for the IPv6 multicast service
Figure 2- 6	
Configuration Steps	● Configure the basic functions of IPv6 multicasting. ● Configure forced forwarding of multicast packets by software.
A	A# configure terminal A(config)#msf6 force-forwarding
Verification	Run **show running-config** to check whether forced forwarding of multicast packets by software is configured.
A	A# show running-config … Msf6 force-forwarding …

2.4.6 Configuring Layer-2 Flow Direction Control for Multicast Streams

Configuration Effect

Configure Layer-2 flow direction control for multicast streams to control the forwarding behavior of multicast streams on ports.

Notes

- The IPv6 multicast basic functions must be configured.

Configuration Steps

- Configure Layer-2 flow direction control for multicast streams on devices unless otherwise specified.

Verification

Make Device A send multicast packets to multicast groups in the network. Multiple user hosts connected to VLAN 1 of Device A receive multicast packets from these multicast groups. Configure Layer-2 flow direction control for multicast streams on Device A so that multicast packets are sent to configured ports.

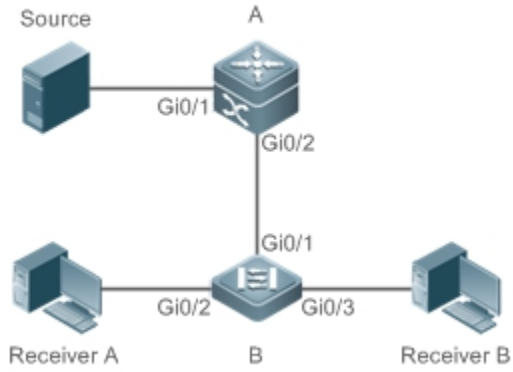
Related Commands

Configuring Layer-2 Flow Direction Control for Multicast Streams

Command	**ipv6 multicast static** *source-address group-address interface-type interface-number*
Parameter Description	*source -address*: Indicates the multicast source address. *group-address*: Indicates the multicast group address. *interface-type interface-number*: Indicates a Layer-2 port that is allowed to forward multicast streams.
Command Mode	Global configuration mode
Usage Guide	Multiple commands can be configured for a multicast stream, that is, multiple ports can be allowed to forward the multicast stream. If flow direction control is configured for a multicast stream, the multicast stream can be forwarded only by the configured ports. Other ports are not allowed to forward the multicast stream. This command controls only the forwarding behavior of multicast streams on ports. It does not directly affect processing of protocol packets by multicast protocols. Some features of multicast protocols (such as PIM-SMv6) are driven by multicast data streams, and therefore, the behavior of the multicast routing protocols may still be affected.

Configuration Example

Creating the IPv6 Multicast Service on an IPv6 Network and Configuring Layer-2 Flow Direction Control for Multicast Streams

Scenario **Figure 2- 7**	
Configuration Steps	● Configure IP multicast basic functions (omitted). ● Configure Layer-2 flow direction control for multicast streams on Device B so that multicast streams are transmitted only to Interface Gi0/2.
B	A# configure terminal A(config)# ipv6 multicast static 2001::1 ff16::16 gigabitEthernet 0/2
Verification	Make Multicast Source (2001: : 1) send packets to G (ff16::16) and make Receiver A and Receive B join G. ● Receiver A should be able to receive multicast packets from G but Receiver B cannot receive multicast packets from G.

Common Errors

- IPv6 unicast routing is incorrectly configured.

2.4.7 Configuring RPF Route Selection According to the Longest Matching Principle

Configuration Effect

Among the three optimal routes selected from the multicast static routing table, MBGP routing table, and unicast routing table, select the optimal route with the longest subnet mask matching as the RPF route.

Notes

- The IP multicast basic functions must be configured.

Configuration Steps

- Configure RPF route selection according to the longest matching principle on each device unless otherwise specified.

Verification

Configure a multicast static route and a unicast static route with the same priority and configure the unicast static route to have the longest subnet mask matching.

- Run the **show ipv6 rpf** *v6source-address* command to check RPF information about a specific source.

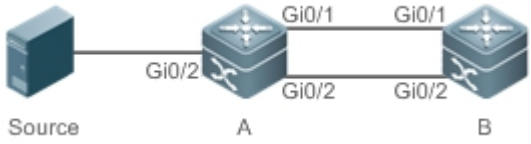
Related Commands

Configuring RPF Route Selection According to the Longest Matching Principle

Command	**ipv6 multicast rpf longest-match**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	The steps of selecting an RPF route are as follows: Select one optimal route used for the RPF check from each of the IPv6 multicast static routing table, IPv6 MBGP routing table, and IPv6 unicast routing table. Select one route out of the three optimal routes as the RPF route. If the command for selecting the RPF route according to the longest matching principle is configured, the route with the longest subnet mask matching is selected out of the three optimal routes as the RPF route. If the three routes share the same subnet mask, the route with the highest priority is selected. If the routes have the same priority, the RPF route is selected according to the sequence of IPv6 multicast static route, IPv6 MBGP route, and IPv6 unicast route. If the command for selecting the RPF route according to the longest matching principle is not configured, the route with the highest priority is selected out of the three optimal routes as the RPF route. If the routes have the same priority, the RPF route is selected according to the sequence of IPv6 multicast static route, IPv6 MBGP route, and IPv6 unicast route.

Configuration Example

Creating the IPv6 Multicast Service on the IPv6 Network and Configuring the RPF Route Selection According to the Longest Matching Principle

Scenario **Figure 2- 8**	
Configuration Steps	● Configure IP multicast basic functions (omitted). ● Configure an IPv6 multicast static route with the subnet mask length smaller than that of the unicast route on Device B. ● Configure the RPF route selection according to the longest matching principle on Device B.
B	B# configure terminal B(config)# ipv6 multicast-routing B(config)# ipv6 mroute 2005::/64 ospfv3 2002::2 B(config)# ipv6 multicast rpf longest-match
Verification	Run the **show ipv6 rpf** command to display the RPF information about the multicast source before and after RPF route selection according to the longest matching principle is configured.
Before Configuration	B# show ipv6 rpf 2005::1 RPF information for 2005::1 RPF interface: GigabitEthernet 0/2 RPF neighbor: 2002::2 RPF route: 2005::/64 RPF type: unicast (ospf) RPF recursion count: 0 Doing distance-preferred lookups across tables Distance: 110 Metric: 1
After Configuration	B# show ipv6 rpf 2005::1 RPF information for 2005::1 RPF interface: GigabitEthernet 0/1 RPF neighbor: fe80::2d0:f8ff:fe22:341b RPF route: 2005::1/128 RPF type: unicast (ospf) RPF recursion count: 0 Doing distance-preferred lookups across tables

	Distance: 110 Metric: 1

Common Errors

- IPv6 unicast routing is incorrectly configured.
- IPv6 multicast routing is not enabled on a router.

2.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the IPv6 multicast forwarding table.	**clear ipv6 mroute** { * | *v6group-address* [*v6source –address*] }
Clears the statistics in the IPv6 multicast forwarding table.	**clear ipv6 mroute statistics** { * | *v6group-address* [*v6source-address*] }

Displaying

Description	Command
Displays the IPv6 multicast forwarding table information.	**show ipv6 mroute** [*group-or-source-address* [*group-or-source-address*]] [**sparse**] [**summary** | **count**]
Displays RPF information about a specific IPv6 source address.	**show ipv6 rpf** *v6source-address*
Displays information the IPv6 static multicast route.	**show ipv6 mroute static**
Displays information about the configured IPv6 multicast interface that takes effect.	**show ipv6 mvif** [*interface-type interface-number*]
Displays the IPv6 Layer-3 multicast forwarding table.	**show ipv6 mrf mfc**
Displays the IPv6 multi-layer multicast forwarding table.	**show msf6 msc**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs all running processes of the IPv6 multicast.	**debug nsm mcast6 all**
Debugs the communication between the IPv6 multicast and the protocol module.	**debug nsm mcast6 fib-msg**

Description	Command
Debugs the interface running of the IPv6 multicast.	**debug nsm mcast6 mif**
Debugs the processing of interfaces and behavior statistics of the IPv6 multicast.	**debug nsm mcast6 stats**
Debugs the Layer-3 multicast forwarding of IPv6.	**debug ipv6 mrf forwarding**
Debugs the operation process of IPv6 Layer-3 multicast forwarding entries.	**debug ipv6 mrf mfc**
Debugs the processing of IPv6 Layer-3 multicast forwarding events.	**debug ipv6 mrf event**
Debugs the forwarding of IPv6 multi-layer multicast packets.	**debug msf6 forwarding**
Debugs the operation process of IPv6 multi-layer multicast forwarding entries.	**debug msf6 mfc**
Debugs the underlying hardware for IPv6 multi-layer multicast forwarding.	**debug msf6 ssp**
Debugs the APIs for IPv6 multi-layer multicast forwarding.	**debug msf6 api**
Debugs the processing of IPv6 multi-layer multicast forwarding events.	**debug msf6 event**

3 Configuring IGMP

3.1 Overview

The Internet Group Management Protocol (IGMP) is a member of TCP/IP protocol family. It manages IP multicast members and is used to establish and maintain multicast group membership between hosts and directly neighboring multicast routers. IGMP behaviors are classified into host behaviors and device behaviors.

- At present, three IGMP versions are available, which are IGMPv1, IGMPv2 and IGMPv3.
- All IGMP versions support the Any-Source Multicast (ASM) model.
- IGMPv3 can be directly used for the Source-Specific Multicast (SSM) model.
- IGMPv1 and IGMPv2 can be used for the SSM model only when the IGMP SSM Mapping technology is supported.

Protocols and Standards

- RFC 1112: Host Extensions for IP Multicasting
- RFC 2236: Internet Group Management Protocol, Version 2
- RFC 3376: Internet Group Management Protocol, Version 3
- RFC 4605: Internet Group Management Protocol (IGMP) / Multicast Listener Discovery

(MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")

3.2 Applications

Application	Description
Local IGMP Service	Implements the IGMP service in a local network.
IGMP Proxy Service	In a simple tree network topology, use the IGMP proxy service instead of the PIM service.

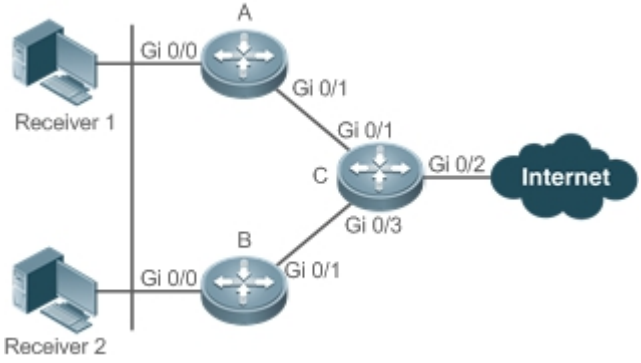
3.2.1 Local IGMP Service

Scenario

As shown in Figure 3- 1, receivers 1 and 2 and routers A and B form a local network.

Query packets sent by router A or B are valid in the LAN, whereas Report packets sent by receivers 1 and 2 are also valid locally.

Figure 3- 1



Remarks	C is the egress gateway (SG) device.

	A and B are core routers.

Deployment

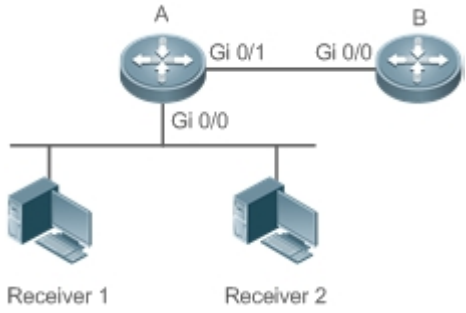
- Routers A, B and C run OSPF.
- The interfaces of A, B and C run multicast protocols (PIM-SM or PIM-DM).

3.2.2 IGMP Proxy Service

Scenario

As shown in Figure 3- 2, router A implements the proxy function working as a host and forms a local network group with router B. Router A forwards Report packets sent by receivers 1 and 2.

Figure 3- 2



Remarks	Router A implements the proxy function. Router B provides the PIM service.

Deployment

- Routers A and B run OSPF.
- The interfaces of A and B run multicast protocols (PIM-SM or PIM-DM).
- The multicast proxy function is implemented on the interfaces Gi0/0 and Gi0/1 of router A.

3.3 Features

Basic Concepts

Host Behavior and Device Behavior

- Layer-3 multicast devices that run multicast management protocols are called devices and their behaviors are called device behaviors.
- PCs or simulated PCs that run multicast management protocols are called hosts and their behaviors are called host behaviors.

Querier

- Devices compete against each other by comparing IP addresses. Devices with lower IP addresses become queriers and send Query packets regularly.

IGMP Proxy-Service Interface

- This interface performs host behaviors, receives Query packets sent by upstream devices (hence also called uplink interface), and sends Report information collected by the router proxy.

IGMP Mroute–Proxy Interface

- This interface implements the router functions, sends packets received by the IGMP PROXY-SERVICE interface (hence also called downlink interface), and collects host information and sends the host information to the IGMP PROXY-SERVICE interface.

IGMP SSM Mapping

- Mapping of the SSM model. IGMPv1 and IGMPv2 do not support the SSM model, but can enable the SSM-MAP function to support the SSM model.

Overview

Feature	Description
IGMP Router	Sends Query packets and obtains local member information.
IGMP Group Filtering	Filters group members and limit the number of group members.
Static IGMP Group	Static group information is available on a router; therefore, it is unnecessary for the host to send a Report packet to obtain the static group information.
Simulating Hosts to Join IGMP Groups	Simulates the host behavior to directly join a multicast group on an interface.
IGMP Proxy	Use this function in a simple tree network topology where no complex multicast route protocols (such as PIM) need to be executed.
IGMP SSM Mapping	Provides the SSM model support for IGMPv1 and IGMPv2. When a host joins a group, you can specify a source to save bandwidth and prevent unwanted and invalid multicast data streams from occupying network bandwidth, especially in a network environment where multiple multicast sources share one multicast address.
Router Alert Option	Checks whether IGMP packets contain the Router Alert option and discards the packets without the Router Alert option. Sends IGMP packets with the Router Alert option.

3.3.1 IGMP Router

- This function is used to send Query packets and obtain local member information.

Working Principle

- In a multicast network running the IGMP, a multicast device periodically sends IGMP Query packets and confirms information about local members based on responses.

- Only one multicast device sends IGMP Query packets in one network segment and this device is called querier. The querier is determined by means of selection. Initially, all multicast devices are in the Querier state. When a device receives a membership query from a device with a lower IP address, the device changes from the Querier state to the Non-querier state. Therefore, only one device is in the Querier state finally. This device has the lowest IP address among all multicast devices in the network.

- The querier sends IGMP packets of different versions based on the IGMP version settings. In addition, the following querier parameters can be modified: frequency for the querier to send IGMP Query packets, query times and query interval for the last member, maximum response time of IGMP Query packets, and keepalive time of the existing querier.

Related Configuration

Enabling IGMP

IGMP is disabled on an interface by default.

You can run the **ip pim** { **sparse-mode**| **dense-mode** } command to enable or disable IGMP for an interface.

IGMP can be enabled only when Sparse Mode (SM) or Dense Mode (DM) is configured on the interface.

Specifying the IGMP Version

IGMPv2 is enabled by default.

You can run the **ip igmp version** { **1** | **2** | **3** } command to set or reset the IGMP version.

Configuring the Last-Member Query Interval

The interval for sending the last-member Query packets is 1s by default.

You can run the **ip igmp last-member-query-interval** *interval* command to set or reset the interval for an interface to send Query packets.

A larger value means a larger interval; a smaller value means a smaller interval.

Configuring the Last-Member Query Times

The number of the last-member query times is 2 by default.

You can run the **ip igmp last-member-query-count** *count* command to set or reset the number of the last-member query times.

A larger value means more last-member query times; a smaller value means fewer last-member query times.

Configuring the Common Member Query Interval

The common member query interval is 125s by default.

You can run the **ip igmp query-interval** *seconds* command to set or reset the common member query interval.

A larger value means a larger common query interval; a smaller value means a smaller common query interval.

Configuring the Maximum Response Time

The maximum response time is 10s by default.

You can run the **ip igmp query-max-response-time** *seconds* command to set or reset the maximum response time.

A larger value means longer response time; a smaller value means shorter response time.

Configuring the Querier Timeout

The querier timeout is 255s by default.

You can run the **ip igmp query-timeout** *seconds* command to set the querier timeout.

A larger value means longer survival time; a smaller value means shorter survival time.

3.3.2 IGMP Group Filtering

Filter group members and limit the number of group members.

Working Principle

To prevent hosts in a network segment where an interface resides from joining a multicast group, you can configure an ACL on this interface as a filter. The interface will filter the received IGMP membership Report packets based on this ACL, maintain group membership only for multicast groups allowed by this ACL and set the maximum number of router members.

Related Configuration

Configuring the IGMP Group ACL

By default, no ACL is used and any group is allowed to join.

You can run the **ip igmp access-group** *access-list-name* command to set or reset the multicast group ACL.

After the ACL is configured, a router receives only packets set in the ACL.

Configuring the Maximum Number of IGMP Group Members

The maximum number of IGMP group members is 1,024 by default.

You can run the **ip igmp limit** *number* command to set or reset the maximum number of multicast group members.

A larger value means more members; a smaller value means fewer members.

3.3.4 Static IGMP Group

When static IGMP groups are available on a router, it is unnecessary for the host to send a Report packet to obtain the static group information. The router can directly exchange group information with a PIM router.

Working Principle

You need to set static group information manually.

Related Configuration

Configuring a Static Group

No static group is configured by default.

You can run the **ip igmp static-group** *group-address* command to configure a static group.

3.3.5 Simulating Hosts to Join IGMP Groups

Simulate the host behavior to directly join a multicast group on an interface.

Related Configuration

Configuring the Join-Group function

No join-group information is set by default.

You can run the **ip igmp join-group** *group-address* command to configure the address of the multicast group to be joined by the simulated host.

3.3.6 IGMP Proxy

Use this function in a simple tree network topology where no complex multicast route protocols (such as PIM) need to be executed. In this way, a downstream proxy host can send IGMP packets and maintain the membership.

Working Principle

When an upstream router is configured as an IGMP proxy-service interface, it is equal to a host that can receive Query packets sent by upstream routers or forward group information sent by downstream hosts. When a downstream router is configured as an IGMP multicast proxy interface, it is equal to a router that can forward Query packets sent by upstream routers or receive Report packets sent by downstream routers.

Related Configuration

Configuring the IGMP Proxy Service

The IGMP proxy service function is disabled by default.

You can run the **ip igmp proxy-service** command to enable the IGMP proxy service.

This function is mandatory when a proxy is to be used.

Configuring the IGMP Mroute Proxy

The IGMP mroute proxy function is disabled by default.

You can run the **ip igmp mroute-proxy** *interfacename* command to enable the IGMP mroute proxy.

This function is mandatory when a proxy is to be used.

3.3.7 IGMP SSM Mapping

Provide the SSM model support for IGMPv1 and IGMPv2. When a host joins a group, you can specify a source to save bandwidth and prevent unwanted and invalid multicast data streams from occupying network bandwidth, especially in a network environment where multiple multicast sources share one multicast address.

Working Principle

Based on IGMP v1/v2, IGMPv3 provides an extra function, namely, the multicast source filter function. In IGMPv1/v2, a host determines to join a group only based on the group address and then receive multicast streams sent to this group address from any source. A host using IGMPv3 advertises the multicast group that the host wants to join and the addresses of multicast sources from which this host wants to receive packets. IGMPv1 and IGMPv2 also implement "source address filtering" in some sense; however, they implement this function on the multicast receivers by enabling the SSM mapping function and configuring the static SSM mapping group.

Related Configuration

Enabling IGMP SSM Mapping

The SSM mapping function is disabled by default.

You can run the **ip igmp ssm-map enable** command to enable the function.

Configuring Static IGMP SSM Mapping

No static SSM mapping is set by default.

You can run the **ip igmp ssm-map static** *access-list-num A.B.C.D* command to configure static SSM mapping.

3.3.8 Router Alert Option

Check whether IGMP packets contain the Router Alert option and discard packets without the Router Alert option.

Support sending IGMP packets containing the Router Alert option.

Working Principle

If a packet contains the Router Alert option, the device needs to check the packet in depth and updates the control data accordingly. If the packet does not contain the option, the device does not check the packet.

After Router Alert option check is enabled, the IGMP packets not containing the Router Alert option are discarded.

After enabled with the function of sending packets with Router Alert option, the device sends IGMP packets with Router Alert option encapsulated.

Related Configuration

Checking Router Alert Option

Router Alert option check is disabled by default.

You can run the **ip igmp enforce-router-alert** command to enable the function.

Sending IGMP Packets with Router Alert Option Encapsulated

Packets are sent without the Router Alert option by default.

You can run the **ip igmp send-router-alert** command to enable the function.

3.4 Configuration

Configuration	Description and Command	
Configuring IGMP Basic Functions	(Mandatory) It is used to set up the multicast service.	
	ip multicast-routing	Enables the IPv4 multicast routing function.
	ip pim { **sparse-mode**| **dense-mode** }	Enables the PIM-SM or PIM-DM function.
Configuring IGMP Routers	**ip igmp version** { **1** | **2** | **3** }	Specifies the IGMP version.
	ip igmp last-member-query-interval interval	Configures the last-member query interval.
	ip igmp last-member-query-count *count*	Configures the last-member query times.
	ip igmp query-interval *seconds*	Configures the membership query interval.
	ip igmp query-max-response-time *seconds*	Configures the maximum response time.
	ip igmp query-timeout *seconds*	Configures the querier timeout.
Configuring IGMP Group Filtering	**ip igmp access-group** *access-list*	Configures the IGMP group ACL.
	ip igmp limit *number* [**except** *access-list*]	Configures the maximum number of IGMP group members.

Configuring IGMP Proxy	**ip igmp proxy-service**	Configures the IGMP proxy service.
	ip igmp mroute-proxy *interface-type interface-number*	Configures the IGMP mroute proxy.
Configuring IGMP SSM Mapping	**ip igmp ssm-map enable**	Enables IGMP SSM mapping.
	ip igmp ssm-map static *access-list source-address*	Configures static IGMP SSM mapping.
Configuring Alert Option	**ip igmp enforce-router-alert**	Checks the Router Alert option.
	ip igmp send-router-alert	Sends IGMP packets containg the Router Alert option.

3.4.1 Configuring IGMP Basic Functions

Configuration Effect

- Enable the multicast routing function of a local network and collect group information of the local network.

Notes

- An interface must be enabled with the PIM-SM or PIM-DM function.

Configuration Steps

Enabling the IPv4 Multicast Routing Function

- Mandatory.
- If there is no special requirement, the IPv4 multicast routing function should be enabled on each router in the local network.

Enabling the PIM-SM or PIM-DM Function

- Mandatory.
- If there is no special requirement, the PIM-SM or PIM-DM function should be directly enabled on an interface of the local network.

Verification

Run the **show ip igmp interface** *interface-type interface-number* command to check whether IGMP is enabled on the interface.

Related Commands

Enabling the IPv4 Multicast Routing Function

Command	**ip multicast-routing**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling the PIM-SM or PIM-DM Function

Command	**ip pim** { **sparse-mode**| **dense-mode** }
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	PIM interfaces must be layer-3 interfaces, including routing interfaces, L3AP, SVI and loopback interfaces. All PIM interfaces should be accessible to IPv4 unicast routes.

Configuration Example

Enabling IGMP for a Local Network

Scenario	● Configure an IPv4 unicast routing protocol (such as OSPF) on a router and ensure that the loopback interface is accessible to a unicast route. ● Enable the IPv4 multicast route function on all routers. ● Enable the PIM-SM or PIM-DM function on interfaces interconnecting devices and interfaces connecting user hosts and multicast sources.
	VSU(config)#ip multicast-routing VSU(config)#int gi 0/5 VSU(config-if-GigabitEthernet 0/5)#ip add 192.168.1.90 255.255.255.0 VSU(config-if-GigabitEthernet 0/5)#ip pim sparse-mode
Verification	Run the **show ip igmp interface** *interface-type interface-number* command to check whether IGMP is enabled on the interface.
	VSU#show ip igmp interface gigabitEthernet 0/5 Interface GigabitEthernet 0/5 (Index 5) IGMP Active, Querier, Version 2 (default) Internet address is 192.168.1.90 IGMP interface limit is 1024 IGMP interface has 1 group-record states IGMP interface has 0 static-group records IGMP activity: 3 joins, 0 leaves IGMP query interval is 125 seconds IGMP querier timeout is 255 seconds IGMP max query response time is 10 seconds Last member query response interval is 10 Last member query count is 2 Group Membership interval is 260 seconds Robustness Variable is 2

Common Errors

- Routers in the network are not enabled with the multicast routing function.
- No multicast interface is available in the network.

3.4.2 Configuring IGMP Routers

Configuration Effect

- Modify the querier timeout and IGMP router parameters will affect the type of packets to be sent and the sending method.

Notes

- The basic functions of IGMP must be configured.

Configuration Steps

Specifying the IGMP Version

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Configuring the Last-Member Query Interval

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Configuring the Last-Member Query Times

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Configuring the Common Member Query Interval

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Configuring the Maximum Response Time

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Verification

Run the **show ip igmp interface** *interface-type interface-number* command to display the interface configurations.

Related Commands

Specifying the IGMP Version

Command	**ip igmp version** { **1** | **2** | **3** }
Parameter	1: Indicates IGMPv 1.

Description	2: Indicates IGMPv 2. 3: Indicates IGMPv 3.
Command Mode	Interface configuration mode
Usage Guide	After this command is configured, IGMP will automatically restart.

Configuring the Last-Member Query Interval

Command	**ip igmp last-member-query-interval** *interval*
Parameter Description	*Interval*: Indicates the interval for sending the Query packets of a specific group. The value ranges from 1 to 255 in the unit of 0.1s, and the default value is 10 (namely, 1s).
Command Mode	Interface configuration mode
Usage Guide	This command applies only to IGMPv2 or IGMPv3. When an interface receives a Leave packet, the interface sends Query packets of the group continually and waits for a response from the host. After timeout occurs, the IGMP router assumes that the group member does not exist in the directly connected network segment and deletes the interface from the IGMP group. The timeout duration is equal to the value of **last-member-query-interval** multiplied by **last-member-query-count** plus 1/2 of **query-max-response-time**.

Configuring the Last-Member Query Times

Command	**ip igmp last-member-query-count** *count*
Parameter Description	*count*: Indicates the times for sending the Query packets of a specific group, ranging from 2 to 7. The default value is 2.
Command Mode	Interface configuration mode
Usage Guide	This command applies only to IGMPv2 or IGMPv3. When an interface receives a Leave packet, the interface sends Query packets of the group continually and waits for a response from the host. After timeout occurs, the IGMP router assumes that the group member does not exist in the directly connected network segment and deletes the interface from the IGMP group. The timeout duration is equal to the value of **last-member-query-interval** multiplied by **last-member-query-count** plus 1/2 of **query-max-response-time**.

Configuring the Common Member Query Interval

Command	**ip igmp query-interval** *seconds*
Parameter Description	*seconds*: Indicates the common member query interval, ranging from 1 to 18,000s. The default value is 125.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Maximum Response Time

Command	**ip igmp query-max-response-time** *seconds*
Parameter	*seconds*: Indicates the maximum response time, ranging from 1 to 25s. The default value is 10.

Description	
Command Mode	Interface configuration mode
Usage Guide	After sending Query packets, the interface waits for a response. If timeout occurs, the IGMP router assumes that the group member does not exist in the directly connected network segment and deletes the group information.

Configuring the Querier Timeout

Command	**ip igmp query-timeout** *seconds*
Parameter Description	*seconds*: Indicates the keepalive time of the querier, ranging from 60s to 300s. The default value is 255s.
Command Mode	Interface configuration mode
Usage Guide	After sending Query packets, an interface waits for Query packets sent by other devices. If timeout occurs, the IGMP router assumes that the querier is unique in the directly connected network segment.

Configuration Example

Configuring Basic Router Parameters

Scenario	● Configure basic functions of IGMP. ● Specify the IGMPv3. ● Configure the last-member query interval to 15 (1.5s). ● Configure the number of the last-member queries to 3. ● Configure the common member query interval to 130s. ● Configure the maximum response time to 15s. ● Configure the querier timeout to 280s.
	VSU(config-if-GigabitEthernet 0/5)#ip igmp version 3 VSU(config-if-GigabitEthernet 0/5)#ip igmp last-member-query-count 3 VSU(config-if-GigabitEthernet 0/5)#ip igmp last-member-query-interval 15 VSU(config-if-GigabitEthernet 0/5)#ip igmp query-interval 130 VSU(config-if-GigabitEthernet 0/5)#ip igmp query-max-response-time 15 VSU(config-if-GigabitEthernet 0/5)#ip igmp query-timeout 280
Verification	Run the **show ip igmp interface** *interface-type interface-number* command to check the IGMP functions of the interface.
	VSU#show ip igmp interface gigabitEthernet 0/5 Interface GigabitEthernet 0/5 (Index 5) IGMP Enabled, Active, Querier, Version 3 Internet address is 192.168.1.90 IGMP interface limit is 1024 IGMP interface has 1 group-record states

	IGMP interface has 0 static-group records IGMP activity: 3 joins, 0 leaves IGMP query interval is 130 seconds IGMP querier timeout is 280 seconds IGMP max query response time is 15 seconds Last member query response interval is 15 Last member query count is 3 Group Membership interval is 275 seconds Robustness Variable is 2 VSU#

Common Errors

- The basic functions of IGMP are not enabled.

3.4.3 Configuring IGMP Group Filtering

Configuration Effect

- A router filters IGMP group members.

Notes

- The basic functions of IGMP must be configured.

Configuration Steps

Configuring the IGMP Group ACL

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Configuring the Maximum Number of IGMP Group Members

Optional.

If there is no special requirement, you can perform this configuration on all router interfaces directly connected to the local network.

Verification

IGMP Group ACL

- Configure an interface to allow only groups in ACL 1 to join. The access addresses of ACL 1 are 225.0.0.1~225.0.0.255.
- Configure the interface to join a group whose address is 225.0.0.5.
- Configure the interface to join a group whose address is 236.0.0.5.
- View the group information of the current interface.

Maximum Number of IGMP Group Members

- Set the maximum member quantity to 5 on an interface.
- Configure the interface to join a group whose address is from 225.0.0.5 to 225.0.0.10.
- View the group information of the interface.

Related Commands

Configuring the IGMP Group ACL

Command	**ip igmp access-group** *access-list*
Parameter Description	*access-list*: Defines a group address range by using a standard IP ACL or an extended ACL. The value ranges from 1 to 199, 1300 to 2699 and characters.
Command Mode	Interface configuration mode
Usage Guide	Configure this command on an interface to control the groups that hosts in a directly connected network segment can join. Use an ACL to limit the group address range. If Report packets denied by the ACL are received, the packets will be discarded. When IGMPv3 is enabled, this command supports an extended ACL. If the received IGMP Report information is (S1,S2,S3…Sn,G), this command will apply the corresponding ACL to the (0,G) information for matching. Therefore, you must configure a (0,G) record explicitly for the extended ACL in order to normally filter (S1,S2,S3…Sn,G).

Configuring the Maximum Number of IGMP Group Members

Command	**ip igmp limit** *number* [**except** *access-list*]
Parameter Description	*number*: Indicates the maximum number of IGMP group members, whose value range varies with devices. The default value is 1,024 for an interface and 65,536 globally. **except** *access-list*: Indicates that the groups in the ACL are not counted. **access-list** indicates a standard IP ACL. The value ranges from 1 to 99, 1300 to 1999 and words.
Command Mode	Interface configuration mode
Usage Guide	Global configuration mode: Limits the maximum quantity of the IGMP group members in a system. Interface configuration mode: limits the maximum quantity of IGMP group members on an interface. If the quantity of group members exceeds the interface or global limit, the Report packets received subsequently will be ignored. If an Except ACL is configured, Report packets within a specified range can be normally processed; therefore, the generated group members are not counted. The interface and global configurations can be performed independently. If the global quantity limit is smaller than that for an interface, the global configuration shall be used.

Configuration Example

Configuring IGMP Group Filtering

Scenario	• Configure the basic functions of IGMP. • Configure the access address range of ACL 1 from 225.0.0.1 to 225.0.0.255. • Set the address of the group to be joined to 225.0.0.5. • Set the address of the group to be joined to 236.0.0.5.

	VSU(config)#access-list 1 permit 225.0.0.1 225.0.0.255 VSU(config-if-GigabitEthernet 0/5)#ip igmp access-group 1 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.5 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 236.0.0.5
Verification	Run the **show ip igmp groups** [*interface-type interface-number*] [*group-address*] [**detail**] command to display the group information of the interface.
	VSU(config-if-GigabitEthernet 0/5)#show ip igmp groups IGMP Connected Group Membership Group Address Interface Uptime Expires Last Reporter 225.0.0.5 GigabitEthernet 0/5 00:14:00 00:02:45 192.168.1.90

Configuring the Maximum Number of IGMP Group Members

Scenario	• Configure the basic functions of IGMP. • Configure the maximum number of IGMP group members for the interface to 5. • Add group information (225.0.0.5~225.0.0.12). • View group information.
	VSU(config-if-GigabitEthernet 0/5)#ip igmp limit 5 VSU(config-if-GigabitEthernet 0/5)# VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.5 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.6 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.7 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.8 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.9 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.10 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.11 VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.12
Verification	Run the **show ip igmp groups** [*interface-type interface-number*] [*group-address*] [**detail**] command to display the group information of the interface.
	VSU(config-if-GigabitEthernet 0/5)#show ip igmp groups IGMP Connected Group Membership Group Address Interface Uptime Expires Last Reporter 225.0.0.5 GigabitEthernet 0/5 00:20:15 00:03:09 192.168.1.90 225.0.0.6 GigabitEthernet 0/5 00:20:24 00:02:58 192.168.1.90 225.0.0.7 GigabitEthernet 0/5 00:00:15 00:04:29 192.168.1.90

	225.0.0.8	GigabitEthernet 0/5	00:00:13	00:04:34	192.168.1.90
	225.0.0.9	GigabitEthernet 0/5	00:00:11	00:04:33	192.168.1.90

Common Errors

- The basic functions of IGMP are not enabled.

3.4.4 Configuring IGMP Proxy

Configuration Effect

- Configure the router proxy function and collect local member information.

Notes

- The basic functions of IGMP must be configured.

Configuration Steps

Configuring the IGMP Proxy Service

- Optional.
- If there is no special requirement, you can perform this configuration on directly connected upstream router interfaces.

Configuring the IGMP Mroute Proxy

- Optional.
- If there is no special requirement, you can perform this configuration on directly connected downstream host interfaces.

Verification

- Set interface 7 for directly connecting to an upstream router as a multicast proxy server.
- Set interface 1 for directly connecting to a downstream host as a multicast proxy.
- Set interface 1 to be joined by groups whose addresses are 225.0.0.6 and 225.5.5.5.
- View the current group information.

Related Commands

Configuring the IGMP Proxy Service

Command	**ip igmp proxy-service**
Parameter Description	*N/A*
Command Mode	Interface configuration mode
Usage Guide	Run the **ip igmp proxy-service** command to set the uplink interface as a Proxy-Service interface. Run the **ip igmp mroute-proxy** command to set the downlink interface as a Mroute-Proxy interface. Forward IGMP Query packets from the Proxy-Service interface to the Mroute-Proxy interface. Forward IGMP Report packets from the Mroute-Proxy interface to the Proxy-Service interface.

	A device allows a maximum of 32 Proxy-Service interfaces. After a Proxy-Service interface receives an IGMP Query packet, the interface sends a response based on the IGMP group member records. If the **switchport** command is executed on the Proxy-Service interface, the **ip igmp mroute-proxy** command configured on the Mroute-Proxy interface will be deleted automatically.

Configuring the IGMP Mroute Proxy

Command	**ip igmp mroute-proxy** *interface-type interface-number*
Parameter Description	*N/A*
Command Mode	Interface configuration mode
Usage Guide	Run the **ip igmp proxy-service** command to set the uplink interface as a Proxy-Service interface. Run the **ip igmp mroute-proxy** command to set the downlink interface as a Mroute-Proxy interface. Forward IGMP Query packets from the Proxy-Service interface to the Mroute-Proxy interface. Forward IGMP Report packets from the Mroute-Proxy interface to the Proxy-Service interface.

Configuration Example

Scenario	● Configure basic functions of IGMP. ● Configure interface 7 as a proxy server. ● Configure interface 1 as a multicast proxy. ● Set interface 1 to be joined by groups whose addresses are 225.0.0.6 and 225.5.5.5.
	VSU(config-if-GigabitEthernet 0/7)#ip igmp proxy-service VSU(config-if-GigabitEthernet 0/7)#exit VSU(config)#int gi 0/1 VSU(config-if-GigabitEthernet 0/1)#ip igmp mroute-proxy gigabitEthernet 0/7 VSU(config-if-GigabitEthernet 0/1)#ip igmp join-group 225.0.0.6 VSU(config-if-GigabitEthernet 0/1)#ip igmp join-group 225.5.5.5
Verification	Run the **show ip igmp groups** [*interface-type interface-number*] [*group-address*] [**detail**] command to display the group information of the interface.
	VSU(config-if-GigabitEthernet 0/1)#show ip igmp groups IGMP Connected Group Membership Group Address Interface Uptime Expires Last Reporter 225.0.0.6 GigabitEthernet 0/1 00:23:05 00:02:40 192.168.36.90 225.5.5.5 GigabitEthernet 0/1 00:22:06 00:02:41 192.168.36.90 IGMP Proxy-server Connected Group Membership Group Address Interface Uptime

	225.0.0.6 GigabitEthernet 0/7 00:23:05 225.5.5.5 GigabitEthernet 0/7 00:22:06 VSU(config-if-GigabitEthernet 0/1)#

Common Errors

- The basic functions of IGMP are not enabled.

3.4.5 Configuring IGMP SSM Mapping

Configuration Effect

- IGMPv3 supports source filtering; however, IGMPv1 and IGMPv2 do not support source filtering, but provides the SSM mapping function to filter sources.

Notes

- The basic functions of IGMP must be configured.

Configuration Steps

Enabling SSM Mapping

(Mandatory) Enable the SSM mapping function.

Enable the SSM mapping function on a router.

Configuring Static SSM Mapping

Optional.

Configure this function on routers enabled with SSM mapping.

Verification

Run the **show ip igmp ssm-mapping** [*group-address*] command to display SSM mapping information.

Related Commands

Enabling SSM Mapping

Command	**ip igmp ssm-map enable**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	Run the **ip igmp ssm-map enable** command to enable the SSM mapping function. Run the **ip igmp ssm-map static** command to set static mapping entries. Run IGMPv3 on the interface. When IGMPv1 or IGMPv2 Report packets are received, source addresses of static mappings can be added.

Configuring Static SSM Mapping

Command	**ip igmp ssm-map static** *access-list source-address*
Parameter Description	*access-list*: Indicates the group address range set by a standard IP ACL. The value ranges from 1 to 99, 1300 to 1999 and words. *source-address*: Indicates the source address.
Command Mode	Global configuration mode
Usage Guide	Run the **ip igmp ssm-map enable** command to enable the SSM mapping function. Run the **ip igmp ssm-map static** command to set static mapping entries. Run IGMPv3 on the interface. When IGMPv1 or IGMPv2 Report packets are received, source addresses of static mappings can be added.

Configuration Example

Scenario	● Configure the basic functions of IGMP. ● Enable SSM mapping. ● Configure static SSM mapping ACL 1.
	VSU(config)#ip igmp ssm-map enable VSU(config)#ip igmp ssm-map static 1 192.168.5.9
Verification	Run the **show ip igmp ssm-mapping** [**group-address**] command to display SSM mapping information.
	VSU#show ip igmp ssm-mapping SSM Mapping : Enabled Database : Static mappings configured

Common Errors

- The basic functions of IGMP are not enabled.

3.4.6 Configuring Alert Option

Configuration Effect

- Check whether IGMP packets contain the Router Alert option and discards the packets without the Router Alert option.
- Support sending IGMP packets with the Router Alert option.

Notes

- The basic functions of IGMP must be configured.

Configuration Steps

Checking Router Alert Option

Optional.

Sending IGMP Packets with Router Alert Option Encapsulated

Optional,

Verification

Checking Router Alert Option

Check whether the IGMP-enabled interface discards the IGMP packets without the Router Alert option.

Sending IGMP Packets with Router Alert Option Encapsulated

Check whether the IGMP-enabled interface sends the IGMP packets containing the Router Alert option.

Related Commands

Checking Router Alert Option

Command	**ip igmp enforce-router-alert**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	Run the **ip igmp enforce-router-alert** command to enable Router Alert option check. Run the **no ip igmp enforce-router-alert** command to disable Router Alert option check.

Sending IGMP Packets with Router Alert Option Encapsulated

Command	**ip igmp send-router-alert**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	Run the **ip igmp send-router-alert** command to enable the function of sending IGMP packets containing Router Alert option. Run the **no ip igmp send-router-alert** command to disable the function.

Configuration Example

Checking Router Alert Option

Scenario	• Configure the basic functions of IGMP. • Configure Router Alert option check.
	VSU(config)#ip igmp enforce-router-alert

Verification	IGMP packets containing Router Alert option 225.1.1.1 are sent to the IGMP-enabled interface and these packets are processed. Run the **show ip igmp groups** command and you will see 225.1.1.1. IGMP packets not containing Router Alert option 225.1.1.1 are sent to the IGMP-enabled interface and these packets are discarded. Run the **show ip igmp groups** command and you will not see 225.1.1.1

Sending IGMP Packets with Router Alert Option Encapsulated

Scenario	● Configure the basic functions of IGMP. ● Configure the function of sending IGMP packets containing router alert option.
	VSU(config)#ip igmp send-router-alert
Verification	Check whether the IGMP-enabled interface sends the IGMP packets containing the Router Alert option.

3.5 Monitoring

Clearing

Description	Command
Clears dynamic group membership from the IGMP buffer.	**clear ip igmp group**
Clears interface information from the IGMP buffer.	**clear ip igmp interface** *interface-type interface-number*

Displaying

Description	Command
Displays all groups in a directly connected subnet.	**show ip igmp groups**
Displays details about all groups in a directly connected subnet.	**show ip igmp groups detail**
Displays specified groups in a directly connected subnet.	**show ip igmp groups** *A.B.C.D*
Displays details about specified groups in a directly connected subnet.	**show ip igmp groups** *A.B.C.D* **detail**
Displays IGMP configurations of a specified interface in a directly connected subnet.	**show ip igmp interface** *interface-type interface-number*
Displays details about all groups of a specified interface in a directly connected subnet.	**show ip igmp groups** *interface-type interface-number* **detail**
Displays information about a specified group of a specified interface in a directly connected subnet.	**show ip igmp groups** *interface-type interface-number A.B.C.D*

Displays details about a specified group of a specified interface in a directly connected subnet.	**show ip igmp groups** *interface-type interface-number A.B.C.D* **detail**
Displays configurations of an IGMP interface.	**show ip igmp interface** [*interface-type interface-number*]
Displays configurations of all IGMP interfaces.	**show ip igmp interface**
Displays configurations of IGMP SSM mapping.	**show ip igmp ssm-mapping**
Displays the information of IGMP SSM mapping to *A.B.C.D*.	**show ip igmp ssm-mapping** *A.B.C.D*

Debugging

Description	Command
Displays whether IGMP debugging is enabled.	**show debugging**
Debugs all IGMP information.	**debug ip igmp all**
Debugs IGMP packet decoding.	**debug ip igmp decode**
Debugs IGMP packet encoding.	**debug ip igmp encode**
Debugs IGMP events.	**debug ip igmp events**
Debugs IGMP FSM.	**debug ip igmp fsm**
Debugs IGMP state machine.	**debug ip igmp tib**
Debugs IGMP warning.	**debug ip igmp warning**

4 Configuring MLD

4.1 Overview

Multicast Listener Discovery (MLD) is a protocol used in the multicast technology.

This protocol receives the multicast member relationship between hosts and routers to determine multicast flow forwarding. Using information obtained from MLD, a device maintains an interface-based multicast listener status table. The multicast listener status table is activated only when at least one host in the link of the interface is a group member.

Currently, MLD has two versions: MLDv1 and MLDv2.

- MLD of both versions supports the Any-Source Multicast (ASM) model.
- MLDv2 can be directly applied to the Source-Specific Multicast (SSM) model.
- MLDv1 can be applied to the SSM model only when MLD SSM mapping is configured.

Protocols and Standards

- RFC2710: Multicast Listener Discovery (MLDv1) for IPv6
- RFC3810: Multicast Listener Discovery Version 2 (MLDv2) for IPv6

4.2 Applications

Application	Description
Configuring the MLD Service on the Local Network	Implements the MLD service on the local network.
Configuring the MLD Proxy Service	In the simple tree topology, the MLD proxy service, instead of the PIM service, is used.

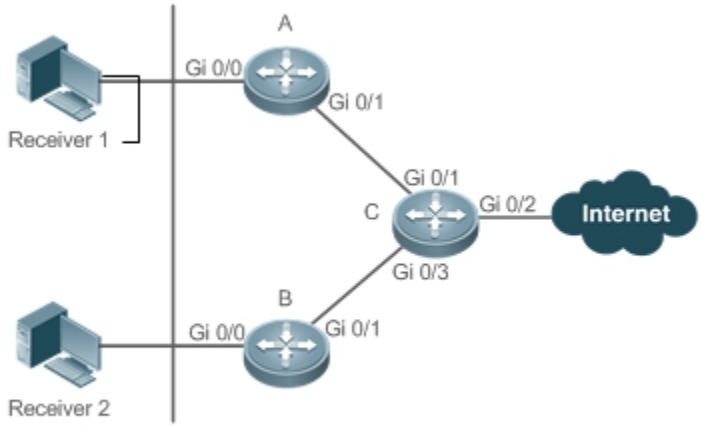
4.2.1 Configuring the MLD Service on the Local Network

Scenario

As shown in Figure 4- 1, the local network consists of receiver 1, receiver 2, router A, and router B.

Query messages sent by router A or router B are valid on the local network, and Report messages sent by receiver A and receiver B are also valid on the local network.

Figure 4- 1



Remarks	Router C is the egress gateway. Routers A and B are local routers.

Deployment

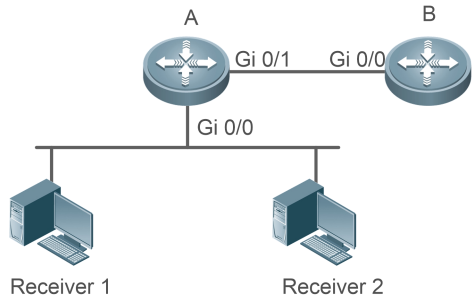
- Routers A, B, and C run the OSPFv6 protocol.
- Interfaces on routers A, B, and C run the multicast protocol (PIM SMv6 or PIM DMv6).

4.2.2 Configuring the MLD Proxy Service

Scenario

As shown Figure 4- 2, the proxy function is enabled on router A. Router A functions as a host and forms a local management group with router B. Router A forwards Report messages from receivers 1 and 2.

Figure 4- 2



Remarks	Router A functions as the proxy. Router B provides the PIM service.

Deployment

- Routers A and B run the OSPFv6 protocol.
- Interfaces on routers A and B run the multicast protocol (PIM SMv6 or PIM DMv6).
- The multicast proxy service is enabled on Gi 0/0 and Gi 0/1 of router A.

4.3 Features

Basic Concepts

Host Behaviors and Device Behaviors

Layer-3 multicast devices running multicast management protocols are referred to as devices and their behaviors are device behaviors.

PCs or simulated PCs running multicast management protocols are referred to as hosts and their behaviors are host behaviors.

Querier

Devices interact and compete with each other. After IP address comparison, the device with a lower IP address becomes the querier and periodically sends Query messages.

MLD PROXY-SERVICE Interface

This interface, also called uplink interface, implements host behaviors. It receives Query messages sent by upstream devices and sends Report messages collected by the router proxy.

MLD MROUTE-PROXY Interface

This interface, also called downlink interface, implements router functions. It sends messages received by the proxy service interface and collects and sends host information to the proxy service interface.

MLD SSM-MAP

SSM mapping refers to mapping of source-specific multicast. MLDv1 does not support the SSM model until the SSM-MAP function is enabled.

Overview

Feature	Description
Setting MLD Router Parameters	Sends Query messages to obtain local member information.
Querier Selection Process or Timeout Mechanism	Selects the unique querier in the current network segment.
Filtering MLD Groups	Filters group members and limits the number of group members.
Supporting Static MLD Groups	Stores static group information on the local router instead of obtaining group information by sending Report messages.
Configuring Simulated Host Group Information	Simulates host behaviors to directly configure group joining information.
Supporting MLD Proxy	Uses this function in the simple tree topology instead of complex multicast routing protocols, such as the PIM.
Supporting SSM-MAP	Provides the SSM model for MLDv1. When a host is added to a group, a specific source can be specified to avoid network bandwidth occupation by unnecessary and invalid multicast data streams. This function is especially useful on a network where multiple multicast sources share the same multicast address.

4.3.1Setting MLD Router Parameters

Sends Query messages to obtain local member information.

Working Principle

A device periodically sends Query messages to ensure that a group has at least one host. If no host is available in a group, the group will be deleted.

Related Configuration

Enabling MLD

By default, MLD is disabled on an interface.

Run the **ipv6 pim** { **sparse-mode**| **dense-mode** } command to enable or disable MLD on an interface.

MLD can be enabled only after PIM SM or PIM DM is enabled on the interface.

Configuring MLD Version

By default, the MLD version is 2.

Run the**ipv6 mld version** { **1** | **2** } command to configure or restore the MLD version of an interface.

Configuring the Query Interval of the Last Member

By default, the interval for sending Query messages is 1s.

Run the **ipv6 mld last-member-query-interval** *interval* command to configure or restore the interval for sending Query messages.

A larger value means a longer interval for sending Query messages.

Configuring the Number of Times for Querying the Last Member

By default, the number of times for querying the last member is 2.

Run the **ipv6 mld last-member-query-count** *count* command to configure or restore the number of times for querying the last member.

A larger value means a larger number of times for querying the last member.

Configuring the Interval for Querying a Common Member

By default, the interval for querying a common member is 125s.

Run the **ipv6 mld query-interval** *seconds* command to configure or restore the interval for querying a common member.

A larger value means a longer interval for querying a common member.

Configuring the Maximum Response Time

By default, the maximum response time is 10s.

Run the **ipv6 mld query-max-response-time** *seconds* command to configure or restore the maximum response time.

A larger value means a longer maximum response time.

4.3.2Querier Selection Process or Timeout Mechanism

Selects the unique querier in the current network segment. The querier sends a Query message to obtain group information on the local network.

Working Principle

On a multicast network running MLD, a multicast device dedicated to query sends MLD Query messages. The device is determined by election. Initially, all devices are in the querier state. When receiving member relationship Query messages from devices with lower IP addresses, the devices switch from the receiver state to non-querier state. Therefore, there is only one device in the query state in the end. This device has the lowest IP address among all multicast devices on the network. When the querier device does not work, MLD also works. Non-querier devices maintain the keepalive interval timer for other queriers. The timer is reset once the device receives a member relationship query message. If the timer times out, the device starts to send Query messages and a new querier election starts.

Related Configuration

Configuring the Keepalive Interval of the Querier

By default, the keepalive interval of the querier is 255s.

Run the **ipv6 mld querier-timeout** *seconds* command to configure or restore the keepalive interval of the querier.

A larger value means a longer keepalive interval of the querier.

4.3.3 Filtering MLD Groups

Filters group members and limits the number of group members.

Working Principle

If you do not want hosts in the network segment where an interface resides to be added to certain multicast groups, you can configure ACL rules on the interface as a filter. The interface will filter received MLD member relationship Report messages based on the ACL rules and maintain member relationships only for multicast groups permitted by the rules. The largest number of router members can also be set.

Related Configuration

Configuring Access Control for Multicast Groups

By default, no access control is configured and hosts can be added to any groups.

Run the **ipv6 mld access-group** *access-list-name* command to configure or restore access control for multicast groups.

After the configuration, the router can receive messages only from hosts in groups specified in the access list.

Configuring the Maximum Number of MLD Group Members

By default, an MLD group has a maximum of 1024 members.

Run the **ipv6 mld limt** *number* command to configure or restore the maximum number of MLD group members.

A larger value means a larger number of group members.

4.3.4 Supporting Static MLD Groups

Stores static group information on a local router instead of obtaining group information by sending Report messages. The local router can directly exchange group information with the PIM router.

Working Principle

Manually configure static group information.

Related Configuration

Configuring Static-Group

By default, no static group information is configured.

Run the **ipv6 mld static-group** *group-address* command to configure or cancel static group information.

4.3.5 Configuring Simulated Host Group Information

Simulates host behaviors to directly configure group joining information.

Related Configuration

Configuring Join-Group

By default, no join-group information is configured.

Run the **ipv6 mld join-group** *group-address* command to configure or cancel join-group information.

4.3.6 Supporting MLD Proxy

In the simply tree topology, it is not necessary to run complex multicast routing protocols (such as PIM). In this case, MLD proxy can be used to send MLD messages for downstream hosts and maintain member relationships.

Working Principle

When an upstream router is configured as an MLD proxy service interface, it functions as a host and can receive Query messages from upstream routers as well as forward group information of downstream hosts. When a downstream router is configured as an MLD multicast proxy interface, it functions as a router and can forward Query messages of upstream routers as well as receive Report messages from downstream routers.

Related Configuration

Configuring MLD PROXY-SERVICE

By default, the MLD proxy service is disabled on an interface.

Run the **ipv6 mld proxy-service** command to configure or cancel the MLD proxy function on an interface.

This function must be configured when proxy is used.

Configuring MLD MROUTE–PROXY

By default, the multicast proxy service is disabled on an interface.

Run the **ipv6 mld mroute-proxy** interfacename command to configure or cancel the multicast proxy function on an interface.

This function must be configured when proxy is used.

4.3.7 Supporting SSM-MAP

This function provides the SSM model for MLDv1. When a host is added to a group, a specific source can be specified to avoid network bandwidth occupation by unnecessary and invalid multicast data streams. This function is especially useful on a network where multiple multicast sources share the same multicast address.

Working Principle

Based on MLDv1, MLDv2 provides an extra function, that is, source filtering multicast. In MLDv1, a host determines to join a group only based on the group address and receives multicast streams sent to the group address from any source. However, an MLDv2 host advertises the multicast group that the host wants to join and the address of the multicast source that it wants to receive. In MLDv1, source address filtering can be implemented to some extent, but filtering is implemented by enabling SSM-MAP and configuring SSM-MAP static groups on multicast flow receivers.

Related Configuration

Enabling MLD SSM-MAP

By default, SSM-MAP is disabled.

Run the **ipv6 mld ssm-map enable** command to enable or disable the SSM-MAP function.

This function must be enabled when SSM-MAP is used.

Configuring MLD SSM-MAP STATIC

By default, no SSM-MAP static link table is configured.

Run the **ipv6 mld ssm-map static** *access-list-num A.B.C.D* command to enable or disable the SSM-MAP static link table.

4.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of MLD	(Mandatory) It is used to configure the multicast service.	
	Ipv6 multicast-routing	Enables the IPv6 multicast routing function.
	Ipv6 pim sparse-mode	Enables the PIM-SM function.
Configuring MLD Router Parameters	**Ipv6 mld version** { **1** | **2** }	Configures the MLD version.
	Ipv6 mld last-member-query-interval *interval*	Configures the interval for querying the last member.
	Ipv6 mld last-member-query-count *count*	Configures the number of times for querying the last member.
	Ipv6 mld query-interval *seconds*	Configures the interval for querying a common member.
	Ipv6 mld query-max-response-time *seconds*	Configures the maximum response interval.
Querier Selection Process or Timeout Mechanism	**ipv6 mld querier-timeout** *seconds*	Configures the keepalive interval of the querier.
Filtering MLD Groups	**Ipv6 mld access-group** *access-list*	Filters MLD group members.
MLD Proxy	**Ipv6 mld proxy-service**	Configures the MLD PROXY-SERVICE.
	ipv6 mld mroute-proxy *interface-type interface-number*	Configures the MLD MROUTE–PROXY.
Supporting SSM-MAP	**Ipv6 mld ssm-map enable**	Enables the SSM-MAP function.
	Ipv6 mld ssm-map static access-list source-address	Configures the SSM-MAP static link table.

4.4.1 Configuring Basic Functions of MLD

Configuration Effect

- Enable the multicast routing function and collect group information on the local network.

Notes

- The PIM SM or PIM DM function must be enabled on an interface.

Configuration Steps

Enabling the IPv6 Multicast Routing Function

- Mandatory.
- The IPv6 multicast routing function should be enabled on all routers on the local network unless otherwise specified.

Enabling the PIM SM or PIM DM Function

- Mandatory.
- The PIM SM or PIM DM function should be directly enabled on an interface on the local network unless otherwise specified.

Verification

Run the **show ipv6 mld interface *interface-type*** *interface-number* command to check whether MLD is enabled on the interface.

Related Commands

Enabling the IPv6 Multicast Routing Function

Command	**Ipv6 multicast-routing**
Parameter Description	**-**
Command Mode	Global configuration mode
Usage Guide	-

Enabling the PIM SM or PIM DM Function

Command	**Ipv6 pim** { **sparse-mode**| **dense-mode** }
Parameter Description	**-**
Command Mode	Interface configuration mode
Usage Guide	PIM interfaces must be layer-3 interfaces, including: routing, L3AP, SVI, and loopback interfaces. IPv6 unicast routes should be accessible to all PIM interfaces.

Configuration Example

Enabling MLD on the Local Network

Configuration Steps	● Configure an IPv6 unicast routing protocol (such as OSPF) on a router and ensure that unicast routes are accessible to the loopback interface. (Omitted) ● Enable the IPv6 multicast routing function on all routers. ● Enable the PIM SM or PIM DM function on device interconnection interfaces and interfaces for connecting user hosts and multicast sources.
	VSU(config)#ipv6 multicast-routing VSU(config)#int gi 0/1 VSU(config-if-GigabitEthernet 0/1)# ipv6 address 2001::1/64 VSU(config-if-GigabitEthernet 0/1)#ipv6 pim sparse-mode
Verification	Run the **show ipv6 mld interface *interface-type*** *interface-number* command to check whether MLD is enabled on the interface.
	VSU#show ipv6 mld interface gigabitEthernet 0/1 Interface GigabitEthernet 0/1 (Index 1) MLD Active, Querier, Version 2 (default) Internet address is fe80::2d0:f8ff:fe22:33b1 MLD interface limit is 1024 MLD interface has 0 group-record states MLD interface has 0 join-group records MLD interface has 0 static-group records MLD activity: 0 joins, 0 leaves MLD query interval is 125 seconds MLD querier timeout is 255 seconds MLD max query response time is 10 seconds Last member query response interval is 10 (1/10s) Last member query count is 2 Group Membership interval is 260 Robustness Variable is 2

Common Errors

- Multicast routing is disabled on routers on the network.
- No multicast interface is available on the network.

4.4.2 Configuring MLD Router Parameters

Configuration Effect

- Modify MLD router parameters to change the message type or sending mode.

Notes

- The basic functions of MLD must be configured.

Configuration Steps

Configuring MLD Version

Optional.

This parameter can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Configuring the Interval for Querying the Last Member

Optional.

This parameter can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Configuring the Number of Times for Querying the Last Member

Optional.

This parameter can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Configuring the Interval for Querying a Common Member

Optional.

This parameter can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Configuring the Maximum Response Interval

Optional.

This parameter can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Verification

Run the **show ipv6 mld interface *interface-type*** *interface-number* command to view the configuration information.

Related Commands

Configuring the MLD Version

Command	**ipv6 mld version** { **1** | **2** }
Parameter Description	1: Indicates version 1. 2: Indicates version 2.
Command Mode	Interface configuration mode
Usage Guide	After this command is executed, MLD will automatically restart.

Configuring the Interval for Querying the Last Member

Command	**ipv6 mld last-member-query-interval** *interval*
Parameter Description	*Interval*: Specifies the interval for sending Query messages of a specified group. The unit is 0.1s, the value ranges from 1 to 255, and the default value is 10 (1s).

Command Mode	Interface configuration mode
Usage Guide	After receiving the Done message, the interface will continuously send Query messages of a specified group and wait for responses from the host. After timeout, it is considered that the no group member exists in the directly-connected network segment and the interface is deleted from the MLD group member record. The timeout interval is calculated as follows: Timeout interval = last-member-query-interval x last-member-query-count + query-max-response-time/2.

Configuring the Number of Times for Querying the Last Member

Command	**ipv6 mld last-member-query-count** *count*
Parameter Description	*count*: Specifies the number of times for sending Query messages of a specified group. The value ranges from 2 to 7. The default value is **2**.
Command Mode	Interface configuration mode
Usage Guide	After receiving the Done message, the interface will continuously send Query messages of a specified group and wait for responses from the host. After timeout, it is considered that the no group member exists in the directly-connected network segment and the interface is deleted from the MLD group member record. The timeout interval is calculated as follows: Timeout interval = last-member-query-interval x last-member-query-count + query-max-response-time/2

Configuring the Interval for Querying a Common Member

Command	**ipv6 mld query-interval** *seconds*
Parameter Description	*seconds*: Specifies the interval for querying a common member. The unit is s, the value ranges from 1 to 18000, and the default value is 125.
Command Mode	Interface configuration mode
Usage Guide	-

Configuring the Maximum Response Interval

Command	**ipv6 mld query-max-response-time** *seconds*
Parameter Description	*seconds*: Specifies the maximum response time. The unit is s, the value ranges from 1 to 25, and the default value is 10.
Command Mode	Interface configuration mode
Usage Guide	After sending Query messages, the interface waits for responses. After timeout, it is considered that no group member exists in the directly-connected network segment and group information is deleted.

Configuration Example

Configuring Basic Router Parameters

Configuration Steps	● Configure basic functions of MLD. (Omitted) ● Configure MLD version 2. ● Configure the interval for querying the last member as 15 (1.5s). ● Configure the number of times for querying the last member as 3. ● Configure the interval for querying the common member as 130s. ● Configure the maximum response time as 15s.
	VSU(config-if-GigabitEthernet 0/1)#ipv6 mld version 2 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld last-member-query-count 3 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld last-member-query-interval 15 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld query-interval 130 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld query-max-response-time 15
Verification	Run the **show ipv6 mld interface** ***interface-type*** *interface-number* command to check whether MLD is enabled on the interface.
	VSU(config-if-GigabitEthernet 0/1)# show ipv6 mld interface gi 0/1 Interface GigabitEthernet 0/1 (Index 1) MLD Enabled, Active, Querier, Version 2 (default) Internet address is fe80::2d0:f8ff:fe22:33b1 MLD interface limit is 1024 MLD interface has 0 group-record states MLD interface has 0 join-group records MLD interface has 0 static-group records MLD activity: 0 joins, 0 leaves MLD query interval is 130 seconds MLD querier timeout is 267 seconds MLD max query response time is 15 seconds Last member query response interval is 15 (1/10s) Last member query count is 3 Group Membership interval is 275 Robustness Variable is 2

Common Errors

- Basic functions of MLD are not enabled.

4.4.3 Querier Selection Process or Timeout Mechanism

Configuration Effect

- Select the unique querier on the local network.

Notes

- The basic functions of MLD must be configured.

Configuration Steps

- This function must be configured if the querier keepalive interval needs to be configured.
- This function can be configured on all MLD-enabled interfaces on the local network.

Verification

Run the **show ipv6 mld interface *interface-type*** *interface-number* command to view the configuration information of the interface.

Related Commands

Configuring the Keepalive Interval of Other Queriers

Command	**ipv6 mld querier-timeout** *seconds*
Parameter Description	*seconds*: Specifies the keepalive interval for other queriers. The unit is s, the value ranges from 60 to 300, and the default value is 255.
Command Mode	Interface configuration mode
Usage Guide	After sending Query messages, the interface waits for Query messages from other devices. After timeout, it is considered that it is the unique querier in the directly-connected network segment.

Configuration Example

Configuring the Keepalive Interval of Other Queriers

Configuration Steps	• Configure the basic functions of MLD. (Omitted) • Configure the keepalive interval of a querier as 280s.
	VSU(config-if-GigabitEthernet 0/1)#ipv6 mld querier-timeout 280
Verification	Run the **show ipv6 mld** [**vrf** *vrf-name*] **interface *interface-type*** *interface-number* command to check whether MLD is enabled on the interface.
	VSU#show ipv6 mld interface gigabitEthernet 0/1 Interface GigabitEthernet 0/1 (Index 1) MLD Enabled, Active, Querier, Version 2 (default) Internet address is fe80::2d0:f8ff:fe22:33b1 MLD interface limit is 1024 MLD interface has 0 group-record states MLD interface has 0 join-group records MLD interface has 0 static-group records MLD activity: 0 joins, 0 leaves MLD query interval is 130 seconds

	MLD querier timeout is 280 seconds MLD max query response time is 15 seconds Last member query response interval is 15 (1/10s) Last member query count is 3 Group Membership interval is 275 Robustness Variable is 2

Common Errors

- The basic functions of MLD are not enabled.

4.4.4 Filtering MLD Groups

Configuration Effect

- A router filters MLD group information.

Notes

- The basic functions of MLD must be configured.

Configuration Steps

Configuring Access Control for Multicast Groups

Optional.

This function can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Configuring the Maximum Number of MLD Group Members

Optional.

This function can be configured on all router interfaces directly connected to the local network unless otherwise specified.

Verification

Filtering MLD Groups

Configure the interface to allow for only groups in link table 1. The access address of link table 1 is (FF66::100/64).

Configure the interface to add a group FF66::05.

Configure the interface to add a group FF65::05.

Check group information on the interface.

Configuring the Maximum Number of MLD Group Members

Configure the number of group members as 5 on the interface.

Configure the interface to add a group (FF66::05 ~ FF65::0B).

Check group information on the interface.

Related Commands

Configuring Access Control for Multicast Groups

Command	**ipv6 mld access-group** *access-list*
Parameter Description	access-list: Specifies the group address range by using IP standard ACLs or IP extended ACLs. The value ranges from 1 to 199, 1300 to 2699, and WORD.
Command Mode	Interface configuration mode
Usage Guide	After running this command on the interface, you can control the groups that hosts in the directly-connected network segment can join. Use ACLs to limit the group address range. Report messages denied by the ACLs will be discarded. When MLDv2 is enabled, this command supports extended ACLs to precisely filter source record information in MLDv2 messages. When the received MLD Report message is (S1,S2,S3…Sn,G), this command will match (0,G) using the corresponding ACLs. Therefore, to normally use this command, you must explicitly configure a (0, G) in the extended ACLs to filter (S1,S2,S3…Sn,G).

Configuring the Maximum Number of MLD Group Members

Command	**ipv6 mld limit** *number* [**except** *access-list*]
Parameter Description	*number*: Specifies the maximum number of MLD group members. The value range depends on the specific device. The interface default value is 1024 and the global one is 65536. **except** *access-list*: Groups in the access list are not counted. The access list is an IP standard ACL. The value ranges from 1 to 99, 1300 to 1999, and WORD.
Command Mode	Interface configuration mode
Usage Guide	Global configuration mode: Limits the number of MLD group members on the whole device. Interface configuration mode: Limits the number of MLD group members of the interface. If the number of group members exceeds the interface limit or global limit, subsequent Report messages will be ignored. If an except list is configured, Report messages in a specified range can be normally processed. Therefore, the group members are not counted. Interface and global limits can be configured separately. If the global limit is smaller than the interface limit, use the global limit.

Configuration Example

Configuring Group Filtering

Configuration Steps	• Configure the basic functions of MLD. (Omitted) • Configure the access address of link table 1 as (FF66::100/64). • Configure the group to join as FF66::05. • Configure the group to join as FF65::05.
	VSU(config)#ipv6 access-list acl VSU(config-ipv6-acl)#permit ipv6 ::/64 ff66::100/64 VSU(config-ipv6-acl)#permit ipv6 2222::3333/64 ff66::100/64 VSU(config-ipv6-acl)#exit

	VSU(config)# VSU(config)#int gi 0/1 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld access-group acl VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::5 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff65::5
Verification	Run the **show ipv6 mld groups** [*interface-type interface-number*] [*group-address*] [**detail**] command to view the group information on the interface.
	VSU#show ipv6 mld groups MLD Connected Group Membership Group Address Interface Uptime Expires Last Reporter ff66::5 GigabitEthernet 0/1 00:05:07 00:03:46 fe80::2d0:f8ff:fe22:33b1

Configuring the Maximum Number of MLD Group Members

Configuration Steps	● Configure the basic functions of MLD. (Omitted) ● Configure the maximum number of group members on the interface as 5. ● Add group information (FF66::5 ~ FF66::0B). ● View the group information.
	VSU(config-if-GigabitEthernet 0/1)#ipv6 mld limit 5 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::5 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::6 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::7 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::8 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::9 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::A VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::B
Verification	Run the **show ipv6 mld groups** [*interface-type interface-number*] [*group-address*] [**detail**] command to view group information on the interface.
	MLD Connected Group Membership Group Address Interface Uptime Expires Last Reporter ff66::5 GigabitEthernet 0/1 00:00:36 00:04:00 fe80::2d0:f8ff:fe22:33b1 ff66::6 GigabitEthernet 0/1 00:00:34 00:04:01 fe80::2d0:f8ff:fe22:33b1 ff66::7 GigabitEthernet 0/1 00:00:22 00:04:13 fe80::2d0:f8ff:fe22:33b1 ff66::8 GigabitEthernet 0/1 00:00:18 00:04:19 fe80::2d0:f8ff:fe22:33b1 ff66::9 GigabitEthernet 0/1 00:00:14 00:04:21 fe80::2d0:f8ff:fe22:33b1

Common Errors

- The basic functions of MLD are not enabled.

4.4.5 MLD Proxy

Configuration Effect

- Configure the router proxy function and collect local member information.

Notes

- The basic functions of MLD must be configured.

Configuration Steps

Configuring MLD PROXY-SERVICE

Optional.

This function can be configured on the interface of routers directly connected to the upstream devices unless otherwise specified.

Configuring MLD MROUTE–PROXY

Optional

This function can be configured on the interface of hosts directly connected to the downstream devices unless otherwise specified.

Verification

- Configure the interface that directly connects interface 7 and upstream router as the multicast proxy service.
- Configure the interface that directly connects interface 1 and downstream host as the multicast proxy.
- Configure groups FF66::05 and FF66::06 to be added to interface 1.
- Check information of the current group.

Related Commands

Configuring MLD PROXY-SERVICE

Command	**ipv6 mld proxy-service**
Parameter Description	-
Command Mode	Interface configuration mode
Usage Guide	Run the **ipv6 mld proxy-service** command to configure the upstream interface as the proxy-service interface. Run the **ipv6 mld mroute-proxy** command to configure the downstream interface as the mroute-proxy interface. Configure the proxy-service interface to forward MLD Query messages to the mroute-proxy interface. Configure the mroute-proxy interface to forward MLD Reports messages to the proxy-service interface. A maximum of 32 proxy-service interfaces can be configured on a device. After receiving MLD Query messages, the proxy-service interface sends a response based on the MLD group member records. If you run **switchport** command on the proxy-service interface, the **ipv6 mld mroute-proxy** command configured on

	the mroute-proxy interface will be automatically deleted.

Configuring MLD MROUTE–PROXY

Command	**ipv6 mld mroute-proxy** *interface-type interface-number*
Parameter Description	-
Command Mode	Interface configuration mode
Usage Guide	Run the **ipv6 mld proxy-service** command to configure the upstream interface as the proxy-service interface. Run the **ipv6 mld mroute-proxy** command to configure the downstream interface as the mroute-proxy interface. Configure the proxy-service interface to forward MLD Query messages to the mroute-proxy interface. Configure the mroute-proxy interface to forward MLD Reports messages to the proxy-service interface.

Configuration Example

Configuration Steps	● Configure MLD basic functions. (Omitted) ● Configure interface 7 as the proxy server. ● Configure interface 1 as the multicast proxy. ● Configure groups FF66::05 and FF66::06 to be added to interface 1.
	VSU(config-if-GigabitEthernet 0/7)#ipv6 mld proxy-service VSU(config-if-GigabitEthernet 0/7)#exit VSU(config)#int gi 0/1 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld mroute-proxy gigabitEthernet 0/7 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::05 VSU(config-if-GigabitEthernet 0/1)#ipv6 mld join-group ff66::06
Verification	Run the **show ipv6 mld groups** [*interface-type interface-number*] [*group-address*] [**detail**] command to view the group information on the interface.
	VSU(config-if-GigabitEthernet 0/1)#show ipv6 mld groups MLD Connected Group Membership Group Address Interface Uptime Expires Last Reporter ff66::5 GigabitEthernet 0/1 00:00:11 00:04:31 fe80::2d0:f8ff:fe22:33b1 ff66::6 GigabitEthernet 0/1 00:00:11 00:04:33 fe80::2d0:f8ff:fe22:33b1 MLD Proxy-server Connected Group Membership Group Address Interface Uptime ff66::5 GigabitEthernet 0/7 00:00:11 ff66::6 GigabitEthernet 0/7 00:00:11

Common Errors

- The basic functions of MLD are not enabled.

4.4.6 Supporting SSM-MAP

Configuration Effect

- MLDv2 supports source filtering while MLDv1 does not. However, MLDv1 provides the SSM-MAP function to implement source filtering.

Notes

- **The basic functions of MLD must be configured**.

Configuration Steps

Enabling SSM-MAP

This function must be configured if SSM-MAP.

This function must be enabled on a router where SSM-MAP is enabled.

Configuring an SSM-MAP Static Link Table

Optional.

This function must be enabled on a router where SSM-MAP is enabled.

Verification

Run the **show ipv6 mld ssm-mapping** [*group-address*] command to display SSM-MAP information.

Related Commands

Enabling SSM-MAP

Command	**ipv6 mld ssm-map enable**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	Run the **ipv6 mld ssm-map enable** command to enable the SSM-MAP function. Run the **ipv6 mld ssm-map static** command to configure static mapping table items. The interface runs MLDv2. When receiving Report messages from MLDv1, the interface adds the static mapping source address.

Configuring an SSM-MAP Static Link Table

Command	**ipv6 mld ssm-map static** access-list source-address
Parameter Description	access-list: Specifies the group address range configured by the ACL. source-address: Source address
Command Mode	Global configuration mode

Usage Guide	Run the **ipv6 mld ssm-map enable** command to enable the SSM-MAP function. Run the **ipv6 mld ssm-map static** command to configure static mapping table items. The interface runs MLDv2. When receiving Report messages from MLDv1, the interface adds the static mapping source address.

Configuration Example

Configuration Steps	● Configure the basic functions of MLD. (Omitted) ● Enable SSM-MAP. ● Configure SSM-MAP static link table 3.
	VSU(config)#ipv6 mld ssm-map enable VSU(config)#ipv6 mld ssm-map static 3 1500::5
Verification	Run the **show ipv6 mld ssm-mapping** [*group-address*] command to view SSM mapping information.
	VSU(config)#show ipv6 mld ssm-mapping SSM Mapping : Enabled Database : Static mappings configured

Common Errors

- The basic functions of MLD are not enabled.

4.5 Monitoring

Clearing

Description	Command
Clears dynamic group member records in the MLD cache.	**clear ipv6 mld group** [*group-address*] [*interface-type interface-number*]
Clears all MLD statistics and group member records on the interface.	**clear ipv6 mld interface** *interface-type interface-number*

Displaying

Description	Command
Displays groups directly connected to the device and group information learned from MLD.	**show ipv6 mld groups** [*group-address* | *interface-type interface-number*] [**detail**]
Displays configurations of the interface.	**show ipv6 mld interface** [*interface-type interface-number*]
Displays SSM-MAP information.	**show ipv6 mld ssm-mapping** [*group-address*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Displays the MLD debugging switch status.	**show debugging**
Debugs all MLD information.	**debug ipv6 mld all**
Debugs MLD packet resolution.	**debug ipv6 mld decode**
Debugs MLD packet encoding.	**debug ipv6 mld encode**
Debugs MLD event information.	**debug ipv6 mld events**
Debugs MLD Finite State Machine (FSM).	**debug ipv6 mld fsm**
Debugs MLD state machine information.	**debug ipv6 mld tib**
Debugs MLD warning.	**debug ipv6 mld warning**

5 Configuring PIM-DM

5.1 Overview

Protocol Independent Multicast (PIM) is an intra-domain multicast routing protocol.

A multicast source sends a packet to a group address. The packet is forwarded by network devices hop by hop and finally reaches the group members. On layer-3 network devices, PIM is used to create and maintain multicast routing entries, so as to support multicast forwarding.

PIM works in two modes: Protocol Independent Multicast - Sparse Mode (PIM-SM) and Protocol Independent Multicast - Dense Mode (PIM-DM).

- PIM-SM is applicable to large-scale networks where group members are sparsely distributed in a wide scope.
- PIM-DM is applicable to small networks where group members are densely distributed.

Protocols and Standards

- RFC3973: Protocol Independent Multicast - Dense Mode (PIM-DM)
- RFC2715: Interoperability Rules for Multicast Routing Protocols

5.2 Applications

Application	Description
Providing the Multicast Service in the Same Network	The multicast service is provided in the same network.
PIM-DM Application in a Hot Backup Environment	The multicast PIM-DM protocol runs in a hot backup environment.

5.2.1Providing the Multicast Service in the Same Network

Scenario

The multicast service is provided in the same network.

The following figure is taken as an example:

- A multicast source sends a multicast packet, and Receiver A and Receiver B in the same network receive the multicast packet.

Figure 5- 1

Remarks	A, B, and C are Layer-3 routers. The multicast source is connected to the Gi0/1 interface of A, Receiver A is connected to the Gi0/1 interface of B, and Receiver B is connected to Gi0/1 of C.

Deployment

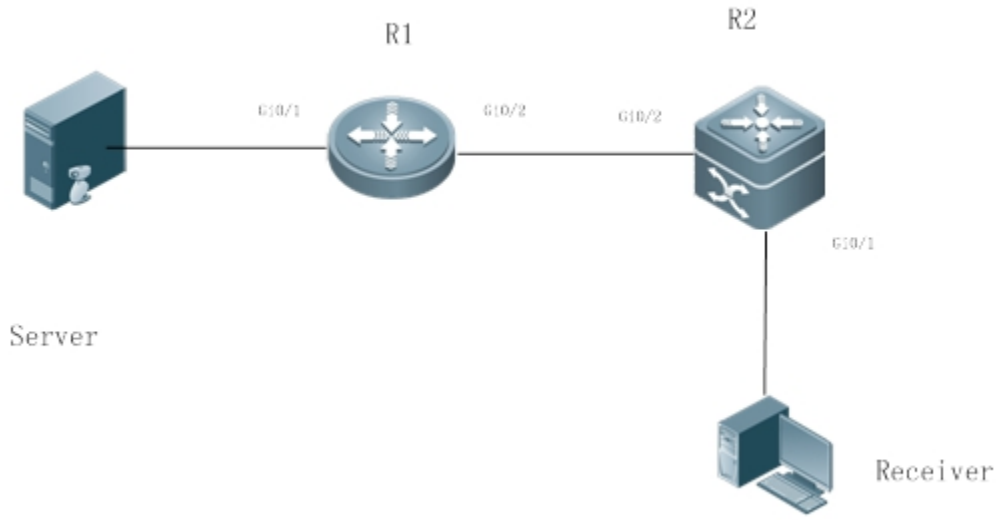
- Run the Open Shortest Path First (OSPF) protocol in the same network to implement unicast routing.
- Run the PIM-DM protocol in the same network to implement multicast routing.
- Run the Internet Group Management Protocol (IGMP) in a user host network segment to implement group member management.

5.2.2 PIM-DM Application in a Hot Backup Environment

Scenario

In a hot backup environment, run PIM-DM. A device performs hot backup switching to ensure that traffic is not interrupted.

Figure 5- 2



Remarks	R1 is connected to the video server, R2 is directly connected to the receiver, and R2 runs in hot backup mode. A Layer-3 multicast protocol runs on R1 and R2.

Deployment

- Run OSPF on R1 and R2 to implement unicast routing.
- Run PIM-DM on R1 and R2 to implement multicast routing.
- Make R2 run in a hot backup environment.

Remarks	R2 may perform hot backup switching in the hot backup environment. In this case, the query interval of PIM Hello packets (the default value is 30 seconds) needs to be adjusted on R2 because the keepalive timer of the neighbor in PIM Hello packets of R1 may have expired (the default value is 3.5 times the query interval, that is, 105 seconds). The multicast function relies on the unicast function currently, and the multicast function starts convergence after the unicast function convergence is complete.

	For example, the default graceful restart (GR) convergence time of the unicast function is 120 seconds. It is recommended that the query interval of PIM Hello packets be set to 60 seconds. The keepalive time of the neighbor in PIM Hello packets is 210 seconds. In this scenario, the query interval of PIM Hello packets need to be set with a reference to the GR convergence time of the unicast function and the value of 3.5 times the query interval of PIM Hello packets must be larger than the GR convergence time of the unicast function. In a hot backup environment, it is recommended that the query interval of PIM Hello packets be larger than the default value (30 seconds). Otherwise, the keepalive timer of the neighbor in PIM Hello packets of the peer end times out during hot backup switching.

5.3 Features

Basic Concepts

PIM Router and PIM Interface

Routers where the PIM protocol is enabled are called PIM Routers. Interfaces where the PIM protocol is enabled are called PIM interfaces.

Multicast packets are forwarded on PIM routers. The PIM interfaces where multicast packets are received are called Upstream Interfaces, and the PIM interfaces where multicast packets are sent are called Downstream Interfaces.

The network segments where upstream interfaces are located are called Upstream Network Segments. The network segments where downstream interfaces are located are called Downstream Network Segments.

PIM Network and PIM Domain

PIM routers are connected through PIM interfaces and form a PIM network.

On some PIM interfaces, borders can be set to divide a large PIM network into multiple PIM domains. The borders are able to reject specified multicast packets or limit the transmission of PIM messages.

Multicast Distribution Tree

Multicast packets are packets transmitted from one point to multiple points. The forwarding path is in a tree structure. This forwarding path is called the Multicast Distribution Tree (MDT).

(*,G), (S,G)

- (*,G): Packets sent from any source to Group G, the corresponding routing entries, and the forwarding path called Rendezvous Point Tree (RPT).
- (S,G): Packets sent from Source S to Group G, the corresponding routing entries, and the forwarding path called Shortest Path Tree (SPT).

Overview

Feature	Description
PIM-DM Neighbor	Neighbor relationships are established between PIM routers to form a PIM network.
PIM-DM SRM	PIM-DM uses a State Refresh Message (SRM) to update the network state.

5.3.1PIM-DM Neighbor

Neighbor relationships are established between PIM routers to form a PIM network. Neighbor relationships must be established between PIM routers before PIM control messages can be exchanged or multicast packets can be forwarded.

Working Principle

A Hello message is sent from a PIM interface. For the IPv4 multicast packet with the Hello message encapsulated, the destination address is 224.0.0.13 (indicating all PIM routers in the same network segment), the source address is the IP address of the PIM interface, and the Time To Live (TTL) value is 1. For the IPv6 multicast packet with the Hello message encapsulated, the destination address is ff02::d.

Function of a Hello message: It is used to discover neighbors, coordinate protocol parameters, and maintain neighbor relationships.

Discovering Neighbors

PIM routers in the same network segment receive multicast packets from the destination address 224.0.0.13 or ff02::d. In this way, the PIM routers obtain neighbor information and establish neighbor relationships.

When a PIM interface is enabled or detects a new neighbor, a Triggered-Hello-Delay message is used to generate a random time. Within the time, the interface sends Hello packets.

Coordinating Protocol Parameters

A Hello message includes multiple protocol parameters, which are described as follows:

– DR_Priority: Router interfaces contend for the designated router (DR) based on their DR priorities. A higher priority means a higher chance of winning.

– Holdtime: Time in which a neighbor is held in the reachable state

– LAN_Delay: LAN delay for transmitting a Prune message in a shared network segment

– Override-Interval: Prune override time carried in a Hello message.

When a PIM router receives a Prune message from an upstream interface, it indicates that downstream interfaces exist in the shared network segment. If the PIM router still needs to receive multicast data, the PIM router must sent a Prune Override message to the upstream interface within the Override-Interval.

LAN_Delay + Override-Interval = PPT (Prune-Pending Timer). After a PIM router receives a Prune message from an downstream interface, the PIM router will not immediately perform pruning until PPT times out. Within the time of PPT, if the PIM router receives a Prune rejection message from the downstream interface, the PIM router cancels pruning.

Maintaining Neighbor Relationships

A Hello message is sent periodically between PIM routers. If a Hello packet is not received from a PIM neighbor within Holdtime, the neighbor is considered unreachable and is deleted from the neighbor list. Any change of PIM neighbors will cause change of the multicast topology in the network. If an upstream or downstream neighbor in an MDT is unreachable, multicast routes converge again and the MDT is reshaped.

Related Configuration

Enabling PIM-DM on an Interface

By default, PIM-DM is disabled on an interface.

Use the **ip pim dense-mode** command to enable or disable PIM-DM on an interface.

PIM-DM must be enabled on an interface to involve the interface in the PIM protocol.

Setting the Interval of Hello Messages on an Interface

By default, a Hello message is sent at an interval of 30 seconds.

The **ip pim query-interval** *interval-seconds* command is used to adjust the interval of Hello messages. The value of the interval ranges from 1 to 65,535.

A Hello message is transmitted less frequently when the value of *interval-seconds* is larger.

5.3.2PIM-DM MDT

The three basic mechanisms dense-mode PIM uses to build multicast forwarding trees are: flood, prune, and graft.

Working Principle

When a multicast source sends multicast packets, the system forwards them to the outgoing interfaces of multicast neighbors and local members. The Reverse Path Forwarding (RPF) check needs to be conducted on all packets received through the upstream interface of the device. Packets that fail the RPF check will be discarded. Multicast packets that pass the RPF check are further forwarded if there is an outgoing interface. If no outgoing interface is available, the device sends a prune packet to the upstream interface. After receiving the prune packet, the upstream interface identifies the source interface of the prune packet as the Pruned state and sets the Pruned Timer (PI). In this way, a multicast forwarding tree with the multicast source as the root is created.

When the system receives a Join message from a local member, if a downstream device in the Pruned state sends a Graft message to the upstream device, the upstream device returns a Graft-Ack message and resumes forwarding of multicast data to the interface of the downstream device after receiving the Graft message.

In environment deployment, when multiple PIM-DM neighbors are created through multiple links between devices and downstream devices need to receive no or few packets, the CPU usage may be high. In this scenario, PIM-SM is recommended for the environment deployment

Related Configuration

Configuring the Prune Override Interval on an Interface

By default, the prune override interval is 500 ms.

Run the **ip pim override-interval** *interval-milliseconds* command to change the prune override interval.

5.3.3 PIM-DM SRM

PIM-DM uses an SRM to refresh the network state.

Working Principle

Devices connected to a multicast source periodically send SRMs to downstream devices to notify changes of the network topology. After receiving the SRMs, the adjacent devices receiving the SRMs add the local topology state information to the messages by modifying some fields in SRMs, and send the messages to downstream devices. When the messages reach leaf devices, the state information of the entire network is updated.

Related Configuration

Disabling the Processing and Forwarding of SRMs

By default, the processing and forwarding of SRMs are enabled.

The **ip pim state-refresh disable** command is used to disable the processing and forwarding of SRMs.

Disabling the SRM function may cause the converged PIM-DM MDT to re-converge, which leads to unnecessary bandwidth waste and multicast routing table flapping. Therefore, it is recommended not to disable SRM in general conditions.

Setting the Interval of SRMs

By default, an SRM is sent at an interval of 60 seconds.

The **ip pim state-refresh origination-interval** *interval-seconds* command is used to adjust the interval of SRMs. The value of the interval ranges from 1 to 100.

SRMs are transmitted less frequently when the value of *interval-seconds* is larger.

Only devices that are directly connected to a multicast source will periodically send a PIM SRM to downstream interfaces. For a device not directly connected to the multicast source, the interval of SRMs on its downstream interfaces is invalid.

5.3.4 MIB

Connected to other agents, the Simple Network Management Protocol (SNMP) manager uses information in the Management Information Base (MIB) to directly manage the PIM-DM function.

Working Principle

The MIB specifies variables (namely information that can be queried and set by the management process) maintained by network elements and directly manages the PIM-DM function.

Related Configuration

Enabling PIM-DM MIB

By default, the PIM-DM MIB function is enabled.

The **ip pim mib dense-mode** command is used to enable the PIM-DM MIB function.

5.4 Configuration

Configuration	Description and Command	
Configuring PIM-DM Basic Functions	(Mandatory) It is used to create the multicast service.	
	ip multicast-routing	Enables IPv4 multicast routing.
	ip pim dense-mode	Enables PIM-DM.
Configuring PIM-DM Neighbors	(Optional) It is used to limit the (S,G) pairs of legitimate multicast packets in Any Source Multicast (ASM) model.	
	ip pim query-interval *interval-seconds*	Sets the Interval of Hello messages on an interface.
	ip pim propagation-delay *interval-milliseconds*	Sets the prune propagation delay on an interface.
	ip pim override-interval *interval-milliseconds*	Sets the prune override interval on an Interface.

Configuration	Description and Command	
	ip pim neighbor-filter *access-list*	Configures neighbor filtering on an interface.
Configuring PIM-DM SRMs	**ip pim state-refresh disable**	Disables the processing and forwarding of SRMs.
	ip pim state-refresh origination-interval *interval-seconds*	Sets the Interval of SRMs on an interface.
Configuring PIM-DM MIB	**ip pim mib dense-mode**	Enables PIM-DM MIB.
Configuring PIM-DM PASSIVE mode	**ip pim dense-mode passive**	Enables PIM-DM PASSIVE mode.
Configuring the PIM-DM Sub VLAN Function	**Ip pim dense-mode subvlan** [**all** | *vid*]	Specifies, on an interface of a super VLAN, the sub VLAN to which packets are sent.

5.4.1 Configuring PIM-DM Basic Functions

Configuration Effect

- Create a PIM-DM network and provide data sources and user terminals in the network with the IPv4 multicast service.

Notes

- PIM-DM needs to use the unicast routes existing in the network. Therefore, IPv4 unicast routing must be configured in the network.

Configuration Steps

Enabling IPv4 Multicast Routing

- Mandatory
- IPv4 multicast routing should be enabled on each router unless otherwise specified.

Enabling PIM-DM

- Mandatory
- PIM-DM should be enabled on the following interfaces unless otherwise specified: interconnected interfaces on routers and interfaces connecting multicast sources and user hosts.

Enabling the PIM-DM PASSIVE Function

- In a PIM network, if an interface needs to receive multicast packets without participating in the PIM network topology construction, the PIM-DM PASSIVE mode can be configured.
- If no special requirements are raised, enable the PIM-DM PASSIVE function on the following interfaces: interfaces of the stub network device in the multicast network for connecting to STAs. After the PIM-DM PASSIVE function is configured on an interface, the interface neither sends nor receives PIM packets.

Configuring the PIM-DM Sub VLAN Function

- In most scenarios on the PIM network, the PIM DM protocol does not need to be enabled on interfaces of a super VLAN. In general, a super VLAN includes many sub VLANs. If the PIM DM protocol is enabled on the interfaces of the super VLAN, multicast packets will be replicated and sent to all sub VLANs. As a result, traffic generated easily exceeds the device processing capability, causing protocol flapping. In some scenarios that require the PIM DM protocol to be enabled on the interfaces of the super VLAN, the PIM-DM sub VLAN function may be configured, to send packets to a specified sub VLAN or all sub VLANs.

- This function is available only on the interfaces of the super VLAN.

Verification

Make multicast sources send multicast packets and make user hosts join the groups.

- Check whether the user hosts can successfully receive packets from each group.
- Check whether correct PIM-DM routing entries are created on routers.

Related Commands

Enabling IPv4 Multicast Routing

Command	**ip multicast-routing**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling PIM-DM

Command	**ip pim dense-mode**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	PIM interfaces must be at Layer-3, including: routing interfaces, aggregate ports(APs), switch virtual interfaces (SVIs), and loopback interfaces. For all PIM interfaces, IPv4 unicast routes should be reachable.

Enabling PIM-DM PASSIVE Mode

Command	**ip pim dense-mode passive**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The PIM interface must be a Layer-3 interface, including: routing interface, aggregate port, switch virtual interface, and loopback interface. For all PIM interfaces, IPv4 unicast routes should be reachable.

Enabling the PIM-DM Sub VLAN Function

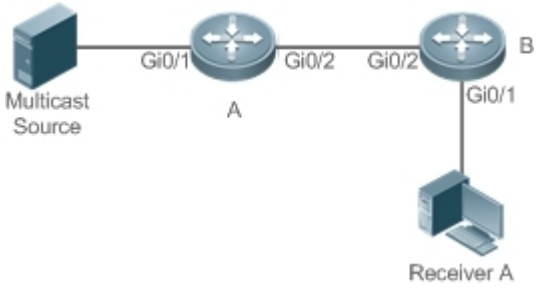
Command	**ip pim dense-mode subvlan [all | *vid*]**
Parameter Description	**all**: sends packets to all sub VLANs. **vid**: sends packets to a specified sub VLAN.
Command Mode	Interface configuration mode
Usage Guide	The PIM interface must be an interface of the super VLAN.

Displaying the PIM-DM Routing Table

Command	**show ip pim dense-mode mroute** [*group-or-source-address* [*group-or-source-address*]] [**summary**]
Parameter Description	*group-or-source-address*: Indicates a group address or source address. *group-or-source-address*: Indicates a group address or source address (The two addresses cannot be group addresses or source addresses at the same time). **summary**: Displays the routing table summary.
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	Check whether sufficient routing entries are provided. Check the upstream and downstream interface lists and ensure that a correct SPT tree is created.

Configuration Example

Enabling IPv4 Multicast Routing on the IPv4 Network

Scenario **Figure 5- 3**	
Configuration Steps	● Configure IPv4 unicast routing protocols (for example, OSPF) on all the routers. ● Enable the IPv4 multicast routing function on all the routers. ● Enable the PIM-DM function on all the interconnected interfaces of the routers, Source, and Receiver..
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if)# ip pim dense-mode A(config-if)# exit A(config)# interface GigabitEthernet 0/2

	A(config-if)# ip pim dense-mode A(config-if)# exit
B	B# configure terminal B(config)# ip multicast-routing B(config)# interface GigabitEthernet 0/1 B(config-if)# ip pim dense-mode B(config-if)# exit B(config)# interface GigabitEthernet 0/2 B(config-if)# ip pim dense-mode B(config-if)# exit
Verification	Configure the multicast source (192.168.1.10) to send packets to G (229.1.1.1). Make Receiver A join G. ● Check whether the multicast packets from Source G are received by Receiver A.. ● Check PIM-DM routing tables on Router A and Router B.
A	A# show ip pim dense-mode mroute PIM-DM Multicast Routing Table (192.168.1.10, 229.1.1.1) MRT lifetime expires in 182 seconds Source directly connected on GigabitEthernet 0/1 State-Refresh Originator State: Originator SRT:57, SAT:147 Upstream IF: GigabitEthernet 0/1 Upstream State: Forwarding Assert State: NoInfo Downstream IF List: GigabitEthernet 0/2, in 'olist': Downstream State: NoInfo Assert State: NoInfo
B	B# show ip pim dense-mode mroute PIM-DM Multicast Routing Table (192.168.1.10, 229.1.1.1) MRT lifetime expires in 130 seconds RPF Neighbor: 192.168.2.1, Nexthop: 192.168.2.1, GigabitEthernet 0/2

	Upstream IF: GigabitEthernet 0/2 Upstream State: Forwarding Assert State: Loser, AT:125 Downstream IF List: GigabitEthernet 0/1, in 'olist': Downstream State: NoInfo Assert State: NoInfo

Common Errors

- IPv4 unicast routing is incorrectly configured.
- IPv4 multicast routing is not enabled on a certain router.
- PIM-DM is not enabled on a certain interface.

5.4.2Configuring PIM-DM Neighbors

Configuration Effect

- Coordinate protocol parameters and adjust parameters in the Hello packet.
- Enable neighbor filtering to improve network security.

Notes

- Basic functions of PIM-DM must be configured.

Configuration Steps

- Set parameters on PIM router interfaces unless otherwise specified.

Verification

- Set parameters in a Hello packet on an interface and run the **debug ip pim dense-mode encode** command to check parameters.
- Enable neighbor filtering and run the **show ip pim dense-mode decode** command to display neighbor filtering information.
- Run the **show running-config interface** [*interface-type interface-number*] command to display configurations on an interface.

Related Commands

Setting the Interval of Hello Messages

Command	**ip pim query-interval** *interval-seconds*
Parameter Description	*interval-seconds*: The value ranges from 1 to 65,535 in the unit of seconds.
Command Mode	Interface configuration mode
Usage Guide	When the Hello interval is set, the holdtime value will be updated as its 3.5 times.

Every time when the interval of Hello messages is updated, the holdtime value is automatically updated as 3.5 times of the interval. If the result of the interval of Hello messages multiplied by 3.5 is greater than 65,535, the holdtime value is updated as 65,535.

Setting the Prune Propagation Delay

Command	**ip pim propagation-delay** *interval-milliseconds*
Parameter Description	*interval-milliseconds*: The value ranges from 1 to 32,767 in the unit of milliseconds.
Command Mode	Interface configuration mode
Usage Guide	Set **propagation-delay** of an interface, that is, configure the prune propagation delay of an interface.

Setting the Prune Override Interval

Command	**ip pim override-interval** *interval-milliseconds*
Parameter Description	*interval-milliseconds*: The value ranges from 1 to 32,767 in the unit of milliseconds.
Command Mode	Interface configuration mode
Usage Guide	Set **override-interval** of an interface, that is, configure the prune override time of an interface.

Configuring PIM-DM Neighbor Filtering

Command	**ip pim neighbor-filter** *access-list*
Parameter Description	*access-list*: The supported ACL ranges from 1 to 99. Naming an ACL is also supported.
Command Mode	Interface configuration mode
Usage Guide	Only addresses that meet ACL filtering conditions can be used as PIM neighbors of the current interface. Otherwise, the addresses filtered out cannot be neighbors. Peering refers to exchange of protocol packets between PIM neighbors. If peering with a PIM device is suspended, the neighbor relationship with it cannot be formed so that PIM protocol packets will not be received from the device.

Configuration Example

Configuring PIM-DM Neighbors on the IPv4 Network

Scenario **Figure 5- 4**	
Configuration Steps	● Configure basic functions of PIM-DM (omitted). ● Set protocol parameters in a Hello packet on the Gi0/1 interface of device A.

A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if)# ip pim query-interval 60 A(config-if)# ip pim propagation-delay 800 A(config-if)# ip pim override-interval 1000 A(config-if)# exit
Verification	● Run the **show running-config interface** [*interface-type interface-number*] command to display configurations on an interface. ● Run the **debug ip pim dense-mode encode** command to debug parameters in a Hello packet.
A	A# (config)#show running-config interface gigabitEthernet 0/1 Building configuration... Current configuration : 245 bytes ! interface GigabitEthernet 0/1 ip pim dense-mode ip pim query-interval 60 ip pim propagation-delay 800 ip pim override-interval 1000
	A# debug ip pim dense-mode encode *Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello Hold-Time 210 *Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello Gen-ID 1362200073 *Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello PD=800 ms, OI=1000 ms *Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello SR-Interval 60 *Dec 22 15:00:58: %7: [ENCODE] Enc Msg Hdr: Hello Checksum=65396, MsgLen=34 Assert State: Loser, AT:125

Configuring PIM-DM Neighbor Filtering on the IPv4 Network

Scenario **Figure 5- 5**	
Configuration	● Configure basic functions of PIM-DM (omitted).

Steps	● Configure an ACL on device A. ● Configure PIM neighbor filtering on the Gi0/1 interface of device A.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1 A(config-if)# ip pim query-interval 60 A(config-if)# ip pim propagation-delay 800 A(config-if)# ip pim override-interval 1000 A(config-if)# exit
Verification	● Run the **show running-config interface** [*interface-type interface-number*] command to display configurations on the interface. ● Run the **debug ip pim dense-mode decode** command to debug parameters in a Hello packet.
A	A#show running-config interface gigabitEthernet 0/2 Building configuration... Current configuration : 187 bytes ! interface GigabitEthernet 0/1 ip pim dense-mode ip pim neighbor-filter pim-dm
	A# debug ip pim dense-mode decode Dec 22 15:15:47: %7: [DECODE] Dec Msg: PIM Hello message, version 2 Dec 22 15:09:47: %7: [DECODE] Dec Msg: Neighbor 192.168.2.2/32 on GigabitEthernet 0/1 denied by access-list pim-dm

Common Errors

- IPv4 unicast routing is incorrectly configured.
- IPv4 multicast routing is not enabled on a certain router.
- PIM-DM is not enabled on a certain interface.

5.4.3Configuring PIM-DM SRMs

Configuration Effect

- Enable or disable the PIM-DM SRM function.
- Adjust the interval of SRMs.

Notes

- Basic functions of PIM-DM must be configured.

Configuration Steps

- The interval of SRMs is only applicable only to the PIM router interfaces that are directly connected to the multicast source.

Verification

- Configure the PIM-DM SRMs and run the **show running-config** command to display the SRM status.
- Run the **show ip pim dense-mode track** command to display the SRM number.
- Run the **show running-config interface** [*interface-type interface-number*] command to display interface configurations.

Related Commands

Disabling the Processing and Forwarding of SRMs

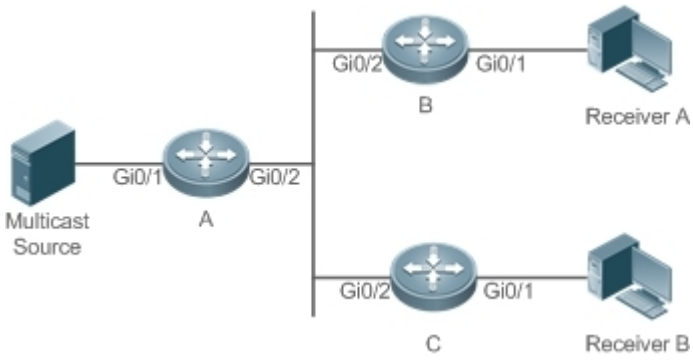
Command	**ip pim state-refresh disable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	When the processing and forwarding of SRMs are disabled, the State Refresh Capable option is not included in a Hello packet, and is not processed when the Hello packet is received. Disabling the SRM function may cause the converged PIM-DM MDT to re-converge, which leads to unnecessary bandwidth waste and multicast routing table flapping. Therefore, it is recommended not to disable this function in general conditions.

Setting the Interval of SRMs

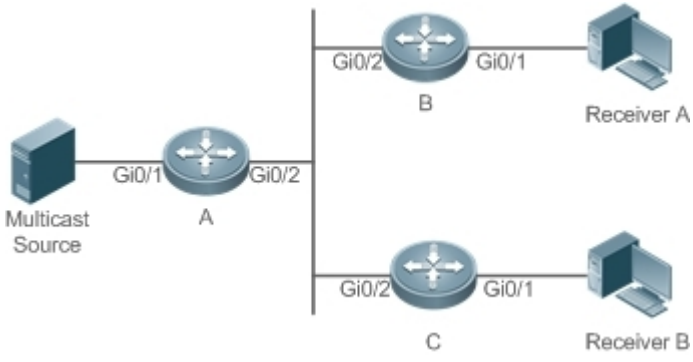
Command	**ip pim state-refresh origination-interval** *interval-seconds*
Parameter Description	*interval-seconds*: The value ranges from 1 to 100 in the unit of second.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Disabling the Processing and Forwarding of SRMs on an Interface on the IPv4 Network

Scenario **Figure 5- 6**	
Configuration Steps	● Configure basic functions of PIM-DM (omitted). ● Disable processing and forwarding of a PIM-DM SRM on an Interface of device A.
A	A# configure terminal A(config)# ip pim state-refresh disable
Verification	● Run the **show running-config** command to check the configuration.
A	A# (config)# show running-config … ! ip pim state-refresh disable ! …

Setting the Interval of SRMs on the IPv4 Network

Scenario **Figure 5- 7**	
Configuration Steps	● Configure basic functions of PIM-DM (omitted). ● Set the interval of PIM-DM SRMs on the Gi0/1 interface of device A.
A	A# configure terminal A(config)# interface GigabitEthernet 0/1

	A(config-if)# ip pim state-refresh origination-interval 5 A(config-if)# exit
Verification	● Run the **show running-config interface** [*interface-type interface-number*] command to display interface configurations. ● Run the **show ip pim dense-mode track** command to display the SRM number.
A	A#show running-config interface gigabitEthernet 0/1 Building configuration... Current configuration : 201 bytes ! interface GigabitEthernet 0/1 ip pim dense-mode ip pim state-refresh origination-interval 5
	A #show ip pim dense-mode track PIM packet counters Elapsed time since counters cleared: 00:18:54 received sent Valid PIMDM packets: 38 102 Hello: 38 76 Join/Prune: 0 0 Graft: 0 0 Graft-Ack: 0 0 Assert: 0 0 State-Refresh: 0 26 PIM-SM-Register: 0 PIM-SM-Register-Stop: 0 PIM-SM-BSM: 0 PIM-SM-C-RP-ADV: 0 Unknown Type: 0 Errors: Malformed packets: 0

	Bad checksums: 0 Unknown PIM version: 0 Send errors: 0

Common Errors

- IPv4 unicast routing is incorrectly configured.
- IPv4 multicast routing is not enabled on a certain router.
- PIM-DM is not enabled on a certain interface.

5.4.4Configuring PIM-DM MIB

Configuration Effect

- Enable the MIB function for PIM-DM.

Verification

- Configure the MIB function of PIM-SM and run the **show running-config** command to check whether the function is configured.

Related Commands

Enabling PIM-DM MIB

Command	**ip pim mib dense-mode**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

5.5 Monitoring

Clearing

Description	Command
Resets the statistic start time and clears the counters of PIM-DM packets.	**clear ip pim dense-mode track**

Displaying

Description	Command
Displays the help information of the commands with IP PIM as the key word.	**ip pim help**
Displays PIM-DM information of an interface.	**show ip pim dense-mode interface** [*interface-type interface-number*] [**detail**]
Displays the PIM-DM neighbors.	**show ip pim dense-mode neighbor** [*interface-type interface-number*]

Description	Command
Displays the PIM-DM next-hop information .	**show ip pim dense-mode nexthop**
Displays the PIM-DM routing table.	**show ip pim dense-mode mroute** [*group-or-source-address* [*group-or-source-address*]] [**summary**]
Displays the number of PIM-DM packets sent and received since the statistic start time.	**show ip pim dense-mode track**

6 Configuring PIM-SM

6.1 Overview

Protocol Independent Multicast (PIM) is an intra-domain multicast routing protocol.

A multicast source sends a packet to a group address. The packet is forwarded by network devices hop by hop and finally reaches the group members. On Layer-3 network devices, PIM is used to create and maintain multicast routing entries, so as to support multicast forwarding.

PIM works in two modes: Protocol Independent Multicast - Sparse Mode (PIM-SM) and Protocol Independent Multicast - Dense Mode (PIM-DM).

- PIM-SM is applicable to large-scale networks where group members are sparsely distributed in a wide scope.
- PIM-DM is applicable to small networks where group members are densely distributed.

Protocols and Standards

- RFC4601: Protocol Independent Multicast -Sparse Mode (PIM-SM)
- RFC5059: Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)
- RFC3962: Protocol Independent Multicast - Dense Mode protocol
- RFC4607: Source-Specific Multicast for IP

6.2 Applications

Application	Description
Enabling ASM for PIM-SM	The receiver receives any multicast source.
Enabling SSM for PIM-SM	The receiver receives only a specific multicast source.

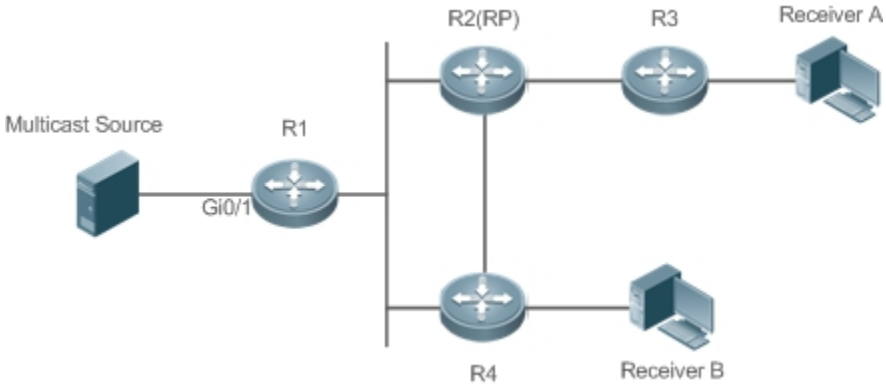
6.2.1 Enabling ASM for PIM-SM

Scenario

Provide multicast services within only one domain.

For example, in the following figure, the receiver receives any multicast source.

Figure 6- 1



Remarks	R 1 is connected directly to the multicast source. R 2 serves as the rendezvous point (RP). R 3 is connected directly to Receiver A. R 4 is connected directly to Receiver B.

Deployment

- Run the Open Shortest Path First (OSPF) protocol to realize unicast routing.
- Run PIM-SM to realize multicast routing.
- Run the Internet Group Management Protocol (IGMP) in the network segment of the user host to manage group members.

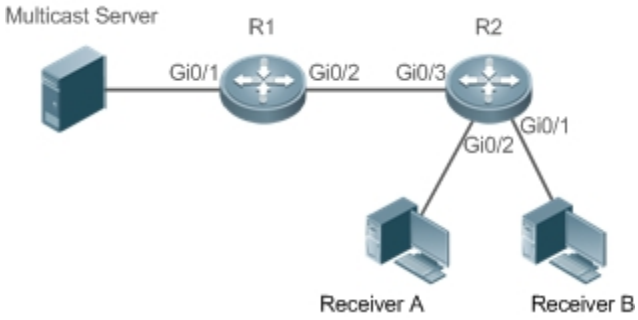
6.2.2 Enabling SSM for PIM-SM

Scenario

Provide multicast services within only one domain.

For example, in the following figure, the receiver receives a specific multicast source.

Figure 6- 2



Remarks	R 1 is connected directly to the multicast source. R 2 serves as the RP. R 2 is connected directly to Receiver A. R 2 is connected directly to Receiver B.

Deployment

- Run the OSPF protocol to realize unicast routing.
- Run PIM-SM to realize multicast routing.
- Run the source-specific multicast (SSM) of PIM-SM within the domain.
- Run IGMPv3 in the network segment of the user host to manage group members.

6.3 Features

Basic Concepts

PIM Router and PIM Interface

A router running PIM is called a PIM router. An interfaces running PIM is called a PIM interface.

Multicast packets are forwarded on PIM routers. The PIM interfaces where multicast packets are received are called upstream interfaces, and the PIM interfaces where multicast packets are sent are called downstream interfaces.

The network segments where upstream interfaces are located are called upstream network segments, and the network segments where downstream interfaces are located are called downstream network segments.

PIM Network and PIM Domain

PIM routers are connected through PIM interfaces to form a PIM network.

On some PIM interfaces, borders can be set to divide a large PIM network into multiple PIM domains. The borders can reject the passage of specific multicast packets or limit the transmission of PIM packets.

Multicast Distribution Tree, DR, and RP

Multicast packets are transmitted from one point to multiple points, forming a tree-shaped forwarding path. Such forwarding path is called the multicast distribution tree (MDT), which includes the following two types:

- RP Tree (RPT): It is rooted at an RP, and uses the designated router (DR) of the member groups connected to it as its leaves.
- Shortest path tree (SPT): It is rooted at a DR that is connected to the multicast source, and uses the RP or the DR of the member groups connected to it as its leaves.

Both the DR and RP are the functions of a PIM router.

- An RP collects the information of a multicast source or multicast member on the network.
- The DR connected to the multicast source advertises the multicast source information to the RP; the DR connected to multicast group members advertises the information of multicast group members to the RP.

(*, G), (S, G)

- (*, G): Indicates the packets sent from any source to a group (G), the corresponding route entries, and the RPT.
- (S, G): Indicates the packets sent from the source (S) to a group (G), the corresponding routing entries, and the SPT.

ASM, SSM

PIM-SM supports both any-source multicast (ASM) and SSM, and it is applicable to different multicast group address segments.

- ASM: In this model, a user is not allowed to select a multicast source. The user host joins a group, and receives the packets sent from all sources.
- SSM: In this model, a user can select a multicast source. The user host joins a group and specifies the source address. Then only the packets sent from this source address is received.

Requirements for using an SSM model: Before selecting a multicast source, you need to learn the address of the multicast source using other network services.

Overview

Feature	Description
PIM-SM Neighbor	Establishes neighbor relationships between RIM routers to form a PIM network.

Feature	Description
DR Election	In the network segment where group member hosts are located, PIM neighbors compete for the DR, and the one wins the election becomes the DR for connecting to the group members. In the network segment where the multicast source is located, PIM neighbors compete for the DR, and the one wins the election becomes the DR for connecting to the multicast source.
BSR Mechanism	On a PIM network, the BSR generates periodic candidate RPs and bootstrap packets of corresponding group addresses.
RP Mechanism	On a PIM network, through static RP configuration or dynamic RP election, the location of the RP can be learned by each PIM router.
Register Information of the Multicast Source	When the multicast source is detected on the network, the source DR sends a register packet to the RP, which obtains the source information and multicast packet.
Creating an RPT	When a group member is detected on the network, the DR connecting to the group members send packets toward the RP to form an RPT. If the multicast source already exists on the network, the packets arrived at the RP can be sent to the group members along the RPT.
Creating an SPT	When data packets arrive at the DR connecting to group members, the DR sends these packets toward the multicast source to form an SPT, and multicast packets are sent to group members along the SPT.
ASM and SSM	A PIM router can provide multicast services of both ASM model and SSM model at the same time. SSM model applies to the groups whose addresses are within the range of the SSM addresses. For other groups, use ASM model.

6.3.1PIM-SM Neighbor

Neighbor relationships are established between PIM routers to form a PIM network. Neighbor relationships must be established between PIM routers before PIM control packets can be exchanged or multicast packets can be forwarded.

Working Principle

A PIM interface sends a Hello packet. For the IPv4 multicast packet whose Hello packet is encapsulated, the destination address is 224.0.0.13 (indicating all PIM routers in the same network segment), the source address is the IP address of the PIM interface, and the Time To Live (TTL) value is 1. For the IPv6 multicast packet whose Hello packet is encapsulated, the destination address is ff02::d.

A Hello packet is used to discover neighbors, coordinate protocol parameters, and maintain neighbor relationships.

↘ Discovering Neighbors

PIM routers in the same network segment receive multicast packets from the destination address 224.0.0.13. In this way, the PIM routers obtain neighbor information and establish neighbor relationships.

When a PIM interface is enabled or detects a new neighbor, a triggered-hello-delay packet is used to generate a random time. Within the time, the interface sends Hello packets.

↘ Coordinating Protocol Parameters

A Hello packet includes multiple protocol parameters, which are described as follows:

– DR_Priority: indicates the priority of a router interface for competing for the DR. A higher priority means a higher chance of winning.

– Holdtime: Indicates the time in which a neighbor is held in the reachable state

– LAN_Delay: Indicates the LAN delay for transmitting a Prune packet in a shared network segment.

– Override-Interval: Indicates the prune override time carried in a Hello packet.

When a PIM router receives a Prune packet from an upstream interface, it indicates that downstream interfaces exist in the shared network segment. If the PIM router still needs to receive multicast data, the PIM router must sent a Prune Override packet to the upstream interface within the override interval.

LAN_Delay + Override Interval = PPT (Prune-Pending Timer). After a PIM router receives a Prune packet from a downstream interface, the PIM router will not immediately perform pruning until PPT times out. Within the time of PPT, if the PIM router receives a Prune rejection packet from the downstream interface, the PIM router cancels pruning.

Maintaining Neighbor Relationships

A Hello packet is sent periodically between PIM routers. If a Hello packet is not received from a PIM neighbor within Holdtime, the neighbor is considered unreachable and is deleted from the neighbor list. Any change of PIM neighbors will cause change of the multicast topology in the network. If an upstream or downstream neighbor in an MDT is unreachable, multicast routes converge again and the MDT is reshaped.

Related Configuration

Enabling PIM-SM on an Interface

By default, PIM-SM is disabled on an interface.

Run **ip pim sparse-mode** to enable or disable PIM-SM on an interface.

PIM-SM must be enabled on an interface to involve the interface in the PIM protocol. If PIM-SM is not enabled for the interface of a DR, static RP, candidate RP (C-RP), or candidate BSR (C-BSR), corresponding roles of the PIM protocol cannot be run.

Setting the Interval of Hello Packets on an Interface

By default, a Hello packet is sent every 30s.

Run **ip pim query-interval** *interval-seconds* to adjust the interval of Hello packets. The value ranges from 1 to 65,535.

A Hello packet is transmitted less frequently when the value of *interval-seconds* is greater.

6.3.2DR Election

In the network segment where group member hosts are located, PIM neighbors compete for the DR, and the one wins the election becomes the DR for connecting to the group members.

In the network segment where the multicast source is located, PIM neighbors compete for the DR, and the one wins the election becomes the DR for connecting to the multicast source.

The DR sends Join/Prune packets toward the MDT, or sends the multicast source data to the MDT.

Working Principle

When creating a PIM neighbor, you can send a Hello packet to obtain the IP address and DR priority of the neighbor to elect a DR.

Two parameters play a key role in winning the DR election: the DR priority of an interface and the IP address of the interface.

DR Priority of an Interface

During the DR election, the RIM router with the highest DR priority will be elected as the DR.

Interface IP Address

During the DR election, if the priority of interfaces is the same, then interface IP addresses will be compared. The interface with the maximum IP address will be elected as the DR.

Related Configuration

Enabling PIM-SM on an Interface

By default, PIM-SM is disabled on an interface.

Run **ip pim sparse-mode** to enable or disable PIM-SM on an interface.

PIM-SM must be enabled on an interface to involve the interface in the PIM protocol. If PIM-SM is not enabled for the interface of a DR, static RP, C-RP, or C-BSR, corresponding protocols cannot be run.

Adjusting the DR Priority of an Interface

By default, the DR priority is 1.

Run **ip pim dr-priority** *priority-value* to adjust the DR priority of the interface. The value ranges from 0 to 4,294,967,294.

The DR priority is used in the DR election in the network segment directly connected the interface. A greater value indicates a higher priority.

6.3.3BSR Mechanism

On a PIM network, the BSR generates periodic candidate RPs and bootstrap packets of corresponding group addresses. These bootstrap packets are sent hop by hop in the domain. All the routers on the entire network will receive these bootstrap packets, and record these candidate RPs and their corresponding group addresses.

Working Principle

One or multiple candidate BSRs are configured in a PIM-SM domain. You need to apply a certain algorithm to select the BSR from these candidate BSRs.

Related Configuration

Configuring Candidate BSRs

By default, candidate BSRs are not configured.

Run **ip pim bsr-candidate** *interface-type interface-number* [*hash-mask-length* [*priority-value*]] to configure or cancel the configuration of candidate BSRs.

Through bootstrap packet (BSM) learning and competition of candidate BSRs, a unique BSR is generated for the PIM-SM domain.

Configuring BSR Boarders

By default, BSR boarders are not configured.

Run **ip pim bsr-border** to configure or cancel the configuration of BSR boarders.

After this command is configured, BSMs received by the interface will be discarded and will not be forwarded by this interface, preventing BSM flooding.

Filtering BSMs

By default, BSMs from the BSR are not filtered.

Run **ip pim accept-bsr list** { <1-99> | <1300-1999> | *WORD* } to configure whether to filter BSMs.

If this function is enabled, only legible BSMs are received by the interface; if this function is disabled, all the external BSMs will be received by the device running PIM-SM.

Configuring Legible C-RP Addresses and the Multicast Groups They Serve for a Candidate BSR

By default, Candidate-RP-Advertisement (C-RP-Adv) packets are not filtered by a candidate BSR.

Run **ip pim accept-crp list** { *<100-199>* | *<2000-2699>* | *WORD* } to configure whether to filter C-RP-Adv packets.

If this function is enabled, C-RP addresses and corresponding multicast groups are filtered by a candidate BSR. If this function is disabled, all external C-RP-Adv packets are received by a candidate BSR.

Allowing a C-BSR to Receive a C-RP-ADV Packet Whose Prefix-Count Is 0

By default, a candidate BSR cannot receive a C-RP-ADV packet whose prefix-count is 0.

Run **ip pim accept-crp-with-null-group** to configure whether to receive a C-RP-ADV packet whose prefix-count is 0.

If this function is enabled, a C-RP-ADV packet whose prefix-count is 0 can be received by a candidate BSR. If this function is disabled, a C-RP-ADV packet whose prefix-count is 0 cannot be received by a candidate BSR.

6.3.4RP Mechanism

On a PIM network, through static RP configuration or dynamic RP election, the location of the RP can be learned by each PIM router. The RP as the root of the RPT, is the point where the RPT is rooted at and RPT data traffic is forwarded from.

Working Principle

All PIM routers in the same PIM domain must be mapped to the same RP as a specific multicast group address. On a PIM network, an RP can be configured as static or dynamic.

Static RP

In static RP configuration, RP addresses are configured directly on PIM routers and these addresses are learnt by the entire PIM network.

Dynamic RP

In a PIM-SM domain, there are candidate RPs that send unicast packets (including RP addresses and the multicast groups they serve) to the BSR, which generates periodic candidate RPs and bootstrap packets of corresponding group addresses. These bootstrap packets are sent hop by hop in the domain, and received and saved by PIM routers, which apply a hash function to map the group addresses to the candidate RP that can provide services. Then the RP corresponds to these multicast group addresses can be confirmed.

Related Configuration

Configuring Static RP Addresses

By default, no RP address is configured.

Run **ip pim rp-address** *rp-address* [*access-list*] to configure a static RP address for a PIM router.

To use static RP addresses, the static RP address of all routers in the PIM-SM domain must be the same, so that the PIM SM multicast routing remains consistent.

Configuring Candidate C-RP Addresses

By default, no C-RP address is configured.

Run **ip pim rp-candidate** *interface-type interface-number* [**priority** *priority-value*] [**interval** *interval-seconds*] [**group-list** *access-list*] to configure or cancel a PIM router as a candidate C-RP.

After a candidate RP is configured, it can send periodic C-RP-Adv packets to the BSR, and the information carried by these C-RP-Adv packets will be advertized to all PIM-SMs in the domain, ensuring the uniqueness of RP mapping.

Ignoring the RP Priority in RP-Set

By default, C-RP of the highest priority is configured.

Run **ip pim ignore-rp-set-priority** to select or deselect the RP priority when selecting the corresponding RP of a multicast group.

If you want to select an RP from multiples RPs that serve the same multicast group address, you can run this command to ignore the RP priority. If this command is not configured, RP priority will be considered when two RPs are compared.

6.3.5Register Information of the Multicast Source

When the multicast source is detected on the network, the source DR sends a register packet to the RP, which obtains the source information and multicast packet.

Working Principle

When a source DR receives a multicast packet from the host directly connected to it, the source DR encapsulates the multicast packet into the register packet, and sends the unicast packet to RP to form an (S, G) entry.

If the RP has an outgoing interface for the forwarding entry, it encapsulates the data packet and forwards the packet to the outgoing interface.

If the RP does not have the forwarding entry of the present group, it generates the (S, G) entry and enables the timer. If the timer times out, the RP sends a Register-Stop packet to the DR to delete the entry. The source DR sends an inspection packet before timeout after it receives the Register-Stop packet.

If no Register-Stop packet is received by the DR, the DR on the timeout data source will encapsulate the multicast data in the register packet and send the unicast packet to the RP.

If a Register-Stop packet is received by the DR, time-delay will be performed once again, and an inspection packet will be sent before time delay.

Related Configuration

Detecting the Reachability of a Register Packet

By default, the reachability of an RP is not detected.

Run **ip pim register-rp-reachability** to configure or cancel the detection of the reachability of an RP.

You can enable this function if you want to detect whether an RP is reachable for a register packet sent from a DR. After this function is enabled, the DR will detect the reachability of a register packet before it is sent to an RP, namely, the DR will check whether a route to the RP exists in the unicast routing entry and static multicast routing entry. If the route does not exist, the register packet will not be sent.

Configuring an RP to Filter the Addresses of Register Packets

By default, all register packets are received an RP.

Run **ip pim accept-register** { **list** *access-list* [**route-map** *map-name*] | **route-map** *map-name* [**list** *access-list*] } to configure an RP to filter or cancel the filtering of the source addresses of received register packets.

You can run this command if you want to filter the source addresses of received register packets. If this function is not enabled, all register packets will be received by the RP. If this function is disabled, only the register packets whose source addresses and multicast group addresses included in access control lists (ACLs) are processed; otherwise, the packets will be filtered.

Limiting the Speed for Sending a Register Packet

By default, the speed for sending a register packet is not limited.

Run **ip pim register-rate-limit** *rate* to limit or cancel the limitation of the speed for sending a register packet.

If the **no** form of this command is configured, the speed is not limited. This command takes effect for only the register packet of each (S, G) packet, but not all the register packets in the entire system.

Calculating the Checksum of the Entire Register Packet Length

By default, the checksum of a register packet is calculated as stipulated by the protocol.

Run **ip pim register-checksum-wholepkt [group-list *access-list*]** to configure the checksum of the register packet length.

You can enable this function if you want to include the length of encapsulated multicast packets into the checksum of the register packet length. If this function is disabled, the checksum of a register packet is calculated as stipulated by the protocol.

Configuring an RP to Forward Multicast Data Packets to Downstream Interfaces After Decapsulating Register Packets

By default, register packets are not decapsulated and multicast packet are not forwarded to interfaces.

Run **ip pim register-decapsulate-forward** to forward or cancel the forwarding of data packets to downstream interfaces.

You can run this command if you want to decapsulate a register packet and forward the multicast packet. If this function is disabled, the multicast packet will not be forwarded.

Configuring the Source IP Address of a Register Packet

By default, the source IP address of a register packet is the same as the interface address of the DR connected to the multicast source.

Run **ip pim register-source** { *local_address* | *Interface-type interface-number* } to configure the source IP address.

You can run this command is you want to configure the source IP address of the register packet sent by a DR. If this function is disabled or the **no** form of this command is used, the source address of the register packet will be the same as the interface address of the DR connected to the multicast source. If you want to configure *local_address*, the configured address must be reachable for a unicast route. *Interface-type interface-number* can be a typical a loopback interface or an interface of other types. The interface address must have been advertised by a unicast route.

Configuring the Suppression Time of a Register Packet

By default, the suppression time of a register packet is 60s.

Run **ip pim register-suppression** *seconds* to configure the suppression time.

If you run this command on a DR, you can change the suppression time of the register packets sent from the DR. If you run this command but does not run **ip pim rp-register-kat** on an RP, the keepalive period of the RP will be changed.

Configuring the Inspection Time of a Null Register Packet

By default, the inspection time is 5s.

Run **ip pim probe-interval** *interval-seconds* to configure the inspection time.

In the time interval before the timeout of register packet suppression, the source DR can send a null register packet to an RP. This time interval is called the inspection time, which is 5s by default.

Configuring the Time of a RP KAT

By default, the default value of a keepalive timer (KAT) is used. The default value is calculated as follows: Suppression time of a register packet x 3 + Inspection time of a null register packet.

Run **ip pim rp-register-kat** *seconds* to configure the KAT time.

You can run this command if you want to configure the keepalive time of (S, G) of a register packet sent from an RP.

6.3.6 Creating an RPT

When a group member is detected on the network, the DR connecting to the group members send packets toward the RP to form an RPT. If the multicast source already exists on the network, the packets arrived at the RP can be sent to the group members along the RPT.

Working Principle

To create an RPT, perform the following steps:

A receiver DR receives an IGMP (*, G) include report packet from the receiving end.

If the DR is not the RP of this group (G), the DR will send a (*, G) Join packet toward the RP. The router receiving this (*, G) Join packet will send the packet hop by hop until it is received by the RP, which means that the RP has joined the RPT.

When the data source host sends the multicast data to a group, the source data is encapsulated in the register packet, and sent from the source DR to the RP in unicast mode. Then the RP decapsulates the register packet, takes the data packets out, and forwards these packets to each group member along the RPT.

The RP sends the (S, G) Join packets along the data source to join the SPT of this source.

After the SPT between the RPs to the source DR is created, the data packets from the data source will be sent decapsulated to the RPs along the SPT.

When the first multicast data packet arrives at an RP along the SPT, the RP sends a Register-Stop packet to the source DR to stop sending a register packet. After the source DR receives the Register-Stop packet, it stops encapsulating a register packet and sends the packet along the SPT to the RP, which will forwards the packet to each group member.

Related Configuration

Configuring the Interval for Sending a Join/Prune Packet

By default, the interval for sending a Join/Prune packet is 60s.

Run **ip pim jp-timer** *seconds* to configure the interval for sending a Join/Prune packet.

You can run this command to configure the interval for sending a Join/Prune packet. If not configured, the value will be a default 60s.

6.3.7 Creating an SPT

When data packets arrive at the DR connecting to group members, the DR sends these packets toward the multicast source to form an SPT, and multicast packets are sent to group members along the SPT. In this way, the burden on RP in the RPT is reduced, and the source DR will arrive at the receiver DR with less hops.

Working Principle

To create an SPT, perform the following steps:

The receiver DR sends (*, G) Join packets toward the source DR along the SPT, and (*, G) Join packets are then send hop by hop until they are received by the source DR, forming an SPT.

Related Configuration

By default, SPT switchover is not enabled.

Run **ip pim spt-threshold** [**group-list** *access-list*] to configure whether to switch to an SPT.

If this function is enabled, upon the reception of the first (S, G) packet, a PIM Join packet is triggered, and an SPT is created. If **group-list** is specified, all the specified groups will be switched to the SPT. If the **no** form of this command is used and **group-list** is not specified, an RPT will not be switched to an SPT, and the DR will remain in the RPT and send a Prune packet toward the source DR; if the **no** form of this command is used and **group-list** is specified, and that the ACLs have been configured, it means that the association between **group-list** and the ACLs is canceled, and all the groups are allowed to switch from an RPT to an SPT.

6.3.8 ASM and SSM

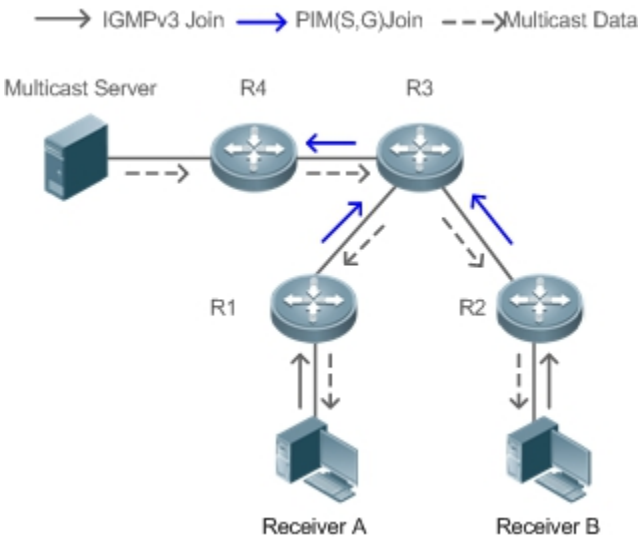
A PIM router can provide multicast services of both ASM model and SSM model at the same time. SSM model applies to the groups whose addresses are within the range of the SSM addresses. For other groups, use ASM model. In an ASM model, only the multicast group (G) is specified for a multicast receiver, and the multicast source (S) is not specified. In an SSM model, both the multicast source (S) and multicast group (G) can be specified for a multicast receiver.

Working Principle

To realize SSM in an IPv4 router, IGMPv3 needs to be applied for managing membership between the host and devices, and PIM-SM needs to be applied to connect to devices.

In an SSM model, as a multicast receiver has learnt the (S, G) of the multicast source through a certain channel (for example, by visiting the server or receiving an advertisement), when a multicast receiver needs to request a multicast service, the multicast receiver can send the IGMP (S, G) Join packet toward the router of last hop. For example, as shown in Figure 6-3, the multicast receiver A sends the IGMP (S, G) Join packet to request the multicast service (S, G). After the router of last hop receives the IGMP (S, G) Join packet, it sends the PIM (S, G) Join packet to the multicast source hop by hop. As shown in Figure 6- 3, when R 1 receives the IGMP (S, G) Join packet sent from multicast Receiver 1, R 1 sends the PIM (S, G) Join packet to R 3, which then sends the packet to R 4, thereby forming an SPT connecting the multicast receiver and multicast source.

Figure 6- 3 SSM Model



To create an SSM model, the following requirements need to be met:

- A multicast receiver needs to learn the (S, G) of the multicast source in advance, and an IGMP (S, G) Join packet needs to be sent if the receiver needs to request a multicast service.
- IGMPv3 must be run on the interface of the last hop router connecting to the multicast receiver. IGMPv1 and IGMPv2 does not support SSM.
- PIM-SM and SSM must be run on the devices connecting the multicast receiver and multicast source.

The default range of SSM groups is 232/8. You can run a command to change the value.

An SSM has the following features:

- A multicast receiver can learn the information of the multicast source through a certain channel (for example, by visiting the server or receiving an advertisement) in advance.
- An SSM model is a specific subnet of PIM-SM. It handles only the PIM (S, G) Join and PIM (S, G) Prune packets and discards the RPT-related packets, for example, PIM (*, G) Join/Prune packets, that are within the scope the SSM. If the SSM detects a register packet within the scope, it will respond immediately with a Register-Stop packet.
- If an RP is not required, the election and distribution of RP information are not performed. The MDTs in an SSM are all SPTs.

Related Configuration

ASM is enabled by default.

Run **ip pim ssm** { **default** | **range** *access-list* } to configure whether to switch to SSM.

In SSM, multicast packets can be received by the multicast source directly but not along the RP tree.

6.4 Configuration

Configuration	Description and Command	
Configuring Basic PIM-SM Functions	(Mandatory) It is used to configure the multicast service.	
	ip multicast-routing	Enables IPv4 multicast routing.
	ip pim sparse-mode	Enables PIM-SM.

Configuration	Description and Command	
	ip pim rp-address	Configures a static RP.
	ip pim rp-candidate	Configures a C-RP.
	ip pim bsr-candidate	Configures a C-BSR.
	ip pim ssm	Enables SSM.
Configuring PIM-SM Neighbors	(Optional) It is used to configure the parameters for sending and receiving the Hello packets between neighbors.	
	ip pim query-interval *interval-seconds*	Configures the interval for sending Hello packets.
	ip pim propagation-delay *milliseconds*	Configures the prune propagation delay.
	ip pim override-interval *milliseconds*	Configures the prune override interval.
	ip pim neighbor-tracking	Enables the suppression capability of an interface for sending Join packets.
	ip pim triggered-hello-delay *interval-seconds*	Configures the delay for sending Hello packets.
	ip pim dr-priority *priority-value*	Configures the DR priority of a Hello packet.
	ip pim neighbor-filter *access_list*	Configures neighbor filtering.
Configuring BSR Parameters	(Optional) It is used to configure a BSR.	
	ip pim bsr-border	Configures BSR boarders.
	ip pim accept-bsr list { <1-99> | <1300-1999> | *WORD* }	Configures BSM packets limit on a PIM router.
	ip pim accept-crp list *access-list*	Configures a C-BSR to inspect the address range of a C-PR.
Configuring RP and DR Parameters	(Optional) It is used to configure the parameters of an RP or a DR.	
	ip pim ignore-rp-set-priority	Ignores the C-RP priority.
	ip pim register-rp-reachability	Enables the source DR to detect the RP reachability.
	ip pim accept-register list *access-list*	Configures the range of source register (S, G) addresses.
	ip pim register-rate-limit *rate*	Limits the speed for sending register packets.
	ip pim register-checksum-wholepkt [**group-list** *access-list*]	Calculates the checksum of the entire register packet.
	ip pim register-decapsulate-forward	Enables an RP to decapsulate a register packet and forwards the multicast packet to interfaces.
	ip pim register-source { *local_address* | *Interface-type interface-number* }	Configures the source IP address of a register packet.

Configuration	Description and Command	
	ip pim register-suppression *seconds*	Configures the suppression time of a register packet.
	ip pim probe-interval *seconds*	Configures the inspection time of a null register packet.
	ip pim rp-register-kat *seconds*	Configures the interval of KATs on an RP.
Configuring the Interval for Sending a Join/Prune Packet	(Optional) It is used to specify the interval for sending a Join/Prune packet.	
	ip pim jp-timer *seconds*	Configures the interval for sending a Join/Prune packet.
Configuring the Router of Last Hop to Switch from an RPT to SPT	(Optional) It is used to switch from SPT to RPT.	
	ip pim spt-threshold [**group-list** *access-list*]	Enables SPT switchover.
Configuring PIM-SM PASSIVE mode	**ip pim sparse-mode passive**	Enables PIM-SM PASSIVE mode.
Configuring the PIM-SM Sub VLAN Function	**Ip pim sparse-mode subvlan** [**all** | *vid*]	Specifies, on an interface of a super VLAN, the sub VLAN to which packets are sent.

6.4.1Configuring Basic PIM-SM Functions

Configuration Effect

- Create a PIM-SM network and provide data sources and user terminals on the network with the IPv4 multicast service.
- Any of ASM or SSM or both models can be configured.

Notes

- PIM-SM needs to use existing unicast routes on the network. Therefore, IPv4 unicast routes must be configured on the network.
- If the PIM network needs to support SSM multicast services, IGMPv3 or SSM mapping must be configured.

Configuration Steps

Enabling IPv4 Multicast Routing

- Mandatory.
- If not specified, IPv4 multicast routing must be enabled on each router.

Enabling PIM-SM

- Mandatory.
- If not specified, PIM-SM must be enabled on the following interfaces: interconnecting router interfaces, interfaces of static RPs, C-RPs, and C-BSRs, and the interfaces connecting to the multicast source and user hosts.

Enabling the PIM-SM PASSIVE Function

- In a PIM network, if an interface needs to receive multicast packets without participating in the PIM network topology construction, the PIM-SM PASSIVE mode can be configured.

- If no special requirements are raised, enable the PIM-SM PASSIVE function on the following interfaces: interfaces of the stub network device in the multicast network for connecting to STAs. After the PIM-SM PASSIVE function is configured on an interface, the interface neither sends nor receives PIM packets.

Configuring an RP

- An RP must be configured if ASM multicast services need to be provided on a PIM network.
- An RP can be configured in three models: configuring only a static RP, configuring only a dynamic RP, and configuring both a static RP and dynamic RP. If both a static RP and dynamic RP are configured, the dynamic RP takes precedence over the static RP.
- Configuring a static RP: If not specified, a static RP should be configured on each router.
- Configuring a dynamic RP: If not specified, a C-RP and C-BSR should be configured on one or multiple routers.

Enabling SSM

- SSM must be enabled if SSM multicast services need to be provided on a PIM network.
- If not specified, SSM must be enabled on every router.

Configuring the PIM-SM Sub VLAN Function

- In general, a super VLAN includes many sub VLANs. If the PIM-SM protocol is enabled on the interfaces of the super VLAN, multicast packets will be replicated and sent to all sub VLANs. As a result, the traffic may exceed the device capability, causing protocol flapping. The Super VLAN interface is disabled with PIM-SM generally. Use this command to enable PIM-SM on the Super VLAN interface to send PIM packets to all sub VLANs or the specified sub VLAN.
- This function is available only on the Super VLAN interface.

Verification

Send multicast packets from the multicast source to the groups within the address rang of ASM and SSM, and join user hosts to these groups.

- Check whether the user hosts can successfully receive packets from each group.
- Check whether PIM-SM routing entries are created on routers correctly.

Related Commands

Enabling IPv4 Multicast Routing

Command	**ip multicast-routing**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling PIM-SM

Command	**ip pim sparse-mode**
Parameter	N/A

Description	
Command Mode	Interface configuration mode
Usage Guide	PIM interfaces must be at Layer-3, including: routing interfaces, aggregate ports(APs), switch virtual interfaces (SVIs), and loopback interfaces. For all PIM interfaces, IPv4 unicast routes should be reachable.

Enabling PIM-SM PASSIVE Mode

Command	**ip pim sparse-mode passive**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The PIM interface must be a Layer-3 interface, including: routing interface, aggregate port, switch virtual interface, and loopback interface. For all PIM interfaces, IPv4 unicast routes should be reachable.

Enabling the PIM-SM Sub VLAN Function

Command	**ip pim sparse-mode subvlan** [**all** | *vid*]
Parameter Description	**all**: sends packets to all sub VLANs. *vid*: sends packets to a specified sub VLAN.
Command Mode	Interface configuration mode
Usage Guide	The PIM interface must be a Layer-3 interface, including: routing interface, aggregate port, switch virtual interface, and loopback interface.

Configuring a Static RP

Command	**ip pim rp-address** *rp-address* [*access_list*]
Parameter Description	*rp-address*: Indicates the address of an RP. *access_list*: Specifies the range of multicast group addresses served by a static RP using an ACL. By default, an RP services all groups.
Command Mode	Global configuration mode
Usage Guide	This command is used to locate a static RP. A static RP should be one with good routing performance. It is recommended that the address of the loopback interface be used as the static RP address.

	The static RP of all routers must be the same (including the RP address and the range of multicast group addresses it serves). It is recommended that the address of the loopback interface be used as the static RP address. The load can be shared if you configure multiple static RPs to serve different multicast group addresses. It is recommended that the address of the loopback interface be used as the static RP address.

Configuring a C-RP

Command	**ip pim rp-candidate** *interface-type interface-number* [**priority** *priority-value*] [**interval** *seconds*] [**group-list** *access_list*]
Parameter Description	*interface-type interface-number*: Uses the address of this interface as the address of the C-RP. **priority** *priority-value*: Competes for the RP priority. A greater value indicates a higher priority. The value ranges from 0 to 255 (192 by default). **interval** *seconds*: Indicates the interval for sending a C-RP packet to a BSR. The value ranges from 1 to 16,383 (60 by default). **group-list** *access_list*: Specifies the range of multicast group addresses served by a C-RP using an ACL. By default, a C-RP services all multicast groups.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure a router as a C-RP. A C-RP should be one with good routing performance. A C-RP and C-BSR can be on the same router or different routers. It is recommended that the address of the loopback interface be used as the C-RP address. If multiple C-RPs serve the same group, redundancy can be realized. If multiple C-RPs serve the different groups, load can be shared.

Configuring a C-BSR

Command	**ip pim bsr-candidate** *interface-type interface-number* [*hash-mask-length* [*priority-value*]]
Parameter Description	*interface-type interface-number*: Uses the address of this interface as the address of the C-BSR. *hash-mask-length*: Indicates the length of hash mask used to competing for the RP. The value ranges from 0 to 32 (10 by default). *priority-value*: Indicates the priority for competing for the BSR. A greater value indicates a higher priority. The value ranges from 0 to 255 (64 by default).
Command Mode	Global configuration mode
Usage Guide	This command is used to configure a router as a C-BSR. A C-BSR should be one with good routing performance. A C-RP and C-BSR can be on the same router or different routers. It is recommended that the address of the loopback interface be used as the C-BSR address. Configuring multiple C-BSRs can realize redundancy.

Enabling SSM

Command	**ip pim ssm** { **default** | **range** *access_list* }
Parameter Description	**default**: Indicates the default range of SSM group addresses, which is 232.0.0.0/8. **range** *access_list*: Specifies the range of SSM group addresses using an ACL.
Command Mode	Global configuration mode

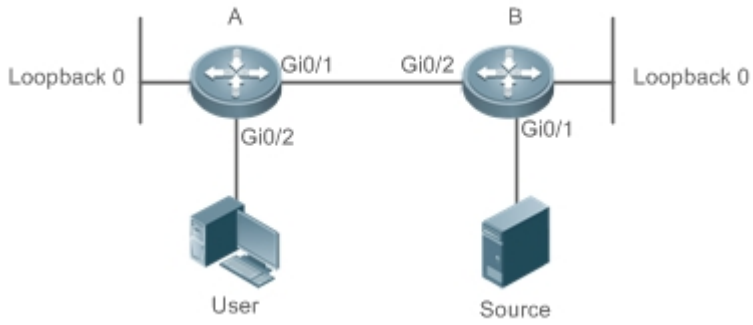
Usage Guide	The SSM group addresses configured on all routers must be the same.

Displaying the PIM-SM Routing Entry

Command	**show ip pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]] [**proxy**]
Parameter Description	*group-or-source-address*: Indicates a multicast group address or source address (the two addresses cannot be multicast group addresses or source addresses at the same time). **proxy**: Indicates the RPF vector carried by an entry.
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	Check whether sufficient routing entries are provided. Check the upstream and downstream interface lists and ensure that a correct SPT tree is created.

Configuration Example

Enabling IPv4 Multicast Routing to Support ASM and SSM

Scenario **Figure 6- 4**	
Configuration Steps	● Configure a IPv4 unicast routing protocol (such as OSPF) on a router, and the router is reachable for the unicast route of a loopback interface. (Omitted) ● Enable IPv4 multicast routing on all the routers. ● Enable PIM-SM on all the interconnected interfaces of the routers, Source, and Receiver. ● Configure C-RP and C-BSR on the loopback interfaces of Router A and Router B, and enable PIM-SM on the loopback interfaces. ● Enable SSM on all routers. ● Enable IGMPv3 on the router interfaces connecting to user terminals. (Omitted)
A	A# configure terminal A(config)# ip multicast-routing A(config)# ip pim ssm default A(config)# interface GigabitEthernet 0/1 A(config-if)# ip pim sparse-mode A(config-if)# exit

	A(config)# interface GigabitEthernet 0/2 A(config-if)# ip pim sparse-mode A(config-if)# exit A(config)# interface loopback 0 A(config-if)# ip pim sparse-mode A(config-if)# exit A(config)# ip pim rp-candidate loopback 0
B	B# configure terminal B(config)# ip multicast-routing B(config)# ip pim ssm default B(config)# interface GigabitEthernet 0/1 B(config-if)# ip pim sparse-mode B(config-if)# exit B(config)# interface GigabitEthernet 0/2 B(config-if)# ip pim sparse-mode B(config-if)# exit B(config)# interface loopback 0 B(config-if)# ip pim sparse-mode B(config-if)# exit B(config)# ip pim bsr-candidate loopback 0
Verification	Send packets from S (192.168.1.10) to G 1 (229.1.1.1) and G2 (232.1.1.1). Add the user to G 1 and G 2, and specify the source when the user joins G 2. ● Check that multicast packets from S (192.168.1.10) to G 1 and G 2 are received by the user. ● Check the PIM-SM routing entries on Router A and Router B. Entries (*, 229.1.1.1), (192.168.1.10, 229.1.1.1), and (192.168.1.10, 232.1.1.1) should be displayed.
A	switch#show ip pim sparse-mode mroute IP Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 3 (S,G) Entries: 2 (S,G,rpt) Entries: 2 FCR Entries: 0 REG Entries: 0

```
(*, 229.1.1.1)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .   .   .   i   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
1   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
Joined
0   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
1   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
Asserted
0   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
1   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
FCR:

(192.168.1.10, 229.1.1.1)
RPF nbr: 192.168.2.1
RPF idx: GigabitEthernet 0/2
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 8 seconds
kat expires in 207 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
1   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
Joined
0   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
1   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
Asserted
0   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
```

```
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . . . o . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

(192.168.1.10, 229.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: PRUNED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Pruned
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . . . o . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

(*, 232.1.1.1)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . i . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Joined
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Asserted
```

```
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
FCR:
(192.168.1.10, 232.1.1.1)
RPF nbr: 192.168.2.1
RPF idx: GigabitEthernet 0/2
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 8 seconds
kat expires in 207 seconds
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Joined
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Asserted
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . . . o . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

(192.168.1.10, 232.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: PRUNED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Pruned
```

	0 . . . . . . 1 . . . . . . Outgoing 0 . . . o . . . . . . 1 . . . . . . (*, 239.255.255.250) RP: 192.168.10.10 RPF nbr: 0.0.0.0 RPF idx: None Upstream State: JOINED 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local 0 . . . i . . . . . . 1 . . . . . . Joined 0 . j . . . . . . 1 . . . . . . Asserted 0 . . . . . . 1 . . . . . . FCR: A#
B	B#show ip pim sparse-mode mroute IP Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 1

```
(192.168.1.10, 229.1.1.1)
RPF nbr: 0.0.0.0
RPF idx: None
SPT bit: 1
Upstream State: JOINED
kat expires in 38 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  .  j  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0  .  .  o  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

(192.168.1.10, 229.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 192.168.2.2
RPF idx: GigabitEthernet 0/2
Upstream State: RPT NOT JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pruned
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

(192.168.1.10, 232.1.1.1)
RPF nbr: 0.0.0.0
RPF idx: None
SPT bit: 1
Upstream State: JOINED
```

```
kat expires in 38 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  .  j  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0  .  .  o  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

(192.168.1.10, 232.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 192.168.2.2
RPF idx: GigabitEthernet 0/2
Upstream State: RPT NOT JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pruned
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
(*, 239.255.255.250)
RP: 192.168.10.10
RPF nbr: 192.168.2.2
RPF idx: GigabitEthernet 0/2
Upstream State: JOINED
jt_timer expires in 15 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  i  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
```

	Asserted 0 . . . . . . FCR:

Common Errors

- IPv4 unicast routing is incorrectly configured.
- IPv4 multicast routing is not enabled on a certain router.
- SSM is not enabled on a router or the SSM group address is different from that of the others'.
- PIM-SM is not enabled on an interface (for example, the interface is configured as a C-RP or C-BSR interface, or is used to connecting to the user host or used as an interface of the multicast source).
- IGMPv3 is not enabled on an interface connecting to the used host.
- RP is not configured on the network.
- A static RP is not configured on a router, or the configured static RP is different from that on other routers.
- C-RPs are configured on the network, but C-BSRs are not.
- Static RPs, C-RPs or C-BSRs are unreachable for unicast routes.

6.4.2 Configuring PIM-SM Neighbors

Configuration Effect

- Coordinate protocol parameters and adjust parameters in the Hello packet.
- A RIM router is used to discover neighbors, coordinate protocol parameters, and maintain neighbor relationships.
- Maintain neighbor relationships and filter the neighbors.

Notes

- Basic PIM-SM functions must be configured.

Configuration Steps

- Configure parameters on PIM router interfaces If not specified.

Verification

Configure the parameters of a Hello packet sent from an interface and run **debug ip pim sparse-mode packet** to display the parameters.

Enable neighbor filtering and run **show ip pim sparse-mode neighbor** to display neighbor information.

Related Commands

↘ Configuring the Interval for Sending Hello Packets

Command	**ip pim query-interval** *interval-seconds*
Parameter Description	Indicates the interval for sending Hello packets, Indicates the suppression time of a register packet in the unit of seconds. The value ranges from 1 to 65,535 (30 by

	default).
Command Mode	Interface configuration mode
Usage Guide	Every time when the interval for sending Hello packets is updated, the holdtime value is automatically updated as 3.5 times of the interval. If the result of the interval for sending Hello packets multiplied by 3.5 is greater than 65,535, the holdtime value is forcibly updated as 18,725.

Configuring the Prune Propagation Delay

Command	**ip pim propagation-delay** *milliseconds*
Parameter Description	*milliseconds*: The unit is ms. The value ranges from 1 to 32,767 (500 by default).
Command Mode	Interface configuration mode
Usage Guide	Once the prune propagation delay or prune override interval is changed, the Join/Prune packet override interval will be changed. As specified by the protocol, the Join/Prune packet override interval must be smaller than the holdtime of a Join/Prune packet; otherwise, short break-up of traffic may be caused. The administrator should maintain such configuration.

Configuring the Prune Override Interval

Command	**ip pim override-interval** *milliseconds*
Parameter Description	*milliseconds*: The unit is ms. The value ranges from 1 to 65,535 (2,500 by default).
Command Mode	Interface configuration mode
Usage Guide	Once the prune propagation delay or prune override interval is changed, the Join/Prune packet override interval will be changed. As specified by the protocol, the Join/Prune packet override interval must be smaller than the holdtime of a Join/Prune packet; otherwise, short break-up of traffic may be caused. The administrator should maintain such configuration.

Enabling Suppression Capability of an Interface for Sending Join Packets

Command	**ip pim neighbor-tracking**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Once Join packets suppression of an interface is enabled, when the present router is to send a Join packet to the upstream neighbor, which has sent a Join packet to its own upstream neighbor, the present router will not send the Join packet; if Join packets suppression is disabled, the Join packet will be sent. When Join packets suppression from downstream receivers are disabled, upstream neighbors will learn how many downstream neighbors are there by counting the Join packets it received, which is called neighbor tracking.

Configuring the Delay for Sending Hello Packets

Command	**ip pim triggered-hello-delay** *interval-seconds*
Parameter Description	*Seconds*: The unit is second. The value ranges from 1 to 5 (5 by default).
Command Mode	Interface configuration mode
Usage Guide	When a PIM interface is enabled or detects a new neighbor, a triggered-hello-delay packet is used to generate a random time. Within the time, the interface sends Hello packets.

Configuring the DR Priority of a Hello Packet

Command	**ip pim dr-priority** *priority-value*
Parameter Description	*priority-value*: Indicates the priority. A greater value indicates a higher priority. The value ranges from 0 to 4,294,967,294 (1 by default).
Command Mode	Interface configuration mode
Usage Guide	A DR may be selected based on the following principles: If all the Hello packets sent from the routers on a local area network (LAN) are configured with priorities, when selecting a DR, the priorities will be compared, and the router with the highest priority will be selected as the DR. If the priority of all routers is the same, their IP addresses will be compared, and the router with the maximum IP address will be selected as the DR. If the priority of the Hello packets sent from a certain router is not configured, the IP addresses of the routers will be compared, and the router with the maximum IP address will be selected as the DR.

Configuring Neighbor Filtering

Command	**ip pim neighbor-filter** *access_list*
Parameter Description	*access_list*: Configures the range of neighbor addresses using a standard IP ACL. The value can be set from 1 to 99 or a string.
Command Mode	Interface configuration mode
Usage Guide	Enabling neighbor filtering can enhance the security of the PIM network and limit the range of legible neighbor addresses. Once a neighbor is filtered out, PIM-SM will not establish peering with it or stop the peering with it.

Displaying the Neighbor Information of an Interface

Command	**show ip pim sparse-mode neighbor** [**detail**]
Parameter Description	**detail**: Displays detailed information.
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuration	● Configure basic PIM-SM functions. (Omitted)

Steps	● Configure the interval for sending Hello packets as 50s. ● Configure the prune propagation delay as 400 ms. ● Configure the prune override interval as 3,000 ms. ● Enable suppression capability of an interface for sending Join packets. ● Configure the delay for sending Hello packets as 3s. ● Configure the DR priority of a hello packet as 5.
	FS# configure terminal FS (config)#int gi 0/1 FS (config-if-GigabitEthernet 0/1)#ip pim query-interval 50 FS (config-if-GigabitEthernet 0/1)#ip pim propagation-delay 400 FS (config-if-GigabitEthernet 0/1)#ip pim override-interval 3000 FS (config-if-GigabitEthernet 0/1)#ip pim triggered-hello-delay 3 FS (config-if-GigabitEthernet 0/1)#ip pim neighbor-tracking
Verification	Run **debug ip pim sparse-mode packet** to display the parameters of a Hello packet.
	FS# debug ip pim sparse-mode packet 00:01:49:43: %7: VRF(0): Hello send to GigabitEthernet 0/1 00:01:49:43: %7: Send Hello packet 00:01:49:43: %7: Holdtime: 175 00:01:49:43: %7: T-bit: on 00:01:49:43: %7: Propagation delay: 400 00:01:49:43: %7: Override interval: 3000 00:01:49:43: %7: DR priority: 5 00:01:49:43: %7: Gen ID: 355154648 00:01:49:43: %7: RPF Vector capable
Configuration Steps	Configure neighbor filtering and set the allowed address range to 192.168.1.0 to 192.168.1.255.
	FS# configure terminal FS (config)#int gi 0/1 FS (config-if-GigabitEthernet 0/1)# ip pim neighbor-filter 1 % access-list 1 not exist FS(config)# access-list 1 permit 192.168.1.0 0.0.0.255 FS(config)#
Verification	Display neighbor information before neighbor filtering is configured.

	FS# show ip pim sparse-mode neighbor Neighbor Address Interface Uptime/Expires Ver DR Priority/Mode 192.168.36.89 GigabitEthernet 0/1 01:12:13/00:01:32 v2 1 / P
	Display neighbor information after neighbor filtering is configured.
	FS# show ip pim sparse-mode neighbor

Common Errors

- Basic PIM-SM functions are not configured or the configuration is not successful.

6.4.3 Configuring BSR Parameters

Configuration Effect

- Configure the address range of BSM packets.

Notes

- Basic PIM-SM functions must be configured.
- C-RPs and C-BSRs must be configured.
- Boarders must be configured on the interfaces between domains.

Configuration Steps

Configuring Boarders

- Boarders must be configures if there are multiple domains.
- Boarders are configured on the interfaces separating two domains.

Configuring BSM Packets Limit on a PIM Router

- Optional.
- If not specified, BSM packets limit can be configured on all PIM routers.

Configuring a C-BSR to Inspect the Address Range of a C-PR

- Optional.
- If not specified, C-PR range inspection can be configured on all C-BSRs.

Allowing a C-BSR to Receive a C-RP-ADV Packet Whose Prefix-Count Is 0

- Optional.
- If not specified, this function can be configured on all C-BSRs.

Verification

Border Inspection

Enable basic PIM-SM functions. Configure two routers to be in different domains, configure Router B as the C-BSR, and Router A to receive BSM packets.

Configure the junction of Router A and Router B as the border so that Router A does not receive BSM packets.

Configuring to Inspect BSM Packets Limit on a PIM Router

When basic PIM-SM functions are enabled, and Router B is set as the C-BSR, Router A can receive BSM packets. When the address range of C-BSRs are limited on Router A, BSM packets will not be received by Router A.

Configuring a C-BSR to Inspect the Address Range of a C-PR

When basic PIM-SM functions are enabled, Router B is set as the C-BSR, and Router A as the C-RP, if the address range of the C-RPs is limited on C-BSR, Router B will not receive the rackets sent from the C-RPs.

Related Commands

Configuring BSR Boarders

Command	**ip pim bsr-border**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	To prevent BSM flooding, you can configure a BSR boarder on an interface, so that the BSM packets arriving at this interface will be discarded but not forwarded.

Configuring BSM Packets Limit on a PIM Router

Command	**ip pim accept-bsr list** { <1-99> | <1300-1999> | *WORD* }
Parameter Description	**list** *access-list*: Configures the range of BSR addresses using a standard IP ACL. The value can be 1 to 99, 1,300 to 1,999, or a string.
Command Mode	Global configuration mode
Usage Guide	After this function is enabled, PIM-SM routers receive only the BSM packets sent from legible BSRs.

Configuring a C-BSR to Inspect the Address Range of a C-PR

Command	**ip pim accept-crp list** *access-list*
Parameter Description	**list** *access-list*: Specifies the range of C-RP addresses and the multicast group addresses they serve using an extended IP ACL. The value can be 100 to 199, 2,000 to 2,699, or a string.
Command Mode	Global configuration mode
Usage Guide	This command should be configured on a C-BSR. When the C-BSR becomes a BSR, it can set the range of legible C-RP addresses and the range of multicast group addresses they serves.

Displaying BSM Packets Information

Command	**show ip pim sparse-mode bsr-router**

Parameter Description	-
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A

Displaying the Packets of All RPs and the Multicast Group Addresses They Serve

Command	**show ip pim sparse-mode rp mapping**
Parameter Description	-
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring BSR Boarders

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● On Router A, configure the address of the C-RP as 192.168.8.8. ● On Router B, configure the address of the C-RP as 192.168.5.5, and the address of the C-BSR as 192.168.6.6. ● Configure a BSR boarder on the junction of Router A and Router B.
	FS# configure terminal FS(config)# int GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ip pim bsr-border FS(config)# end
Verification	Before configuring the boarder, display the BSM information on Router A.
	FS# show ip pim sparse-mode bsr-router PIMv2 Bootstrap information This system is the Bootstrap Router (BSR) BSR address: 192.168.6.6 Uptime: 01:14:25, BSR Priority: 64, Hash mask length: 10 Next bootstrap packet in 00:00:52 Role: Candidate BSR Priority: 64, Hash mask length: 10 State: Elected BSR Candidate RP: 192.168.8.8(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:06

	Candidate RP: Indicates all the C-RPs configured on the existing router. It does not include the C-RPs configured on other routers.
	After the boarder is configured, display the BSM information on Router A.
	FS# show ip pim sparse-mode bsr-router

Configuring BSM Packets Limit on a PIM Router, Filtering BSM Source Addresses, and Configuring the Range of BSM Source Addresses to 192.168.1.1 to 192.168.1.255

Configuration Steps	• Configure basic PIM-SM functions. (Omitted) • On Router A, configure the address of the C-RP as 192.168.8.8. • On Router B, configure the address of the C-RP as 192.168.5.5, and the address of the C-BSR as 192.168.6.6. • On Router A, configure the range of allowed BSM source addresses to 192.168.1.1 to 192.168.1.255.
	FS# configure terminal FS(config)# ip pim accept-bsr list 1 % access-list 1 not exist FS(config)# access-list 1 permit 192.168.1.0 0.0.0.255 FS(config)#
Verification	Before configuring BSM packets limit, display the BSM information on Router A.
	FS#show ip pim sparse-mode bsr-router PIMv2 Bootstrap information BSR address: 192.168.6.6 Uptime: 00:00:11, BSR Priority: 64, Hash mask length: 10 Expires: 00:01:59 Role: Non-candidate BSR Priority: 0, Hash mask length: 10 State: Accept Preferred Candidate RP: 192.168.8.8(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:06
	After BSM packets limit is configured, display the BSM information on Router A.
	FS# show ip pim sparse-mode bsr-router Candidate RP: 192.168.8.8(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:06

Configuring a C-BSR to Inspect the Address Range of a C-PR, Filtering C-RP Addresses, and Configuring the Range of C-RP Addresses to 192.168.1.1 to 192.168.1.255

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● On Router A, configure the address of the C-RP as 192.168.8.8. ● On Router B, configure the address of the C-RP as 192.168.5.5, and the address of the C-BSR as 192.168.6.6. ● On Router B, configure the range of allowed C-RP source addresses to 192.168.1.1 to 192.168.1.255.
	FS# configure terminal FS(config)# ip pim accept-crp list 100 % access-list 1 not exist FS(config)# access-list 1 permit 192.168.1.0 0.0.0.255 FS(config)#
Verification	Before configuring C-RP filtering, display the information of all RP groups on Router B.

	FS#show ip pim sparse-mode rp mapping
	PIM Group-to-RP Mappings
	This system is the Bootstrap Router (v2)
	Group(s): 224.0.0.0/4
	  RP: 192.168.8.8(Not self)
	  Info source: 192.168.8.8, via bootstrap, priority 192
	  Uptime: 00:15:16, expires: 00:02:18
	  RP: 192.168.5.5(Self)
	  Info source: 192.168.6.6, via bootstrap, priority 192
	  Uptime: 18:52:30, expires: 00:02:00

	After C-RP filtering is configured, display the information of all RP groups on Router B.

	FS#show ip pim sparse-mode rp mapping
	PIM Group-to-RP Mappings
	This system is the Bootstrap Router (v2)
	Group(s): 224.0.0.0/4
	  RP: 192.168.5.5(Self)
	  Info source: 192.168.6.6, via bootstrap, priority 192
	  Uptime: 21:38:20, expires: 00:02:10

	⚠ After C-RP filtering is configured on a router, only the C-RP packets sent from other routers are filtered, and those sent from the present router are not filtered.

Common Errors

- Basic PIM-SM functions are not configured or the configuration is not successful.
- C-BSRs are not configured.
- The BSR border is not configured on the interfaces of different domains.

6.4.4 Configuring RP and DR Parameters

Configuration Effect

- Ignore the C-RP priority and reselect an RP.
- Detect the reachability of an RP for the source DR.
- Configure the range of (S, G) addresses of source register packets, and allow the ASM to serve only the multicast packets within the range.
- Limit the speed of the source DR for sending register packets.
- Configure the checksum of the register packet length.
- Configure an RP to decapsulate register packets and forward the multicast packets to downstream interfaces.
- Configure the source IP address of a register packet.
- Configure the suppression time of a register packet.
- Configure the inspection time of a null register packet.
- Configure the (S, G) lifetime based on the register packet received by an RP.

Notes

- Basic PIM-SM functions must be configured.

Configuration Steps

Ignoring the C-RP Priority and Reselecting an RP

- Optional.
- If not specified, the C-RP priority can be disabled on every router.

Detecting the Reachability of an RP for the Source DR

- Optional.
- If not specified, this function can be enabled on the DR connected directly to the data source.

Configuring the Range of Source Register (S, G) Addresses

- Optional.
- If not specified, source register address filtering can be enabled on all C-RPs or static RPs.

Limiting the Speed of the Source DR for Sending Register Packets

- Optional.
- If not specified, this function can be enabled on the source DR.

Configuring the Checksum of the Register Packet Length

- Optional.
- If not specified, this function can be enabled on all C-RPs or static RPs.

Configuring Whether to Forward the Multicast Packet After Decapsulating a Register Packet

- Optional.
- If not specified, this function can be enabled on all C-RPs or static RPs.

Configuring the Source IP Address of a Register Packet

- Optional.
- If not specified, the source IP address of a register packet can be configured on the DR connected directly to the data source.

Configuring the Suppression Time of a Register Packet

- Optional.
- If not specified, the suppression time of a register packet can be configured on the DR connected directly to the data source.

Configuring the Inspection Time of a Null Register Packet

- Optional.
- If not specified, the inspection time of a null register packet can be configured on the DR connected directly to the data source.

Configuring the (S, G) Lifetime Based on the Register Packet Received by an RP

- Optional.
- If not specified, the (S, G) lifetime can be configured on all C-RPs or static RPs.

Verification

Ignoring the C-RP priority

On Router A, configure the C-RP address as 192.168.8.8, and default priority as 192. On Router B, configure the C-RP address as 192.168.5.5, priority as 200, and C-BSR address as 192.168.6.6.

- Run **show ip pim sparse-mode rp** 233.3.3.3 to display the RPs of the present group.

Enabling the Source DR to Detect RP Reachability

On Router A, configure the C-RP address as 192.168.8.8, and default priority as 192. On Router B, configure the C-RP address as 192.168.5.5, priority as 192, and C-BSR address as 192.168.6.6. Enable Router B to detect RP reachability.

- Run **show running-config** to check whether the preceding configurations take effect.

Configuring the Range of Source Register (S, G) Addresses

On Router A, configure the C-RP address as 192.168.8.8, and default priority as 192. On Router B, configure the address of the C-BSR as 192.168.6.6. Configure the source address a 192.168.1.100 and the multicast group address as 233.3.3.3. On Router A, configure the range of allowed source multicast group addresses to 192.168.2.0 to 192.168.2.255.

- Run **show ip pim sparse-mode mroute** to display the (S, G) entry.

Limiting the Speed of the Source DR for Sending Register Packets

Configure the speed of Router B for sending register packets, and run **show ip pim sparse-mode track** to display the number of packets that has been sent.

Configuring the Checksum of the Register Packet Length

On Router A, configure to calculate the checksum of the entire register packet length but not just the packet header. Run **show running-config** to check the configuration.

Forwarding an RP Register Packet After It Is Decapsulated

On Router A, configure to forward a register packet after it is decapsulated. Run **show running-config** to display the configuration.

Configuring the Source IP Address of a Register Packet

Configure the source address of a register packet on Router B, and run **show running-config** to display the configuration.

Configuring the Suppression Time of a Register Packet and the Inspection Time of a Null Register Packet

On Router B, configure the suppression time and inspection time of a register packet, and run **show ip pim sparse-mode track** to display the configuration.

Configuring an RP to Receive Register Packets and the (S, G) Lifetime

On Router A, configuring an RP to receive register packets and the (S, G) lifetime, and run **show ip pim sparse-mode mroute** to display the maximum (S, G) lifetime.

Related Commands

Ignoring the C-RP priority

Command	**ip pim ignore-rp-set-priority**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Displaying the RP Corresponding to a Group

Command	**show ip pim sparse-mode rp-hash** *group-address*
Parameter Description	*group-address*: Indicates the parsed multicast group address.
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	N/A

Enabling the Source DR to Detect RP Reachability

Command	**ip pim register-rp-reachability**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	After this function is enabled, the source DR will detect the RP reachability before sending a register packet. If the RP is unreachable, the packet will not be sent.

Configuring the Range of Source Register (S, G) Addresses

Command	**ip pim accept-register** { **list** *access-list* [**route-map** *map-name*] | **route-map** *map-name* [**list** *access-list*] }
Parameter Description	**list** *access-list*: Configures the range of (S, G) addresses using an extended IP ACL. The value can be 100 to 199, 2,000 to 2699, or a string. **route-map** *map-name*: Configures the range of (S, G) addresses using a route map.
Command Mode	Global configuration mode
Usage Guide	This command is run on a static RP or a C-RP to specify the source address and multicast group address of a register packet.

Displaying a Multicast Routing Entry

Command	**show ip pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]]
Parameter Description	*group-or-source-address*: Indicates a multicast group address or source address (the two addresses cannot be multicast group addresses or source addresses at the same time).
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	You can specify either a multicast group address or source address, or both a multicast group address and source address; or you can specify neither a multicast group address nor source address. The two addresses cannot be multicast group addresses or source addresses at the same time.

Limiting the Speed of the Source DR for Sending Register Packets

Command	**ip pim register-rate-limit** *rate*
Parameter Description	*Rate*: Indicates the maximum number of register packets that can be sent each second. The value ranges from 1 to 65,535.
Command Mode	Global configuration mode
Usage Guide	This command takes effect for only the register packet of each (S, G) packet, but not all the register packets in the entire system. Enabling this command can reduce the burden on the source DR and RPs. Only the packets within the speed limit can be sent.

Displaying the Counters of PIM-SM Packets

Command	**show ip pim sparse-mode track**
Parameter	-

Description	
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	The start time for counting PIM-SM packets is automatically enabled upon system startup. Run **clear ip pim sparse-mode track** to reset the start time and clear the PIM-SM packet counters.

Calculating the Checksum of the Entire Register Packet Length

Command	**ip pim register-checksum-wholepkt** [**group-list** *access-list*]
Parameter Description	**group-list** *access-list*: Configures the multicast group addresses applicable to this configuration using an ACL. access-list: The value can be set to 1 to 99, and 1300 to 1999. It also supports the naming of the ACL.
Command Mode	Global configuration mode
Usage Guide	You can enable this function if you want to calculate the length of the entire PIM-SM packet, including that of the multicast packet encapsulated in the register packet, but not just the length of the PIM-SM packet header. If **group-list** *access-list* is specified, this configuration takes effect for all multicast group addresses.

Enabling an RP to Decapsulate a Register Packet and Forward the Multicast Packet to Interfaces

Command	**ip pim register-decapsulate-forward**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	This command is configured on a static RP or a C-RP. It is used to decapsulate a register packet with multicast packet and forward the multicast packet to interfaces. If there are too many register packets to be decapsulated, the CPU will be greatly burdened. In this case, this function is recommended to be disabled.

Configuring the Source IP Address of a Register Packet

Command	**ip pim register-source** { *local_address* | *Interface-type interface-number* }
Parameter Description	*local_address*: Specifies the source IP address of a register packet. *interface-type interface-number*: Specifies the IP address of this interface as the source IP address of the register packet.
Command Mode	Global configuration mode
Usage Guide	The specified address must be reachable. When an RP sends a Register-Stop packet, the PIM router corresponds to this address need to respond. Therefore, it is recommended that a loopback address (or other physical addresses) be used. This configuration does not require the enabling of PIM.

Configuring the Suppression Time of a Register Packet

Command	**ip pim register-suppression** *seconds*
Parameter Description	*Seconds*: Indicates the suppression time of a register packet in the unit of seconds. The value ranges from 1 to 65,535 (60 by default).

Command Mode	Global configuration mode
Usage Guide	If you configure this parameter on a DR, the suppression time of a register packet sent from the DR will be changed. If **ip pim rp-register-kat** is not configured and if you configure this parameter on an RP, the RP keepalive will be changed.

Configuring the Inspection Time of a Null Register Packet

Command	**ip pim probe-interval** *seconds*
Parameter Description	**vrf** *vid*: Specifies VRF. *Seconds*: Indicates the inspection time of a null register packet in the unit of seconds. The value ranges from 1 to 65,535 (5 by default).
Command Mode	Global configuration mode
Usage Guide	The inspection time of a null register packet indicates the period of time for sending a null register packet to an RP before the timeout of suppression time. The inspection time cannot exceed half of the suppression time; otherwise, the configuration will not take effect, and a warning message will be displayed. Meanwhile, the result of suppression time multiplied by 3 plus the inspection time cannot exceed 65,535, otherwise, a warning will be displayed.

Configuring the Interval of KATs on an RP

Command	**ip pim rp-register-kat** *seconds*
Parameter Description	*Seconds*: Indicates the interval of a KAT in the unit of second. The value ranges from 1 to 65,535 (210 by default).
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the RPs of Corresponding Multicast Group Addresses When the C-RP Priority is Considered or Not Considered

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● On Router A, configure the address of the C-RP as 192.168.8.8. ● On Router B, configure the address of the C-RP as 192.168.5.5, priority as 200, and the address of the C-BSR as 192.168.6.6. ● Display the group corresponding to 233.3.3.3. ● Configure to ignore the C-RP priority on Router B.
	FS# configure terminal FS(config)# ip pim ignore-rp-set-priority
Verification	Display the information before you configure to ignore the C-RP priority.
	FS# show ip pim sparse-mode rp-hash 233.3.3.3 RP: 192.168.8.8

	Info source: 192.168.8.8, via bootstrap PIMv2 Hash Value 10(mask 255.192.0.0) RP 192.168.8.8, via bootstrap, priority 192, hash value 1084558102 RP 192.168.5.5, via bootstrap, priority 200, hash value 1094656709
	Display the information after you configure to ignore the C-RP priority.
	FS# show ip pim sparse-mode rp-hash 233.3.3.3 RP: 192.168.5.5 Info source: 192.168.6.6, via bootstrap PIMv2 Hash Value 10(mask 255.192.0.0) RP 192.168.8.8, via bootstrap, priority 192, hash value 1084558102 RP 192.168.5.5, via bootstrap, priority 200, hash value 1094656709

Configuring to Inspect the Reachability of a Source RP

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Configure to inspect the reachability of a source RP.
	FS(config)# ip pim register-rp-reachability
Verification	Run **show running-config** to check whether the following information is displayed.
	FS(config)#show running-config ip pim register-rp-reachability

Configuring the Range of Source Register (S, G) Addresses

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Configure source address filtering on Router A. The allowed address range is from 192.168.2.0 to 192.168.2.255.
	FS#show ip pim sparse-mode mroute FS(config)#ip pim accept-register list 101 % access-list 101 not exist FS(config)#access-list 101 permit ip 192.168.2.0 0.0.0.255 any FS#show ip pim sparse-mode mroute
Verification	Before enabling source address filtering, run **show ip pim sparse-mode mroute** to display the multicast entry, and check whether the (S, G) entry and (S, G, RPT) entry exist.
	FS#show ip pim sparse-mode mroute IP Multicast Routing Table

```
(*,*,RP) Entries: 0
(*,G) Entries: 1
(S,G) Entries: 1
(S,G,rpt) Entries: 1
FCR Entries: 0
REG Entries: 0

(192.168.1.100, 233.3.3.3)
RPF nbr: 192.168.36.90
RPF idx: VLAN 1
SPT bit: 0
Upstream State: NOT JOINED
kat expires in 187 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

(192.168.1.100, 233.3.3.3, rpt)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: RPT NOT JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pruned
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
```

	Outgoing 0 . . . . . . (*, 239.255.255.250) RP: 192.168.8.8 RPF nbr: 0.0.0.0 RPF idx: None Upstream State: JOINED 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local 0 . . . . . . Joined 0 . j . . . . . . Asserted 0 . . . . . . FCR:
	After source address filtering is enabled, run **show ip pim sparse-mode mroute** to display the multicast entry, and check whether the (S, G) entry and (S, G, RPT) entry exist.
	FS#show ip pim sparse-mode mroute IP Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 0 (S,G,rpt) Entries: 0 FCR Entries: 0 REG Entries: 0 (*, 239.255.255.250) RP: 192.168.8.8 RPF nbr: 0.0.0.0 RPF idx: None Upstream State: JOINED

```
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  j  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
FCR:
```

Limiting the Speed of the Source DR for Sending Register Packets

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Check the number of PIM-SM packets sent by Router B. ● Check the number of PIM-SM packets sent by Router B in 1s. ● Configure the speed of Router B for sending register packets. ● Check the number of PIM-SM packets sent by Router B in 1s.
	FS (config)#ip pim register-rate-limit 1
Verification	Display the number of PIM-SM packets sent by Router B before you configure the speed. The information should be displayed as follows:

```
FS#show ip pim sparse-mode track
PIM packet counters track
Elapsed time since counters cleared: 04d01h01m

                          received                    sent
Valid PIM packets:        18754                       29771
Hello:                    11149                       17842
Join-Prune:               0                           3234
Register:                 0                           3211
Register-Stop:            3192                        0
Assert:                   0                           0
BSM:                      0                           5484
C-RP-ADV:                 4413                        0
PIMDM-Graft:              0
PIMDM-Graft-Ack:          0
PIMDM-State-Refresh:      0
Unknown PIM Type:         0
```

```
Errors:
Malformed packets:                    0
Bad checksums:                        0
Send errors:                          0
Packets received with unknown PIM version: 0
FS#
```

Display the number of PIM-SM packets sent by Router B in 1s before the speed is configured. The information should be displayed as follows:

```
FS #show ip pim sparse-mode track
PIM packet counters track
Elapsed time since counters cleared: 04d01h04ms

                        received                sent
Valid PIM packets:      18765                   29789
Hello:                  11154                   17852
Join-Prune:             0                       3236
Register:               0                       3214
Register-Stop:          3195                    0
Assert:                 0                       0
BSM:                    0                       5487
C-RP-ADV:               4416                    0
PIMDM-Graft:            0
PIMDM-Graft-Ack:        0
PIMDM-State-Refresh:    0
Unknown PIM Type:       0

Errors:
Malformed packets:                    0
Bad checksums:                        0
Send errors:                          0
Packets received with unknown PIM version: 0
FS#
```

Display the number of PIM-SM packets sent by Router B after the speed is configured. The information should be

	displayed as follows:
	FS#show ip pim sparse-mode track PIM packet counters track Elapsed time since counters cleared: 04d01h06m received sent Valid PIM packets: 18777 29808 Hello: 11159 17862 Join-Prune: 0 3239 Register: 0 3215 Register-Stop: 3196 0 Assert: 0 0 BSM: 0 5489 C-RP-ADV: 4419 0 PIMDM-Graft: 0 PIMDM-Graft-Ack: 0 PIMDM-State-Refresh: 0 Unknown PIM Type: 0 Errors: Malformed packets: 0 Bad checksums: 0 Send errors: 0 Packets received with unknown PIM version: 0 FS#

Configuring the Checksum of the Register Packet Length

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Calculate the checksum of the entire register packet length. ● Run **show running-config** to check whether the preceding configurations take effect.
	FS(config)#ip pim register-checksum-wholepkt
Verification	Display the configurations on Router A, which should be as follows:
	FS#show running-config …

	! ip pim register-checksum-wholepkt ip pim rp-candidate Loopback 0 ! …

Enabling an RP to Decapsulate a Register Packet and Forward the Multicast Packet to Interfaces

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Enable Router A to forward a register packet. ● Run **show running-config** to check whether the preceding configurations take effect.
	FS(config)#ip pim register-decapsulate-forward
Verification	Display the configurations on Router A, which should be as follows:
	FS#show running-config … ! ! ip pim register-decapsulate-forward ip pim register-checksum-wholepkt ip pim rp-candidate Loopback 0 ! ! ! …

Configuring the Source IP Address of a Register Packet

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Configure the source address of Loop 2 as 192.168.2.2. ● Configure source address interface for the register packet of Router B as Loop 2. ● Run **show running-config** to check whether the preceding configurations take effect.
Verification	Display the configurations on Router B, which should be as follows:
	FS#show running-config ! ! ! ip pim register-source Loopback 1

	ip pim bsr-candidate Loopback 0 ! ! ! !

Configuring the Suppression Time of a Register Packet and the Inspection Time of a Null Register Packet

Configuration Steps	● Configure basic PIM-SM functions. (Omitted) ● Configure the suppression time of a register packet on Router B to 20s. ● Configure the inspection time of a null register packet on Router B to 2s. ● Run **show ip pim sparse-mode track** to display number of register packets.
	FS(config)#ip pim register-suppression 20 FS(config)#ip pim probe-interval 2
Verification	Display the number of register packets on Router B. The information should be displayed as follows:
	FS#show ip pim sparse-mode track PIM packet counters track Elapsed time since counters cleared: 04d23h15m received sent Valid PIM packets: 23788 43249 Hello: 13817 23178 Join-Prune: 0 4568 Register: 0 8684 Register-Stop: 4223 0 Assert: 0 0 BSM: 0 6819 C-RP-ADV: 5748 0 PIMDM-Graft: 0 PIMDM-Graft-Ack: 0 PIMDM-State-Refresh: 0 Unknown PIM Type: 0 Errors:

	Malformed packets: 0 Bad checksums: 0 Send errors: 0 Packets received with unknown PIM version: 0 FS# FS#
	In 18s, display the number of register packets on Router B. The information should be displayed as follows:
	FS#show ip pim sparse-mode track PIM packet counters track Elapsed time since counters cleared: 04d23h17m received sent Valid PIM packets: 23798 43263 Hello: 13820 23184 Join-Prune: 0 4569 Register: 0 8685 Register-Stop: 4224 0 Assert: 0 0 BSM: 0 6820 C-RP-ADV: 5749 0 PIMDM-Graft: 0 PIMDM-Graft-Ack: 0 PIMDM-State-Refresh: 0 Unknown PIM Type: 0 Errors: Malformed packets: 0 Bad checksums: 0 Send errors: 0 Packets received with unknown PIM version: 0 FS#

Configuring an RP to Receive Register Packets and the (S, G) Lifetime

Configuration	● Configure basic PIM-SM functions. (Omitted)

Steps	● Configure Router A to receive register packets and the (S, G) lifetime is 60s. ● Run **show ip pim sparse-mode mrout** to display number of register packets.
	FS(config)#ip pim rp-register-kat 60
Verification	After the lifetime is configured, check that the (S, G) lifetime on Router A does not exceed 60s.
	FS(config)#show ip pim sparse-mode mroute IP Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 0 (192.168.1.100, 233.3.3.3) RPF nbr: 192.168.36.90 RPF idx: VLAN 1 SPT bit: 0 Upstream State: NOT JOINED kat expires in 49 seconds 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local 0 . . . . . . Joined 0 . . . . . . Asserted 0 . . . . . . Outgoing 0 . . . . . . (192.168.1.100, 233.3.3.3, rpt) RP: 192.168.8.8 RPF nbr: 0.0.0.0 RPF idx: None

```
Upstream State: RPT NOT JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pruned
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

(*, 239.255.255.250)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0   .  j  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
FCR:

FS(config)#
FS(config)#show ip pi
```

Common Errors

- Basic PIM-SM functions are not configured or the configuration is not successful.
- The (S, G) of register packets is not configured on a C-RP or static RP, or the configuration is not successful.
- The ACL for limiting the (S, G) of register packets is not configured or the range of (S, G) in this ACL is not correctly configured.
- The range of (S, G) of register packets on each C-RP or static RP is not the same.

6.4.5 Configuring the Interval for Sending a Join/Prune Packet

Configuration Effect

- Change the interval for sending a Join/Prune packet to form an RPT or SPT.

Notes

- Basic PIM-SM functions must be configured.

Configuration Steps

- Configure the interval for sending a Join/Prune packet.

Verification

On Router B, configure the interval for sending a Join/Prune packet as 120s. Run **show ip pim sparse-mode mroute** to display the lifetime of the entry.

Related Commands

Configuring the Interval for Sending a Join/Prune Packet

Command	**ip pim jp-timer** *seconds*
Parameter Description	*Seconds*: Indicates the interval for sending a Join/Prune packet. The unit is second. The value ranges from 1 to 65,535 (60 by default).
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Interval for Sending a Join/Prune Packet

Configuration Steps	• Configure basic PIM-SM functions. (Omitted) • Configure the interval for sending a Join/Prune packet.
	FS(config)#ip pim jp-timer 120
Verification	Run **show ip pim sparse-mode mroute** to display the maximum timeout time of a Join/Prune packet.
	FS(config)#show ip pim sparse-mode mroute IP Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 1

```
(192.168.1.100, 233.3.3.3)
RPF nbr: 0.0.0.0
RPF idx: None
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 96 seconds
kat expires in 92 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  o  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
.

(192.168.1.100, 233.3.3.3, rpt)
RP: 192.168.8.8
RPF nbr: 192.168.36.89
RPF idx: GigabitEthernet 0/1
Upstream State: RPT NOT JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pruned
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
```

```
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .


(*, 239.255.255.250)
RP: 192.168.8.8
RPF nbr: 192.168.36.89
RPF idx: GigabitEthernet 0/1
Upstream State: JOINED
jt_timer expires in 119 seconds
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . i . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Joined
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Asserted
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
FCR:


VSU(config)#
```

Common Errors

- Basic PIM-SM functions are not configured or the configuration is not successful.

6.4.6 Configuring the Router of Last Hop to Switch from an RPT to SPT

Configuration Effect

- Switch from an RPT to SPT

Notes

- Basic PIM-SM functions must be configured.

Configuration Steps

- Configure the router of last hop to switch from an RPT to SPT.

Verification

Configure basic PIM-SM functions first. Configure the source DR to sends the data traffic (*, 233.3.3.3), and the receiving end to join group 233.3.3.3 forcibly to form an RPT. Configure the receiver DR to switch from the RPT to SPT forcibly. Run **show running-config** to display the result.

Related Commands

Enabling SPT switchover

Command	**ip pim spt-threshold** [**group-list** *access-list*]
Parameter Description	**group-list** *access-list*: Specifies the range of multicast group addresses allowed for SPT switchover using an ACL. access-list: The supported value ranges from 1 to 99 or 1,300 to 1,999. Naming an ACL is also supported.
Command Mode	Global configuration mode
Usage Guide	If **group-list** *access-list* is not specified, all groups are allowed to perform SPT switchover.

Configuration Example

Configuring the Router of Last Hop to Switch from an RPT to SPT

Configuration Steps	• Configure basic PIM-SM functions. (Omitted) • Configure the source DR to send the data traffic of group 233.3.3.3. • Configure the receiver DR to receive the data traffic of group 233.3.3.3. • Configure the receiver DR of last hop to switch from an RPT to SPT.
	FS(config)#ip pim spt-threshold
Verification	Run **show running-config** to display the configuration.
	! ! ip pim jp-timer 120 ip pim spt-threshold ip pim rp-candidate Loopback 0 ! ! !

6.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears multicast routing entries.	**clear ip mroute** { * | *group-address* [*source-address*] }
Clears the counters of multicast routes.	**clear ip mroute statistics** {* | *group-address* [*source-address*] }
Clears the information about dynamic RPs.	**clear ip pim sparse-mode bsr rp-set ***
Clears the counters of PIM-SM packets.	**clear ip pim sparse-mode track**

Displaying

Description	Command
Displays the details of BSR information.	**show ip pim sparse-mode bsr-router**
Displays the PIM-SM information of an interface.	**show ip pim sparse-mode interface** [*interface-type interface-number*] [**detail**]
Displays the local IGMP information about a PIM-SM interface.	**show ip pim sparse-mode local-members** [*interface-type interface-number*]
Displays the information about a PIM-SM multicast routing entry, and displays the RPF vector of a PIM-SM entry using **proxy**.	**show ip pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]]
Displays the information about PIM-SM neighbors.	**show ip pim sparse-mode neighbor** [**detail**]
Displays the information about the next hop of PIM-SM obtained from the NSM.	**show ip pim sparse-mode nexthop**
Displays the information about the RP corresponding the multicast group address *group-address*.	**show ip pim sparse-mode rp-hash** *group-address*
Displays the information about all the RPs and the groups they serve.	**show ip pim sparse-mode rp mapping**
Displays the number of PIM-SM packets sent and received since the statistic start time.	**show ip pim sparse-mode track**

7 Configuring PIM-SMv6

7.1 Overview

Protocol Independent Multicast (PIM) is a multicast routing protocol.

PIM does not rely on a specific unicast routing protocol. It uses the unicast routing table established by any unicast routing protocol to complete the reverse path forwarding (RPF) check and establish multicast routes. PIM does not need to transmit and receive multicast route updates. Therefore, the overhead of PIM is much lower than that of other multicast routing protocols.

PIM defines two modes: dense mode and sparse mode. Protocol Independent Multicast Spare Mode (PIM-SM) is applicable to various network environments.

PIM-SM running on IPv6 is called PIM-SMv6.

Protocols and Standards

- RFC4601: Protocol Independent Multicast -Sparse Mode (PIM-SM)
- RFC5059: Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)
- RFC3962: Protocol Independent Multicast - Dense Mode protocol
- RFC4607: Source-Specific Multicast for IP

7.2 Applications

Application	Description
ASM Implementation by Using PIM-SMv6	A receiver receives packets from any multicast source.
SSM Implementation by Using PIM-SMv6	A receiver selects a multicast source.
Application Example of an Embedded RP	An embedded RP address is configured within the IPv6 multicast group address.
PIM-SMv6 Application in a Hot Backup Environment	The multicast PIM-SMv6 protocol runs in a hot backup environment.

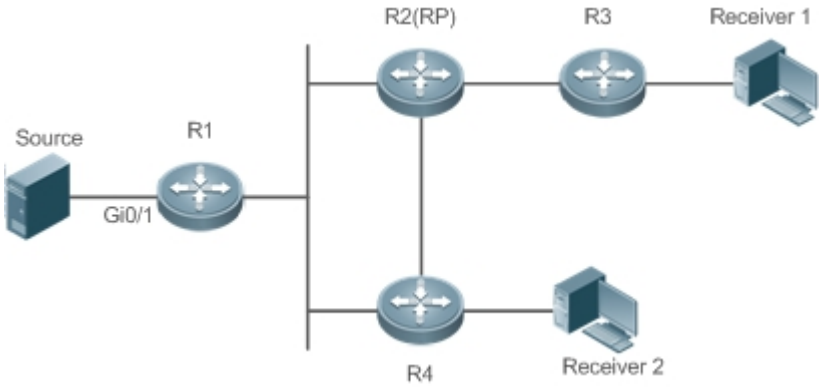
7.2.1ASM Implementation by Using PIM-SMv6

Scenario

The multicast service is provided only in one domain.

As shown in the following figure, receivers receive packets from any multicast source.

Figure 7- 1



Remarks	R1 is directly connected to the multicast source. R2 is configured as a rendezvous point (RP). R3 is directly connected to Receiver A. R4 is directly connected to Receiver B.

Deployment

- Run the Open Shortest Path First for IPv6 (OSPFv6) protocol in the domain to implement unicast routing.
- Run the PIM-SMv6 protocol in the domain to implement multicast routing.
- Run the Internet Group Management Protocol version 6 (IGMPv6) protocol in a user host network segment to implement group member management.

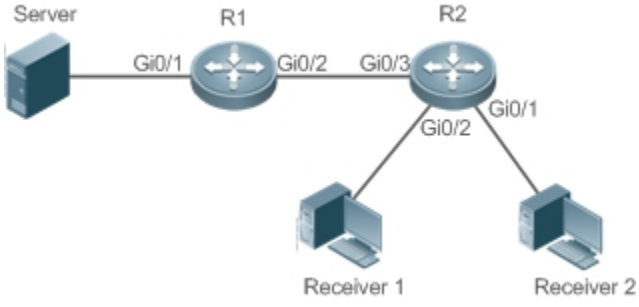
7.2.2 SSM Implementation by Using PIM-SMv6

Scenario

The multicast service is provided only in one domain.

As shown in the following figure, receivers receive packets from a specific multicast source.

Figure 7- 2



Remarks	R1 is directly connected to the multicast source. R2 is configured as an RP. R2 is directly connected to Receiver A. R2 is directly connected to Receiver B.

Deployment

- Run the OSPFv6 protocol in the domain to implement unicast routing.
- Run the PIM-SMv6 protocol in the domain to implement multicast routing.
- Enable the source-specific multicast (SSM) function of the PIM-SMv6 protocol to implement the SSM function.
- Run the Internet Group Management Protocol version 3 (IGMPv3) in a user host network segment to implement group member management.

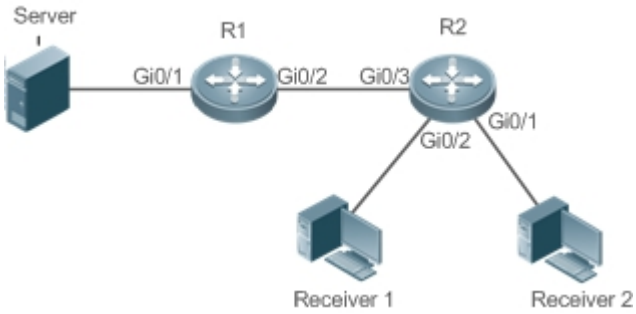
7.2.3 Application Example of an Embedded RP

Scenario

The multicast service is provided only in one domain.

As shown in the following figure, an RP address is configured for R2 to make the router become an embedded RP.

Figure 7- 3



Remarks	R1 is directly connected to the multicast source. R2 is configured as an RP. R2 is directly connected to Receiver A. R2 is directly connected to Receiver B. R2 is configured as an embedded RP.

Deployment

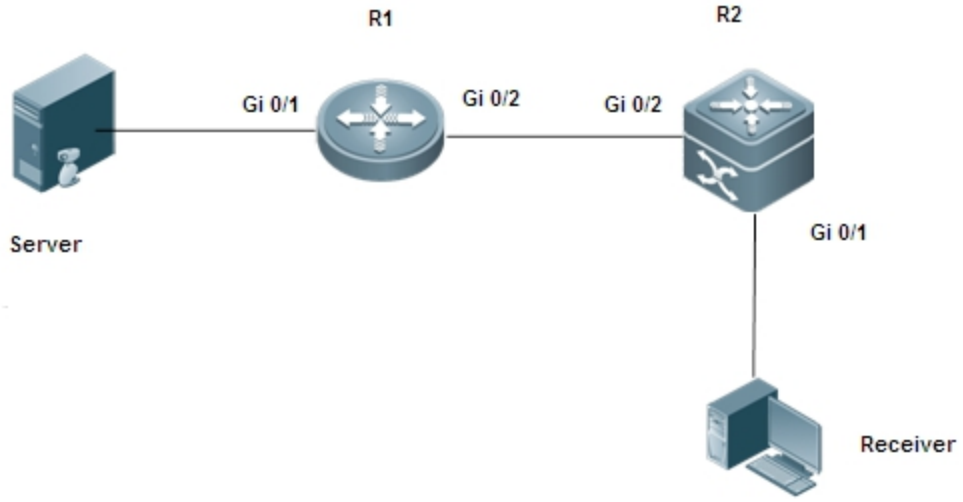
- Run the OSPFv6 protocol in the domain to implement unicast routing.
- Run the PIM-SMv6 protocol in the domain to implement multicast routing.
- Enable the SSM function of the PIM-SMv6 protocol to implement the SSM function.
- Run the IGMPv3 protocol in a user host network segment to implement group member management.
- Configure the RP address and embedded RP on R2.

7.2.4 PIM-SMv6 Application in a Hot Backup Environment

Scenario

In a hot backup environment, run PIM-SMv6. A device performs hot backup switching to ensure that traffic is not interrupted.

Figure 7- 4



Remarks	R1 is connected to the video server, R2 is directly connected to the receiver, and R2 runs in hot backup mode. A Layer-3 multicast protocol runs on R1 and R2.

Deployment

- Run OSPF on R1 and R2 to implement unicast routing.
- Run PIM-SMv6 on R1 and R2 to implement multicast routing.
- Make R2 run in two-node cluster hot backup mode.

Remarks	R2 may perform hot backup switching in the hot backup environment. In this case, the query interval of PIM Hello packets (the default value is 30 seconds) needs to be adjusted on R2 because the keepalive timer of the neighbor in PIM Hello packets of R1 may have expired (the default value is 3.5 times the query interval, that is, 105 seconds). The multicast function relies on the unicast function currently, and the multicast function starts convergence after the unicast function convergence is complete. For example, the default graceful restart (GR) convergence time of the unicast function is 120 seconds. It is recommended that the query interval of PIM Hello packets be set to 60 seconds. The keepalive time of the neighbor in PIM Hello packets is 210 seconds. In this scenario, the query interval of PIM Hello packets need to be set with a reference to the GR convergence time of the unicast function and the value of 3.5 times the query interval of PIM Hello packets must be larger than the GR convergence time of the unicast function. In addition, if the convergence time of the unicast function is long, the transmission interval of PIM Join/Prune packets also need to be adjusted, because the keepalive time of PIM Join/Prune packets is 3.5 times the transmission interval of PIM Join/Prune packets. The default keepalive time of PIM Join/Prune packets is 210 seconds. If R2 is configured as a dynamic RP, the interval for a candidate RP (C-RP) to transmit C-RP notifications also needs to be adjusted. The default transmission interval is 60 seconds and the keepalive time is 2.5 times the transmission interval of C-RP notifications. For example, if the convergence time of the unicast function is longer than 150 seconds, the transmission interval of C-RP notifications needs to be adjusted. In a hot backup environment, it is recommended that the query interval of PIM Hello packets be larger than the default value (30 seconds). Otherwise, the keepalive timer of the neighbor in PIM Hello packets of the peer end times out during hot backup switching.

7.3 Features

Basic Concepts

PIM Router and PIM Interface

Routers where the PIM protocol is enabled are called PIM routers. Interfaces where the PIM protocol is enabled are called PIM interfaces.

Multicast packets are forwarded by PIM routers. The PIM interfaces for receiving multicast packets are called upstream interfaces, and the PIM interfaces for transmitting multicast packets are called downstream interfaces.

Network segments where upstream interfaces are located are called upstream network segments. Network segments where downstream interfaces are located are called downstream network segments.

PIM Network and PIM Domain

PIM routers are connected through PIM interfaces and form a PIM network.

On some PIM interfaces, borders are set to divide a large PIM network into multiple PIM domains. The borders may reject specific multicast packets or limit transmission of PIM messages.

Multicast Distribution Tree, DR, RP

Multicast packets are transmitted from one point to multiple points. The forwarding path presents a tree structure. This forwarding path is called a multicast distribution tree (MDT). MDTs are classified into two types:

- Rendezvous point tree (RPT): Uses the rendezvous point (RP) as the root and designated routers (DRs) connected to group members as leaves.
- Shortest path tree (SPT): Use the DR connected to a multicast source as the root and the RPs or DRs connected to group members as leaves.

DRs and RPs are function roles of PIM routers.

- RPs collect information about multicast sources and group members in the network.
- The DR connected to a multicast source reports multicast source information to the RP and the DRs connected to group members report the group member information to the RP.

(*,G), (S,G)

- (*,G): Indicates the packets transmitted from any source to Group G, routing entries corresponding to the packets, and forwarding path (RPT) corresponding to the packets.
- (S,G): Indicates the packets transmitted from Source S to Group G, routing entries corresponding to the packets, and forwarding path (SPT) corresponding to the packets.

ASM, SSM

PIM-SM supports two multicast service models: any-source multicast (ASM) and source-specific multicast (SSM), which are applicable to different multicast address segments.

- ASM: In the ASM model, a user host cannot select a multicast source. The user host joins a multicast group and receives all packets sent from all sources to the multicast group.
- SSM: In the SSM model, a user host can select a multicast source. The user host specifies the source address when joining a multicast group, and then receives packets only from the specified source to the multicast group.

SSM model requirement: Other network services must be used to enable a user host to know the position of a multicast source in advance so that the user host selects the multicast source.

Overview

Feature	Description
Establishment of PIM Neighbor Relationships	Neighbor relationships are established between PIM routers to form a PIM network.
DR Election	In the shared network segment connected to group members, DR election is conducted among PIM neighbors to elect the DR connected to group members. In the shared network segment connected to a multicast source, DR election is conducted among PIM neighbors to elect the DR connected to the multicast source.
RP Mechanism	In a PIM network, the RP is statically configured or dynamically elected so that each PIM router knows the position of the RP.
Registration Information About a Multicast Source	When a multicast source arises in a network, the DR connected to the multicast source transmits the Register packet to the RP so that the RP obtains information about the multicast source and multicast packets.
RPT Establishment	When a group member arises in a network, the DR connected to the group member transmits the Join packet in the RP direction to establish an RPT. If there is a multicast source in the network, the multicast packet transmitted to the RP can reach the group member along the RPT.
SPT Establishment	When a data packet reaches the DR connected to a group member, the DR connected to the group member transmits the Join packet in the multicast source direction to establish an SPT. Then, multicast packets are forwarded along the SPT.
ASM and SSM Models	PIM routers provide multicast services of the ASM model and SSM model. The SSM model is used for groups within the SSM address range, and the ASM model is used for other groups.

7.3.1 Establishment of PIM Neighbor Relationships

Neighbor relationships are established between PIM routers to form a PIM network. Neighbor relationships must be established between PIM routers before other PIM control messages are exchanged or multicast packets are forwarded.

Working Principle

A Hello message is sent by a PIM interface. For the multicast packet for encapsulating the Hello message, the destination address is ff02::d (indicating all PIM routers in the same network segment), the source address is the IP address of the PIM interface, and the time to live (TTL) value is 1.

Hello messages are used to discover neighbors, negotiate about protocol parameters, and maintain neighbor relationships.

➘ Discovering PIM Neighbors

PIM routers in the same network segment receive multicast packets with the destination address of ff02::d. In this way, the PIM routers obtain neighbor information and establish neighbor relationships.

When a PIM interface is enabled or detects a new neighbor, the Triggered-Hello-Delay message is used to generate a random time period. Within the time period, the interface sends Hello packets.

➘ Negotiating About Protocol Parameters

A Hello message contains multiple protocol parameters, which are described as follows:

– DR_Priority: Indicates the priority of each router interface for DR election. A higher priority means a higher possibility of being elected as the DR.

– Holdtime: Indicates the timeout time in which a neighbor is held in the reachable state.

– LAN_Delay: Indicates the delay for transmitting a Prune message in a shared network segment.

– Override-Interval: Indicates the prune override time carried in a Hello message.

When a PIM router receives a Prune message from an upstream interface, it indicates that other downstream interfaces exist in the shared network segment. If the PIM router still needs to receive multicast data, it must send a Prune Override message to the upstream interface within the time of **Override-Interval**.

LAN_Delay + Override-Interval = PPT (Prune-Pending Timer). After a PIM router receives a Prune message from a downstream interface, it does not immediately perform pruning but waits for PPT timeout. After the PPT times out, the PIM router performs pruning. Within the time of PPT, if the PIM router receives a Prune Override message from the downstream interface, it cancels pruning.

Maintaining Neighbor Relationships

A Hello message is sent periodically between PIM routers. If a Hello packet is not received from a PIM neighbor within Holdtime, the neighbor is considered unreachable and is deleted from the neighbor list. Any changes in PIM neighbors will cause multicast topology changes in the network. If an upstream neighbor or a downstream neighbor in an MDT is unreachable, multicast routing re-convergence is performed again and the MDT is migrated.

Related Configuration

Enabling the PIM-SMv6 Function on an Interface

By default, the PIM-SMv6 function is disabled on an interface.

Run the **ipv6 pim sparse-mode** command to enable or disable the PIM-SMv6 function on an interface.

The PIM-SMv6 function must be enabled on an interface so that the interface participates in the operation of PIM protocols. If the PIM-SMv6 function is disabled on an interface that functions as a DR, static RP, candidate - rendezvous point (C-RP), or candidate – bootstrap router (C-BSR), the corresponding protocol role does not take effect.

Adjusting the Transmission Interval of Hello Messages on an Interface

By default, Hello messages are transmitted at an interval of 30 seconds.

Run the **ipv6 pim query-interval** *seconds* command to adjust the transmission interval of Hello messages on an interface. The value ranges from 1 to 65,535.

A larger value of *interval-seconds* means a larger transmission interval of Hello messages and a smaller value of *interval-seconds* means a smaller transmission interval of Hello messages.

7.3.2 DR Election

In the shared network segment connected to group members, DR election is conducted among PIM neighbors to elect the DR connected to the group members.

In the shared network segment connected to a multicast source, DR election is conducted among PIM neighbors to elect the DR connected to the multicast source.

The DR transmits the Join/Prune message in the MDT root node direction for the directly connected group members, or transmits data of the directly connected multicast source to the MDT.

Working Principle

The neighbor IP address and DR priority are obtained from Hello packets of neighbors during establishment of PIM neighbor relationships, so as to elect the DR.

The key of DR election is the DR priorities and IP addresses of interfaces.

Interface DR Priority

A higher interface DR priority means a higher probability that a PIM router is successfully elected as the DR during the DR election.

Interface IP Address

If interfaces of PIM routers share the same DR priority during DR election, IP addresses of neighbors are compared. A larger IP address means a higher probability that a PIM router is successfully elected as the DR.

Related Configuration

Setting IP Addresses of Interfaces

By default, no IP addresses are configured for interfaces.

Run the **ipv6 address** command to set an IP address for an interface.

When PIM routers share the same DR priority, the PIM router with a larger IP address is elected as the DR.

Enabling the PIM-SMv6 Function on an Interface

By default, the PIM-SMv6 function is disabled on an interface.

Run the **ipv6 pim sparse-mode** command to enable or disable the PIM-SMv6 function on an interface.

The PIM-SMv6 function must be enabled on an interface so that the interface participates in the operation of PIM protocols. If the PIM-SMv6 function is disabled on an interface that functions as a DR, static RP, C-RP, or C-BSR, the corresponding protocol role does not take effect.

Adjusting the DR Priority of an Interface

By default, the DR priority is 1.

Run the **ipv6 pim dr-priority** *priority-value* command to adjust the DR priority of an interface. The priority value ranges from 0 to 4,294,967,294.

The DR priority of an interface is used to elect the DR in the directly connected network segment of the interface. A larger priority value means a higher probability that a PIM router is elected as the DR.

7.3.3 BSR Mechanism

In a PIM network, the bootstrap router (BSR) periodically generates bootstrap messages (BSMs) including information about a series of C-RPs and relevant group addresses. BSMs are transmitted hop by hop in the entire domain. PIM routers throughout the network receive BSMs and record information about C-RPs and the relevant group addresses.

Working Principle

One or more C-BSRs are configured in the PIM-SMv6 domain and the BSR is elected from the candidate BSRs according to certain rules.

Related Configuration

Configuring a C-BSR

By default, no C-BSR is configured.

Run the **ipv6 pim bsr-candidate** *interface-type interface-number* [*hash-mask-length* [*priority-value*]] command to configure or cancel a C-BSR.

C-BSRs elect the globally unique BSR in the PIM-SM domain by means of BSM learning and election. The BSR transmits BSMs.

Configuring the BSR Border

By default, no BSR border is configured.

Run the **ipv6 pim bsr-border** command to configure or cancel the BSR border.

After this command is configured for an interface, the interface immediately discards the received BSMs and does not forward BSMs, thereby preventing BSM flooding. No BSR border is configured if this command is not configured.

Defining the Valid BSR Range

By default, the BSMs of BSRs are not filtered.

Run the **ipv6 pim accept-bsr list** *ipv6_access-list* command to define or cancel the BSR range.

After this command is configured, the valid BSR range is defined. If this command is not configured, the device with the PIM-SMv6 function enabled will receive all BSMs.

Configuring a C-BSR to Restrict the Address Range of Valid C-RPs and the Range of Multicast Groups Served by the C-RPs

A C-BSR receives notifications from all C-RPs.

Run the **ipv6 pim accept-crp list** *ipv6_access-list* command to configure whether to filter notifications from C-RPs.

After this command is configured, the C-BSR restricts the address range of valid C-RPs and the range of multicast groups served by the C-RPs. If this command is not configured, the C-BSR receives notifications from all C-RPs.

Configuring a C-BSR to Receive C-RP-ADV Packets with prefix-count of 0

By default, a C-BSR does not receive C-RP-ADV packets with prefix-count of 0.

Run the **ipv6 pim accept-crp-with-null-group** command to configure whether to receive C-RP-ADV packets with prefix-count of 0.

After this command is configured, the C-BSR can receive C-RP-ADV packets with prefix-count of 0. If this command is not configured, the C-BSR does not process C-RP-ADV packets with prefix-count of 0.

7.3.4 RP Mechanism

In a PIM network, the RP is statically configured or dynamically elected so that each PIM router knows the position of the RP. The RP serves as the root of the RPT. The RPT establishment and the forwarding of RPT data streams must use the RP as the forwarding point.

Working Principle

All PIM routers in a PIM domain must be able to be mapped to the same RP through a specific multicast group address. RPs are classified into static RPs and dynamic RPs in a PIM network.

Static RP

In static RP configuration, the RP address is directly configured on each PIM router so that all PIM routers in the PIM network know the RP address.

Dynamic RP

C-RPs are also configured in the PIM-SMv6 domain. These C-RPs transmit data packets that contain their addresses and information about multicast groups served by them to the BSR in unicast mode. The BSR periodically generates BSMs that contain information about a series of C-RPs and their group addresses. BSMs are transmitted hop by hop in the entire domain. Devices receive and store these BSMs. The DR at the receive end uses a hash algorithm to map a group address to the C-RP that can serve the group. Then, the RP corresponding to the group address can be determined.

Related Configuration

Setting a Static RP Address

By default, no RP address is configured.

Run the **ipv6 pim rp-address** *ipv6_rp-address* [*ipv6_access-list*] command to configure or cancel a static RP address for a PIM router.

An RP must be configured so as to implement ASM in a PIM-SMv6 network. You can configure a static RP or dynamic RP.

If a static RP is configured in a PIM-SMv6 network, the static RP configuration on all devices in the PIM-SMv6 domain must be consistent to prevent multicast route ambiguity in the PIM-SMv6 domain.

Configuring a C-RP Address

By default, no C-RP address is configured.

Run the **ipv6 pim rp-candidate** *interface-type interface-number* [**priority** *priority-value*] [**interval** *interval-seconds*] [**group-list** *ipv6_access-list*] command to configure or cancel a PIM router as a C-RP.

C-RPs periodically transmit C-RP notifications to the BSR. Information contained in these C-RP notifications is dispersed to all PIM-SMv6 devices in the domain, thereby ensuring the uniqueness of RP mapping.

Ignoring the RP Priority in RP Setting

By default, a C-RP with a higher priority is selected preferentially.

Run the **ipv6 pim ignore-rp-set-priority** command to specify or ignore the RP priority when selecting the RP for a group.

When one RP needs to be selected for a multicast address and multiple RPs can serve this multicast address, use this command if the RP priority needs to be ignored during the RP comparison. If this command is not configured, the RP priority will be considered during the RP comparison.

Configuring the Static RP First

By default, a dynamic C-RP is adopted preferentially.

Run the **ipv6 pim static-rp-preferred** command to select the static RP first during RP selection.

After this command is configured, the static RP is adopted first. If this command is not configured, a C-RP is adopted first.

Configuring the Embedded RP Function

By default, the embedded RP function is enabled for all IPv6 multicast group addresses where the RP address is embedded.

Run the **ipv6 pim rp embedded** [**group-list** *ipv6_acl_name*] command to enable the embedded RP function.

The embedded RP function is the peculiar RP discovery mechanism of IPv6 PIM. This mechanism uses the IPv6 multicast address where the RP address is embedded, to enable a multicast device to directly extract the RP address from the multicast address. By default, the embedded RP function is enabled for all IPv6 multicast group addresses where the RP address is embedded.

7.3.5 Registration Information About a Multicast Source

When a multicast source arises in a network, the DR connected to the multicast source transmits the Register packet to the RP so that the RP obtains information about the multicast source and multicast packets.

Working Principle

The DR at the data source end receives a multicast data packet from the directly connected host, and encapsulates the multicast data into a Register message. Then, it transmits the Register message to the RP in unicast mode. The RP generates the (S,G) entry.

If the forwarding entry contains an outgoing interface on the RP, the RP forwards the encapsulated data packet to the outgoing interface.

If the RP does not have the forwarding entry of the current group, it starts the (S,G) entry start timer. After the timer expires, the RP transmits the Register-Stop message to the DR and deletes the entry. After the DR at the data source end receives the Register-Stop message, the DR transmits the probing packet before the Register-Stop message timer expires.

If the DR does not receive the Register-Stop message, after the timer expires, the DR at the data source end encapsulates the multicast data into the Register message and transmits it to the RP in unicast mode.

If the DR receives the Register-Stop message, it re-starts the delay and re-transmits the probing packet before the delay expires.

Related Configuration

Configuring Reachability Detection of RP Register Packets

By default, the RP reachability is not detected.

Run the **ipv6 pim register-rp-reachability** command to set or cancel the RP reachability detection.

If the RP reachability needs to be detected for the Register packet transmitted from the DR to the RP, you can configure this command. After this command is configured, the RP reachability is detected before the DR transmits the Register packet to the RP. That is, the DR queries the unicast routing table and static multicast routing table to check whether a route reachable to the RP exists. If no, the DR does not transmit the Register packet.

Configuring the RP to Filter Register Packets

By default, the RP allows every received Register packet.

Run the **ipv6 pim accept-register** { **list** *ipv6_access-list* [**route-map** *map-name*] | **route-map** *map-name* [**list** *ipv6_access-list*] } command to enable or disable the RP to filter received Register packets.

To filter received Register packets on the RP, configure this command. If this command is not configured, the RP allows every received Register packet. If this command is configured, only Register packets whose source addresses and group addresses are allowed by the ACL are processed. Otherwise, the Register packets are filtered out.

Configuring the Transmission Rate Limit for Register Packets

By default, the transmission rate of Register packets is not limited.

Run the **ipv6 pim register-rate-limit** *rate* command to configure whether to limit the transmission rate of Register packets.

If **no** is set in this command, the transmission rate is not limited. This command is used to configure the transmission rate of Register packets from the (S,G) multicast group address rather than the transmission rate of Register packets of the entire system.

Configuring the Checksum Calculation of a Register Packet Based on the Entire Register Packet

By default, the checksum in a Register packet is calculated in default mode specified in the protocol.

Run the **ipv6 pim register-checksum-wholepkt [group-list** *ipv6_access-list* **]** command to set the packet length for checksum calculation.

If the entire PIM protocol packet including the encapsulated multicast data packet is used for checksum calculation of a Register packet, use this command. If this command is not configured, the checksum in a Register packet is calculated in default mode specified in the protocol.

Configuring the Source Address of Register Packets

By default, the source address of Register packets uses the address of the DR interface connected to a multicast source.

Run the **ipv6 pim register-source** { *ipv6_local_address* | *interface-type interface-number* } command to configure the source address of Register packets.

To configure the source address of Register packets transmitted from the DR, use this command. If this command is not configured or **no** is set in this command, the source address of Register packets uses the address of the DR interface connected to a multicast source. If the address parameter of this command is used, the configured address must be a reachable unicast route. If the interface parameter of this command is used, this interface may be a loopback interface or an interface of other types and the interface address must be an advertised unicast route.

Configuring the Suppression Time of Register Packets

The default suppression time of Register packets is 60 seconds.

Run the **ipv6 pim register-suppression** *seconds* command to configure the suppression time.

If this command is used to configure the suppression time of Register packets, configuring the value on the DR will change the suppression time of Register packets on the DR. If the **ipv6 pim rp-register-kat** *seconds* command is not configured, defining the value on the RP will change the keepalive time on the RP.

Configuring the Probing Time of NULL Register Packets

The default probing time is 5 seconds.

Run the**ipv6 pim probe-interval** *interval-seconds* command to set the probing time.

The source DR transmits the NULL-Register packet to the RP within a certain interval prior to the timeout of the suppression time of the Register packet. This interval is the probing time. The default probing time is 5 seconds.

Configuring the Time Value of the RP KAT Timer

By default, the KAT default value is used. KAT default value = Registration suppression time x 3 + Registration detection time.

Run the **ipv6 pim rp-register-kat** *seconds* command to set time of the KAT timer.

To configure the keepalive time of Register packets from the (S,G) multicast group address on the RP, use this command.

7.3.6 RPT Establishment

When a group member arises in a network, the DR connected to the group member transmits the Join packet in the RP direction to establish an RPT. If there is a multicast source in the network, the multicast packet transmitted to the RP can reach the group member along the RPT.

Working Principle

The RPT establishment process is as follows:

1. The DR at the receive end receives an MLD (*,G)Include report packet from a receiver.

2. If the DR at the receive end is not the RP of Group G, the DR at the receive end transmits one (*.G)join packet in the RP direction. The upstream device that receives the (*.G)join packet transmits the (*.G)join packet in the RP direction. The (*.G)join packet is transmitted hop by hop till the RP of Group G receives the (*.G)join packet, indicating that the DR at the receive end joins the RPT.

3. When the data source host transmits multicast data to a group, the source data is encapsulated into the Register message and is transmitted to the RP in unicast mode by the DR at the data source end. The RP decapsulates the Register message, retrieves the data packet, and then forwards it to each group member along the RPT.

4. The RP transmits the (S,G)join packet to the DR at the data source end to join the SPT of this data source.

5. After the SPT from the RP to the DR at the data source end is established, data packets from the data source are transmitted to the RP along the SPT without encapsulation.

6. When the first multicast data packet reaches the RP along the SPT, the RP transmits the Register-Stop message to the DR at the data source end to enable the DR to stop the encapsulation of Register packets. After the DR at the data source end receives the Register-Stop message, it does not encapsulate the Register packets but transmits the Register packets to the RP along the SPT of the data source. The RP forwards the Register packets to each group member along the RPT.

Related Configuration

Configuring the Transmission Interval of Join/Prune Packets

The default transmission interval of Join/Prune packets is 60 seconds.

Run the **ipv6 pim jp-timer** *seconds* command to set the transmission interval of Join/Prune packets.

To change the default transmission interval of Join/Prune packets, configure this command. If this command is not configured, the default transmission interval of Join/Prune packets is 60 seconds.

7.3.7 SPT Establishment

When a data packet reaches the DR connected to a group member, the DR connected to the group member transmits the Join packet in the multicast source direction to establish an SPT. Then, multicast packets are forwarded along the SPT, thereby relieving the load of the RP in the RPT and reducing the number of hops from the DR at the data source end to the receive end.

Working Principle

The SPT establishment process is as follows:

The DR at the receive end transmits the (*,G)join packet to the DR at the source end along the SPT. The (*,G)join packet is transmitted hop by hop till the DR at the source end receives the (*,G)join packet, forming an SPT.

Related Configuration

By default, SPT switching is disabled.

Run the **ipv6 pim spt-threshold** [**group-list** *ipv6_access-list*] command to configure whether to start SPT switching.

After this command is configured, when the DR receives the (S,G) packet from the first group member, one PIM Join message is generated and forwarded to the RP to establish a SPT tree. If **group-list** is defined, the defined group is switched from the RPT to the SPT. If **no** is set in this command and **group-list** is not defined, the switching from the RPT to the SPT is disabled and the device redirects to the RPT and transmits one Prune packet to the source. If **no** is set in this command, **group-list** is defined, and the defined ACL is a configured ACL, the ACL associated with **group-list** is cancelled and all groups are allowed to switch from the RPT to the SPT.

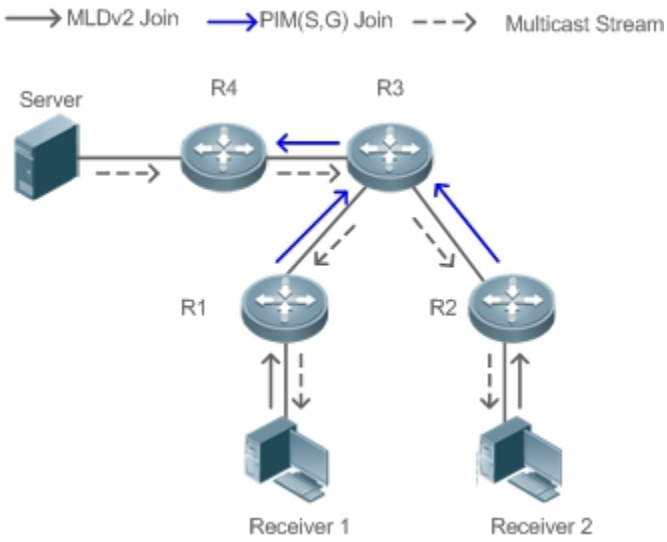
7.3.8 ASM and SSM Models

PIM-SM supports two multicast models: ASM and SSM. In the ASM model, a multicast data receiver specifies only to join a multicast group G but does not specify the multicast source S. In the SSM model, a multicast data receiver can specify both the multicast source S and multicast group G.

When the SSM model is implemented over IPv6, MLDv2 needs to be used to manage the member relationship between hosts and devices and PIM-SMv6 needs to be used to connect devices.

In the SSM model, a multicast receiver learns about the multicast source (S,G) information by means of some channels (such as accessing the server or receiving advertisements) in advance. When the multicast receiver needs to order a multicast service, it directly transmits the MLD(S,G) Join packet to the last-hop device, for example, as shown in the following figure, Multicast Receiver 1 transmits the MLD(S,G) Join packet to order the multicast service (S,G). After receiving the MLD(S,G) Join packet from the multicast receiver, the last-hop device transmits the PIM(S,G) Join packet to the multicast source hop by hop, for example, as shown in the following figure, after receiving the MLD(S,G) Join packet from Multicast Receiver 1, R1 transmits the PIM(S,G) Join packet to R3, which transmits the PIM (S,G) Join packet to R4. As a result, the SPT from the multicast receiver to the multicast source is established.

Figure 7- 5



The following conditions need to be met for the implementation of the SSM model:

- A multicast receiver learns about the multicast source (S,G) information beforehand by means of some channels. The multicast receiver initiates the MLD(S,G) Join packet to the desired multicast service.
- MLDv2 must be enabled on the interface of the last-hop device connected to the multicast receiver. MLDv1 does not support SSM.
- PIM-SM and SSM must be enabled on the intermediate devices between the multicast receiver and the multicast source.

After the SSM function is enabled, the default group range of SSM is FF3x::/32. You can run a command to change the group range of SSM.

The SSM model has the following features:

- In the SSM model, a multicast receiver can learn about the multicast source information in advance by means of some channels (for example, receiving advertisements or accessing a specified server).
- The SSM model is a specific subset of PIM-SM and processes only PIM(S,G) Join and PIM(S,G) Prune messages. It discards RPT-relevant messages within the SSM range, for example, PIM(*,G) Join/Prune messages. For Register packets within the SSM range, it immediately responds with the Register-Stop packet.
- In the SSM model, no RP is required and the election and distribution of RP messages are not required. The established MDT is the SPT in SSM.

7.4 Configuration

7.4.1 Configuring Basic Functions of PIM-SMv6

Configuration Effect

- Create a PIM network to provide the IPv6 multicast service for data sources and user terminals in the network.
- Both or either of the two multicast service models (ASM and SSM) can be supported.

Notes

- PIM-SMv6 needs to use the IPv6 unicast routing function.
- If the PIM network needs to support the multicast service of the SSM model, MLDv3 or SSM Mapping needs to be configured.

Configuration Steps

Enabling the IPv6 Multicast Routing Function

- Mandatory.
- The IPv6 multicast routing function should be enabled on each router unless otherwise specified.

Enabling the PIM-SMv6 Function

- Mandatory.
- The PIM-SMv6 function should be enabled on the following interfaces unless otherwise specified: router interconnection interfaces, interface that function as a static RP, C-RP, or C-BSR, interface for connecting to a multicast source, and interface for connecting to a user host.

Enabling the PIM-SMv6 PASSIVE Function

- In a PIM network, if an interface needs to receive multicast packets without participating in the PIM network topology construction, the PIM-SMv6 PASSIVE mode can be configured.
- If no special requirements are raised, enable the PIM-SMv6 PASSIVE function on the following interfaces: interfaces of the stub network device in the multicast network for connecting to STAs. After the PIM-SMv6 PASSIVE function is configured on an interface, the interface neither sends nor receives PIM packets.

Configuring the PIM-SMv6 Sub VLAN Function

- In general, a super VLAN includes many sub VLANs. If the PIM-SMv6 protocol is enabled on the interfaces of the super VLAN, multicast packets will be replicated and sent to all sub VLANs. As a result, the traffic may exceed the device capability, causing protocol flapping. The Super VLAN interface is disabled with PIM-SMv6 generally. Use this command to enable PIM-SMv6 on the Super VLAN interface to send PIM packets to all sub VLANs or the specified sub VLAN.
- This function is available only on the interfaces of the super VLAN.

Configuring an RP

- If a PIM network needs to support the multicast service of the ASM model, an RP must be configured.
- There are three methods of configuring an RP: configuring only a static RP, configuring only a dynamic RP, and configuring both a static RP and a dynamic RP. If both a static RP and a dynamic RP are configured, the dynamic RP is preferred.
- Configuring a static RP: A static RP should be configured on each router unless otherwise specified.
- Configuring a dynamic RP: A C-RP or C-BSR should be configured on one or more routers unless otherwise specified.

Enabling the SSM

- If a PIM network needs to support the multicast service of the SSM model, the SSM must be enabled.
- The SSM should be enabled on each router unless otherwise specified.

Verification

Make a multicast source in the network send packets to groups within the range of ASM and SSM and make a user host join the groups.

- Check whether the user host can successfully receive packets from each group.
- Check whether correct PIM-SMv6 routing entries are created on routers.

Related Commands

Enabling the IPv6 Multicast Routing Function

Command	**ipv6 multicast-routing**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling the PIM-SMv6 Function

Command	**ipv6 pim sparse-mode**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Before enabling the PIM-SMv6 function, enable the multicast routing forwarding function in global configuration mode. Otherwise, multicast data packets cannot be forwarded even if the PIM-SMv6 function is enabled. When the PIM-SMv6 function is enabled, MLD is automatically enabled on each interface without manual configuration.

	If the message "Failed to enable PIM-SMv6 on <*interface name*>, resource temporarily unavailable, please try again" is displayed during the configuration of this command, try to configure this command again. If the message "PIM-SMv6 Configure failed! VIF limit exceeded in NSM!!!" is displayed during the configuration of this command, the configured number of multicast interfaces reaches the upper limit of multicast interfaces that can be configured on the device. If the PIM-SMv6 function still needs to be enabled on an interface, delete some unnecessary PIM-SMv6 or PIM-DMv6 interfaces. If an interface is of the tunnel type, only the 6Over4 configuration tunnel, 6Over4 GRE tunnel, 6Over6 configuration tunnel, and 6Over6 GRE tunnel support the IPv6 multicast function. The multicast function can also be enabled on tunnel interfaces that do not support the multicast function but no prompts are displayed and multicast packets are neither received nor transmitted. Multicast tunnels can be established only on Ethernet ports. Embedded tunnels and QoS/ACL of multicast data are not supported.

Enabling PIM-SMv6 PASSIVE Mode

Command	**ipv6 pim sparse-mode passive**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Before enabling the PIM-SMv6 function, enable the multicast routing forwarding function in global configuration mode. Otherwise, multicast data packets cannot be forwarded even if the PIM-SMv6 PASSIVE function is enabled. When the PIM-SMv6 function is enabled, MLD is automatically enabled on each interface without manual configuration. Interfaces with the PIM-SMv6 PASSIVE function enabled neither receive nor transmit PIM packets but can forward multicast packets. Therefore, the PIM-SMv6 PASSIVE mode is generally configured on the interface of the stub network device connected to a user host, so as to prevent Layer-2 flooding of PIM Hello packets

Enabling the PIM-SMv6 Sub VLAN Function

Command	**ipv6 pim sparse-mode subvlan** [**all** | *vid*]
Parameter Description	**all**: sends packets to all sub VLANs. *vid*: sends packets to a specified sub VLAN.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring a Static RP

Command	**ipv6 pim rp-address** *ipv6_rp-address* [*ipv6_access-list*]

Parameter Description	*ipv6_rp-address*: Indicates the IPv6 address of an RP. *ipv6_access_list*: References an IPv6 ACL to restrict the group address range served by the static RP. A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	Multicast static RPs can be configured. A static RP and a C-RP can coexist. Notes: 1. If both the BSR mechanism and RP static configuration are effective, the dynamic configuration is preferred. 2. A control list can be used to statically configure the address of an RP for multiple multicast groups (using the ACL) or all multicast groups (without using the ACL), but one static RP address cannot be used multiple times. 3. If multiple static RPs serve the same group, the static RP with a larger IPv6 address is used preferentially. 4. Only multicast groups with the addresses allowed by the ACL are effective. By default, all multicast groups are allowed. 5. After the configuration is complete, the static RP source address will be inserted into the group range-based static RP group tree structure. The multicast static group in each group range maintains the linked list structure of one static RP group. This linked list is arranged in descending order by IPv6 address. When an RP is selected for a group range, the RP with the largest IPv6 address will be selected. 6. When a static RP address is deleted, this address is deleted from all existing groups and an address is selected from the existing static RP tree structure as the RP address.

Configuring a C-RP

Command	**ipv6 pim rp-candidate** *interface-type interface-number* [**priority** *priority-value*] [**interval** *interval-seconds*] [**group-list** *ipv6_access-list*]
Parameter Description	*interface-type interface-number*: Indicates an interface name. This interface address is used as the C-RP address. **priority** *priority-value*: Specifies the priority of the C-RP. The value ranges from 0 to 255 and the default value is **192**. **interval** *seconds*: Indicates the interval for transmitting C-RP messages to the BSR, with the unit of seconds. The value ranges from 1 to 16,383 and the default value is **60**. **group-list** *ipv6_access-list*: References an IPv6 ACL to restrict the group address range served by the C-RP. A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	In the PIM-SMv6 protocol, the RPT created by the multicast routing uses an RP as the root node. After the BSR is elected, all C-RPs periodically transmit C-RP messages to the BSR in unicast mode and then the BSR disperses the messages in the entire PIM domain. To specify an interface as the C-RP of a specific group range, contain the ACL option in this command. Note that the calculation of the group range is based only on the permitted access control entries (ACEs) and denied ACEs are not involved in the calculation. If **group-list** ipv6_access-list is not carried in the command, all groups are served.

Configuring a C-BSR

Command	**ipv6 pim bsr-candidate** *interface-type interface-number* [*hash-mask-length* [*priority-value*]]
Parameter Description	*interface-type interface-number*: Indicates an interface name. This interface address is used as the C-BSR address. *hash-mask-length*: Indicates the hash mask length. The value ranges from 0 to 128 and the default value is **126**.

	priority-value: Indicates the priority. The value ranges from 0 to 255 and the default value is **64**.
Command Mode	Global configuration mode
Usage Guide	A unique BSR must exist in a PIM-SMv6 domain. The BSR collects and advertises RP information. The unique well-known BSR is elected from multiple C-BSRs by means of BSMs. All C-BSRs consider that they are the BSR before knowing the BSR information. They periodically transmit BSMs that contain the BSR address and priority in the PIM-SMv6 domain. This command can be used to enable a device to transmit one BSM to all PIM neighbors by using the allocated BSR address. Each neighbor compares the original BSR address with the address in the received BSM. If the IPv6 address in the received BSM is equal to or larger than its BSR address, the neighbor stores this address as the BSR address and forwards the BSM. Otherwise, the neighbor discards the BSM. The current device deems that it is the BSR till it receives a BSM from another C-BSR and learns that the C-BSR has a higher priority (or the same priority but a larger IPv6 address).

Enabling the SSM

Command	**ipv6 pim ssm** { **default** | **range** *ipv6_access-list* }
Parameter Description	**default**: The default group address range of SSM is FF3x::/32. **range** *ipv6_access-list*: References an IPv6 ACL to restrict the SSM group address range. A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	To apply SSM in a PIM-SMv6 network, you must configure this command.

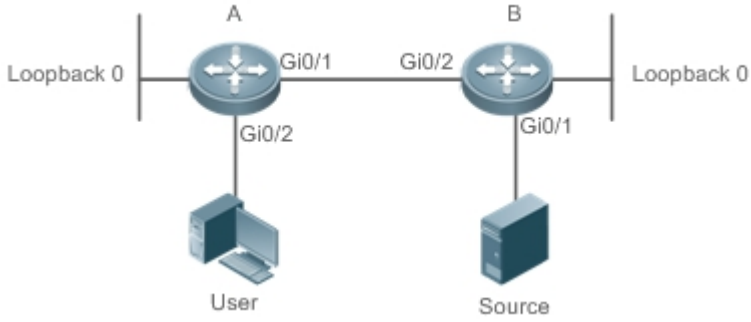
Displaying the PIM-SM Routing Table

Command	**show ipv6 pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]]
Parameter Description	*group-or-source-address*: Indicates the group address or source address. The two addresses cannot be group addresses or source addresses at the same time.
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	A group address, a source address, or both group address and source address can be specified each time. You can also not specify a specific group address or source address but you cannot specify two group addresses or two source addresses at the same time.

Configuration Example

Creating the IPv6 Multicast Service on an IPv6 Network to Support ASM and SSM

Scenario **Figure 7- 6**	

	
Configuration Steps	● Configure an IPv6 unicast routing protocol (such as OSPFv6) on the routers and ensure that the unicast routes of the loopback interfaces are reachable. (Omitted) ● Enable the IPv6 multicast routing function on all routers. ● Enable the PIM-SMv6 function on device interconnection interfaces, interface for connecting to the user host, and interface for connecting to the multicast source. ● Configure a C-RP and a C-BSR on the loopback interfaces of Router A and Router B. Enable the PIM-SMv6 function on the loopback interfaces. ● Enable SSM on all routers. ● Enable MLDv3 on the router interface for connecting to the user host. (Omitted)
A	switch(config)#ipv6 multicast-routing switch(config)#ipv6 pim ssm default switch(config)#int gi 0/2 switch(config-if-GigabitEthernet 0/2)#ipv6 add 2000::2/64 switch(config-if-GigabitEthernet 0/2)#ipv6 pim sparse-mode switch(config-if-GigabitEthernet 0/2)#exit switch(config)#int gi 0/1 switch(config-if-GigabitEthernet 0/1)#ipv6 add 1000::1/64 switch(config-if-GigabitEthernet 0/1)#ipv6 pim sparse-mode switch(config-if-GigabitEthernet 0/1)#exit switch(config)#int Loopback 0 switch(config-if-Loopback 0)#ipv6 add 3000::5/64 switch(config-if-Loopback 0)#ipv6 pim sparse-mode switch(config-if-Loopback 0)#exit switch(config)#ipv6 pim rp-candidate Loopback 0
B	FS(config)#ipv6 multicast-routing FS(config)#ipv6 pim ssm default

	FS(config)#int gi 0/2 FS(config-if-GigabitEthernet 0/2)#ipv6 add 2000::1/64 FS(config-if-GigabitEthernet 0/2)#ipv6 pim sparse-mode FS(config-if-GigabitEthernet 0/2)#exit FS(config)#int gi 0/1 FS(config-if-GigabitEthernet 0/1)#ipv6 add 1100::1/64 FS(config-if-GigabitEthernet 0/1)#ipv6 pim sparse-mode FS(config-if-GigabitEthernet 0/1)#exit FS(config)#int Loopback 0 FS(config-if-Loopback 0)#ipv6 add 5000::5/64 FS(config-if-Loopback 0)#ipv6 pim sparse-mode FS(config-if-Loopback 0)#exit FS(config)#ipv6 pim bsr-candidate Loopback 0
Verification	Make Source(2000::2/64) send packets to G1(ff16::1) and make User join G1. ● Check the multicast packets received by the User. The User should be able to receive multicast packets from G1. ● Check PIM-SMv6 routing tables on Router A and Router B. Entries should exist on the PIM-SMv6 routing tables.
A	switch(config)# show ipv6 pim sparse-mode mroute IPv6 Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 0 (*, ff16::1) RP: 3000::5 RPF nbr: :: RPF idx: None Upstream State: JOINED 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local

```
0 . . i . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Joined
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Asserted
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
FCR:

(1100::2, ff16::1)
RPF nbr: fe80::21a:a9ff:fe3a:6355
RPF idx: GigabitEthernet 0/2
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 44 seconds
kat expires in 194 seconds
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Joined
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Asserted
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . . o . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

(1100::2, ff16::1, rpt)
RP: 3000::5
RPF nbr: ::
```

	RPF idx: None Upstream State: PRUNED 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local 0 . . . . . . 1 . . . . . . Pruned 0 . . . . . . 1 . . . . . . Outgoing 0 . . o . . . . . . 1 . . . . . .
B	FS#show ipv6 pim sparse-mode mroute IPv6 Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 0 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 1 (1100::2, ff16::1) RPF nbr: :: RPF idx: None SPT bit: 1 Upstream State: JOINED kat expires in 20 seconds 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local 0 . . . . . . Joined 0 . j . . . . .

```
Asserted
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . o . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .


(1100::2, ff16::1, rpt)
RP: 3000::5
RPF nbr: fe80::2d0:f8ff:fe22:341b
RPF idx: GigabitEthernet 0/2
Upstream State: RPT NOT JOINED
    00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Pruned
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Outgoing
0 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
```

Common Errors

- IPv6 unicast routing is incorrectly configured.
- IPv6 multicast routing is not enabled on a router.
- SSM is not enabled on a router or the SSM group address range of the router is different from that of other routers.
- PIM-SMv6 is not enabled on an interface (for example, interface that is specified as a C-RP or C-BSR, or interface that functions as the gateway of a user host or multicast source).
- MLDv3 is not enabled on an interface connected to a user host.
- No RP is configured in the network.
- No static RP is configured on a router or the configured static RP is different from that on other routers.
- A C-RP is configured but no C-BSR is configured in the network.
- The unicast route to the static RP, C-RP, or C-BSR is unreachable.

7.4.2Configuring PIM Neighbor Parameters

Configuration Effect

- Negotiate about protocol parameters and adjust parameters in a Hello packet.
- PIM routers discover neighbors, negotiate about protocol parameters, and maintain neighbor relationships.
- Protect neighbor relationships to restrict neighbors.

Notes

- The basic functions of PIM-SMv6 must be configured.

Configuration Steps

- Set parameters on each PIM router interface unless otherwise specified.

Verification

Set parameters in a Hello packet on an interface and run the **debug ipv6 pim sparse-mode packet** command to check parameters in the Hello packet.

Set neighbor filtering and run the **show ipv6 pim sparse-mode neighbor** command to check the neighbor relationship.

Related Commands

Configuring the Transmission Interval of Hello Messages

Command	**ipv6 pim query-interval** *seconds*
Parameter Description	Indicates the transmission interval of Hello packets. The unit is seconds. The value ranges from 1 to 65,535 and the default value is **30**.
Command Mode	Interface configuration mode
Usage Guide	Each time the transmission interval of Hello messages is updated, the Holdtime of Hello messages is accordingly updated according to the following rule: The Holdtime of Hello messages is updated to 3.5 times transmission interval of Hello messages. If the transmission interval of Hello messages multiplied by 3.5 is larger than 65,535, the transmission interval of Hello messages is forcibly updated to 18,725.

Configuring the Propagation Delay for Hello Messages

Command	**Ipv6 pim propagation-delay** *milliseconds*
Parameter Description	*milliseconds*: The unit is milliseconds. The value ranges from 1 to 32,767 and the default value is **500**.
Command Mode	Interface configuration mode
Usage Guide	Changing the propagation delay or prune override delay will affect J/P-override-interval. According to the protocol, J/P-override-interval must be smaller than the Holdtime of Join-Prune packets. Otherwise, short flow interruption will be incurred. This must be maintained and guaranteed by network administrators.

Configuring the Prune Override Interval for Hello Messages

Command	**ipv6 pim override-interval** *milliseconds*
Parameter Description	*milliseconds*: The unit is milliseconds. The value ranges from 1 to 65,535 and the default value is **2500**.
Command Mode	Interface configuration mode
Usage Guide	Changing the propagation delay or prune override delay will affect J/P-override-interval. According to the protocol, J/P-override-interval must be smaller than the Holdtime of Join-Prune packets. Otherwise, short flow interruption will be

	incurred.

Configuring the Interface Joining Suppression Capability for Hello Messages

Command	**ipv6 pim neighbor-tracking**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	When the joining suppression capability of an interface is enabled and the local router needs to transmit a Join packet to an upstream neighbor, the Join packet of the local router is suppressed and is not transmitted if the local router receives a Join packet from a neighbor to the upstream router. If the joining suppression capability of the interface is disabled, the local router transmits the Join packet. When the joining suppression capability of a downstream receiver is disabled, the upstream neighbor can accurately know the number of receivers connected to the downstream neighbor through the received Join packet, thereby implementing neighbor tracking.

Configuring the Delay of Sending Out Hello Messages

Command	**Ipv6 pim triggered-hello-delay** *seconds*
Parameter Description	*Seconds*: The unit is seconds. The value ranges from 1 to 5 and the default value is **5**.
Command Mode	Interface configuration mode
Usage Guide	When an interface is enabled or detects a new neighbor, the Triggered-Hello-Delay message is used to generate a random time period. Within the time period, the interface sends Hello packets.

Configuring the DR Priority for Hello Messages

Command	**Ipv6 pim dr-priority** *priority-value*
Parameter Description	*priority-value*: Indicates the priority. A larger value means a higher priority. The value ranges from 0 to 4,294,967,294 and the default value is **1**.
Command Mode	Interface configuration mode
Usage Guide	The process of selecting a DR is as follows: The priority parameter is set for Hello packets of devices in the same LAN. The priority is compared for the selection of a DR. The device with a higher priority is the DR. If multiple devices share the same priority, the device with a larger IP address is the DR. When the priority parameter is not set for Hello packets of a device in a LAN, the device with a larger IP address is elected as the DR in the LAN.

Configuring Neighbor filtering

Command	**ipv6 pim neighbor-filter** *ipv6_access-list*
Parameter Description	*ipv6_access-list*: References an IPv6 ACL to restrict the neighbor address range.

Command Mode	Interface configuration mode
Usage Guide	This command can be used to filter neighbors to strengthen the security of the PIM network and restrict the address range of legitimate neighbors. If a neighbor is rejected by an ACL, PIM-SMv6 will not establish a peering relationship with this neighbor or suspend the peering relationship with this neighbor.

Displaying Neighbor Information About an Interface

Command	**show ipv6 pim sparse-mode neighbor** [**detail**]
Parameter Description	**detail**: Displays details.
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	N/A

Configuration Example

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Set the transmission interval of Hello packets of PIM-SMv6 to 50 seconds. ● Set the propagation delay of Hello packets of PIM-SMv6 to 400 milliseconds. ● Set the prune override interval of Hello packets of PIM-SMv6 to 3,000 milliseconds. ● Configure the interface joining suppression capability for Hello messages of PIM-SMv6. ● Set the delay of sending out Hello messages of PIM-SMv6 to 3 seconds. ● Set the DR priority of Hello messages of PIM-SMv6 to 5.
	switch # configure terminal switch (config)#int gi 0/1 switch (config-if-GigabitEthernet 0/1)#ipv6 pim query-interval 50 switch (config-if-GigabitEthernet 0/1)#ipv6 pim propagation-delay 400 switch (config-if-GigabitEthernet 0/1)#ipv6 pim override-interval 3000 switch (config-if-GigabitEthernet 0/1)#ipv6 pim triggered-hello-delay 3 switch (config-if-GigabitEthernet 0/1)# ipv6 pim dr-priority 5
Verification	Run the **debug ipv6 pim sparse-mode packet** command to check parameters in a Hello packet.

	switch # debug ipv6 pim sparse-mode packet *Jan 2 02:37:55: %7: Hello send to GigabitEthernet 0/2 *Jan 2 02:37:55: %7: Send Hello message *Jan 2 02:37:55: %7: Holdtime: 175 *Jan 2 02:37:55: %7: T-bit: off *Jan 2 02:37:55: %7: Propagation delay: 400 *Jan 2 02:37:55: %7: Override interval: 3000 *Jan 2 02:37:55: %7: DR priority: 5 *Jan 2 02:37:55: %7: Gen ID: 99572792 *Jan 2 02:37:55: %7: Secondary Addresses: *Jan 2 02:37:55: %7: 2000::2
Configuration Steps	Configure neighbor filtering on an interface to receive neighbor packets with the address of (8000::1/64).
	switch(config-if-GigabitEthernet 0/2)#ipv6 pim neighbor-filter acl % access-list acl not exist switch(config-if-GigabitEthernet 0/2)#exit switch(config)#ipv6 access-list acl switch(config-ipv6-acl)#permit ipv6 8000::1/64 any
Verification	Before neighbor filtering is configured, display the neighbor information.
	switch#show ipv6 pim sparse-mode neighbor Neighbor Address Interface Uptime/Expires DR Pri/Mode fe80::21a:a9ff:fe3a:6355 GigabitEthernet 0/2 00:32:29/00:01:16 1 /
Verification	After neighbor filtering is configured, the neighbor information is blank.
	switch#show ipv6 pim sparse-mode neighbor

Common Errors

- The basic functions of PIM-SMv6 are not configured or fail to be configured.

7.4.3 Configuring BSR Parameters

Configuration Effect

- Restrict the range of BSMs.

Notes

- The basic functions of PIM-SMv6 must be configured.
- A C-RP and a C-BSR must be configured.
- The border must be configured on the interface between domains.

Configuration Steps

Configuring the Border

- The border must be configured if there are multiple domains.
- Configure the border the interface between two domains.

Configuring a PIM Router to Restrict BSMs

- Optional.
- This configuration can be performed on a PIM router unless otherwise specified.

Configuring a C-BSR to Restrict the C-PR Range

- Optional.
- This configuration can be performed on all C-BSRs unless otherwise specified.

Configuring a C-BSR to Receive C-RP-ADV Packets with prefix-count of 0

- Optional.
- This configuration can be performed on all C-BSRs unless otherwise specified.

Verification

Verifying the Border

Enable the basic functions of PIM-SMv6, set two routers in different domains, and set Router B as a C-BSR. Router A can normally receive BSMs.

Set the common border between Router A and Router B as a border interface. Router A cannot receive BSMs.

Verifying a PIM Router to Restrict BSMs

Enable the basic functions of PIM-SMv6 and set Router B as a C-BSR. Router A can normally receive BSMs. Restrict the C-BSR range on Router A. Router A cannot receive BSMs.

Verifying a C-BSR to Restrict the C-PR Range

Enable the basic functions of PIM-SMv6, set Router B as a C-BSR, set Router A as a C-RP, and restrict the C-RP range on the C-BSR. Router B cannot receive packets from the C-RP.

Related Commands

Configuring the BSR Border

Command	**ipv6 pim bsr-border**

Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The BSR border can be configured on an interface to restrict flooding of BSMs. When this interface receives BSMs, it immediately discards them and BSMs are not forwarded by this interface.

Configuring a PIM Router to Restrict BSMs

Command	**ipv6 pim accept-bsr list** *ipv6_access-list*
Parameter Description	**list** *ipv6_access-list*: References an IPv6 ACL to restrict the BSR address range. A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a C-BSR to Restrict the C-PR Range

Command	**ipv6 pim accept-crp list** *ipv6_access-list*
Parameter Description	**list** *ipv6_access-list*: References an IPv6 ACL to restrict the address range of the C-RP and the group address range served by the C-RP. A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	Configure this command on a C-BSR. When this C-BSR is elected as the BSR, it can restrict the address range of the valid C-RP and the multicast group range served by the C-RP.

Displaying BSMs

Command	**show ipv6 pim sparse-mode bsr-router**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	N/A

Displaying All RPs Configured on the Local Device and the Multicast Groups Served by the RPs

Command	**show ipv6 pim sparse-mode rp mapping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring the BSR Border

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Configure the BSR border on the juncture interface between Router B and Router A.
	FS(config-if-GigabitEthernet 0/2)#ipv6 pim bsr-border
Verification	Before the BSR border is configured, the BSM information of Router A is displayed as follows:
	switch#show ipv6 pim sparse-mode bsr-router PIMv2 Bootstrap information BSR address: 5000::5 Uptime: 00:05:42, BSR Priority: 64, Hash mask length: 126 Expires: 00:01:28 Role: Non-candidate BSR Priority: 0, Hash mask length: 126 State: Accept Preferred Candidate RP: 3000::5(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:24 switch#
	Candidate RP: Indicates all C-RPs configured on the local router, excluding other routers.
	After the BSR border is configured, the BSM information of Router A is displayed as follows:
	switch#show ipv6 pim sparse-mode bsr-router Candidate RP: 3000::5(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:53

Configuring a PIM Router to Restrict the Source Address Range of BSMs to (8000::5/64)

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Configure PIM Router A to restrict BSMs. The restricted source address range is (8000::5/64).
	switch(config)#ipv6 access-list acl switch(config-ipv6-acl)#permit ipv6 8000::5/64 any switch(config-ipv6-acl)#exit switch(config)#ipv6 pim accept-crp list acl

Verification	Before the BSM restriction is configured, the BSM information of Router A is displayed as follows:
	switch#show ipv6 pim sparse-mode bsr-router PIMv2 Bootstrap information BSR address: 5000::5 Uptime: 00:05:42, BSR Priority: 64, Hash mask length: 126 Expires: 00:01:28 Role: Non-candidate BSR Priority: 0, Hash mask length: 126 State: Accept Preferred Candidate RP: 3000::5(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:24 switch#
	After the BSM restriction is configured, the BSM information of Router A is displayed as follows:
	switch#show ipv6 pim sparse-mode bsr-router Candidate RP: 3000::5(Loopback 0) Advertisement interval 60 seconds Next Cand_RP_advertisement in 00:00:34

Configuring a C-BSR to Restrict the Source Address Range of C-PR Packets to (9000::5/64)

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Configure Router B to restrict C-RP packets. The restricted source address range is (9000::5/64).
	FS(config)#ipv6 access-list acl FS(config-ipv6-acl)#permit ipv6 9000::5/64 any FS(config-ipv6-acl)#exit FS(config)#ipv6 pim accept-crp list acl
Verification	Before C-RP packet filtering is configured, information about all RP groups on Router B is displayed as follows:

	FS#show ipv6 pim sparse-mode rp mapping PIM Group-to-RP Mappings This system is the Bootstrap Router (v2) Group(s): ff00::/8 RP: 3000::5(Not self) Info source: 3000::5, via bootstrap, priority 192 Uptime: 00:02:26, expires: 00:02:08 FS#
	After C-RP packet filtering is configured, information about all RP groups on Router B is displayed as follows:
	FS#show ipv6 pim sparse-mode rp mapping PIM Group-to-RP Mappings This system is the Bootstrap Router (v2)

Configuring the Static RP First

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Set the address of Interface Loopback0 of Router A to 3000::5. (Omitted) ● Set the address of Interface Loopback1 of Router A to 4000::5. (Omitted) ● Set the static address of Router A to 3300::5. (Omitted) ● Set the static address of Router B to 3300::5. ● Configure the static RP first on Router A.
	switch(config)#ipv6 pim rp-address 3300::5 switch(config)#ipv6 pim static-rp-preferred
Verification	Before static RP first is configured, display information about the RP corresponding to FF16::1.
	switch#show ipv6 pim sparse-mode rp ff16::1 RP: 4000::5 Info source: 5000::5, via bootstrap PIMv2 Hash Value 126 RP 4000::5, via bootstrap, priority 56, hash value 892666309 RP 3000::5, via bootstrap, priority 200, hash value 1161101765 RP 3300::5, static (hash value 204800453 not used)
Verification	After static RP first is configured, display information about the RP corresponding to FF16::1.

```
switch(config)#show ipv6 pim sparse-mode rp ff16::1
RP: 3300::5    (Static)

PIMv2 STATIC RP PREFERRED
PIMv2 Hash Value 126
RP 4000::5, via bootstrap, priority 56, hash value 892666309
RP 3000::5, via bootstrap, priority 200, hash value 1161101765
RP 3300::5, static (hash value 204800453 not used)

switch(config)#
```

Common Errors

- The basic functions of PIM-SMv6 are not configured or fail to be configured.
- No C-BSR is configured.
- The BSR border is not configured on an interface between different domains.

7.4.4 Configuring RP and DR Parameters

Configuration Effect

- Configure the ignorance of the C-RP priority for the RP reselection.
- Configure the DR at the data source end to detect the RP reachability.
- Restrict the (S,G) multicast group address of the data source so that the ASM model provides the multicast service only for multicast packets within the allowable range.
- Configure the rate limit for the DR at the data source end to transmit Register packets.
- Configure the checksum length of Register packets.
- Configure the source address of Register packets.
- Configure the suppression time of Register packets.
- Configure the probing time of NULL packets.
- Configure the TTL of Register packets received by the RP from the (S,G) multicast group address.
- Configure the static RP first.

Notes

- The basic functions of PIM-SMv6 must be configured.

Configuration Steps

Configuring the Ignorance of the C-RP Priority for the RP Reselection

- Optional.

- The ignorance of the C-RP priority can be enabled on each router unless otherwise specified.

Configuring the DR at the Data Source End to Detect the RP Reachability

- Optional.
- The reachability detection can be enabled on the DR that is directly connected to the data source unless otherwise specified.

Restricting the (S,G) Address Range of Register Packets at the Data Source End

- Optional.
- The (S,G) address range of Register packets at the data source end can be restricted on all routers that function as C-RPs or static RPs unless otherwise specified.

Restricting the Rate for the DR at the Data Source End to Transmit Register Packets

- Optional.
- The transmission rate limit of Register packets can be enabled on the DR that is directly connected to the data source unless otherwise specified.

Configuring the Checksum Length of Register Packets

- Optional.
- The checksum length of Register packets can be configured on all C-RPs or static RPs unless otherwise specified.

Configuring the Source Address of Register Packets

- Optional.
- The source address of Register packets can be configured on the DR that is directly connected to the data source unless otherwise specified.

Configuring the Suppression Time of Register Packets

- Optional.
- The suppression time of Register packets can be configured on the DR that is directly connected to the data source unless otherwise specified.

Configuring the Probing Time of NULL Packets

- Optional.
- The probing time of NULL packets can be configured on the DR that is directly connected to the data source unless otherwise specified.

Configuring the TTL of Register Packets Received by the RP from the (S,G) Multicast Group Address

- Optional.
- The TTL of Register packets from the (S,G) multicast group address can be configured on all routers that function as C-RPs or static RPs unless otherwise specified.

Configuring the Static RP First

- Optional.

- The static RP first can be configured on all routers unless otherwise specified.

Verification

Verifying the Ignorance of the C-RP Priority

Set the address to 3000::5 and priority to 200 for Interface Loopback0 on Router A. Set the address to 4000: : 5 and priority to 56 for Interface Loopback1 on Router A. Set the C-BSR address to 5000: : 5 on Router B.

- Run the **show ipv6 pim sparse-mode rp** ff16::2 command to display information about the RP that serves the current group.

Verifying the DR at the Data Source End to Detect the RP Reachability

Set the address to 3000::5 and priority to 200 for Interface Loopback0 on Router A. Set the address to 4000: : 5 and priority to 56 for Interface Loopback1 on Router A. Set the C-BSR address to 5000: : 5 on Router B. Configure the RP reachability detection on Router B.

- Run the **show running-config** command to check whether the RP reachability detection is configured.

Verifying the Restriction of the (S,G) Address Range of Register Packets at the Data Source End

Set the address to 3000::5 and priority to 200 for Interface Loopback0 on Router A. Set the address to 4000: : 5 and priority to 56 for Interface Loopback1 on Router A. Set the C-BSR address to 5000: : 5 on Router B. The address of the multicast group is FF16::2. Set Router A to receive packets only from the multicast source with the source address of (1300::1/64).

- Run the **show ip pim sparse-mode mroute** command to display the (S,G) entries.

Verifying the Rate Limit for the DR at the Data Source End to Transmit Register Packets

- Set the rate of transmitting Register packets for Router B and then run the **show ip pim sparse-mode track** command to check the number of transmitted Register packets for confirmation.

Verifying the Checksum Length of Register Packets

- Set Router A to check a Register packet based on the entire packet rather than based only on the packet header and Register packet header. Run the **show running-config** command to check the configuration.

Verifying the Source Address of Register Packets

- Configure the source address of Register packets on Router B and run the **show running-config** command to check the configuration on Router A.

Verifying the Suppression Time and Probing Time of Register Packets

- Configure the suppression time and probing time of Register packets on Router B and run the **show running-config** command to check the configuration.

Verifying the TTL of Register Packets Received by the RP from the (S,G) Multicast Group Address

- Configure the TTL of Register packets from the (S,G) multicast group address on Router A and run the **show ip pim sparse-mode mroute** command to display the maximum (S,G) TTL.

Verifying the Static RP First

- Configure a static RP and a C-RP on Router A, configure the static RP first, and then run the **show ipv6 pim sparse-mode rp** ff16::2 command to display information about the current RP.

Related Commands

Ignoring the C-RP Priority

Command	**ipv6 pim ignore-rp-set-priority**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Displaying Information About the RP That Serves a Group

Command	**show ipv6 pim sparse-mode rp-hash** *group-address*
Parameter Description	*group-address*: Indicates the parsed group address.
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	N/A

Configuring the DR Directly Connected to the Data Source to Detect RP Reachability

Command	**ipv6 pim register-rp-reachability**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After this command is configured, the RP reachability is detected before Register packets are transmitted. If the RP is reachable, Register packets are transmitted. If the RP is unreachable, Register packets are not transmitted.

Restricting the (S,G) Address Range of Register Packets at the Data Source End

Command	**ipv6 pim accept-register** { **list** *ipv6_access-list* [**route-map** *map-name*] | **route-map** *map-name* [**list** *ipv6_access-list*] }
Parameter Description	**list** *ipv6_access-list*: References an IP extended ACL to restrict the (S,G) address range. The value range is 100-199, 2000-2699, and Word. **route-map** *map-name*: Uses a route map to restrict the (S,G) address range.
Command Mode	Global configuration mode
Usage Guide	After this command is configured, when receiving a Register packet from an unauthorized source, the RP immediately returns the Register-Stop packet.

Displaying Multicast Routing Entries

Command	**show ipv6 pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]]
Parameter	*group-or-source-address*: Indicates the group address or source address. The two addresses cannot be group addresses or

Description	source addresses at the same time.
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	A group address, a source address, or both addresses can be specified each time. You can also not specify a specific group address or source address but you cannot specify two group addresses or two source addresses at the same time.

Configuring the Rate Limit for the DR to Transmit Register Packets

Command	**ipv6 pim register-rate-limit** *rate*
Parameter Description	*Rate*: Indicates the number of Register packets that are allowed to be transmitted per second. The value ranges from 1 to 65,535.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the transmission rate of Register packets from the (S,G) multicast group address rather than the Register packets of the entire system. After this command is configured, the load of the source DR and RP will be relieved and Register packets whose rate does not exceed the limit will be transmitted.

Displaying the Statistics on PIM Packets

Command	**show ipv6 pim sparse-mode track**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	When the system is started, the statistics start time point is first set. Each time **clear ip pim sparse-mode track** is called, the statistics start time point is set again and the PIM packet counter is cleared.

Configuring the Checksum Calculation of a Register Packet Based on the Entire Packet

Command	**ipv6 pim register-checksum-wholepkt** [**group-list** *ipv6_access-list*]
Parameter Description	**group-list** *access-list*: Uses an ACL to restrict the group addresses that use this configuration. access-list: Supports digits <1,99> and <1300,1999>. A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	The device calculates the checksum of a Register packet based on the entire PIM protocol packet including the encapsulated multicast data packet, rather than the PIM header of the Register packet. If **group-list** ipv6_access-list is not carried in this command, all group addresses apply this configuration.

Configuring the Source Address of Register Packets

Command	**ipv6 pim register-source** { *ipv6_local_address* | *interface-type interface-number* }
Parameter Description	*local_address*: Specifies an IPv6 address as the source address of Register packets. *interface-type interface-number*: Specifies the IPv6 address of an interface as the source address of Register packets.
Command Mode	Global configuration mode

Usage Guide	The configured address must be reachable. When the RP receives a Register packet, it transmits the Register-Stop packet with the source IPv6 address of the Register packet as the destination address. PIM-SMv6 does not need to be enabled on associated interfaces.

Configuring the Suppression Time of Register Packets

Command	**ipv6 pim register-suppression** *seconds*
Parameter Description	*Seconds*: Indicates the suppression time of Register packets. The unit is seconds. The value ranges from 1 to 65,535 and the default value is **60**.
Command Mode	Global configuration mode
Usage Guide	Configuring this value on the DR will change the suppression time of Register packets defined on the DR. If the **ipv6 pim rp-register-kat** command is not configured, configuring this value on the RP will change the keepalive time of the RP.

Configuring the Probing Time of Register Packets

Command	**ipv6 pim probe-interval** *seconds*
Parameter Description	*Seconds*: Indicates the probing time of Register packets. The unit is seconds. The value ranges from 1 to 65,535 and the default value is **5**.
Command Mode	Global configuration mode
Usage Guide	Probing time of Register packets is the interval for the source DR to transmit the NULL-Register packet to the RP prior to the timeout of the suppression time of Register packets. The probing time of Register packets cannot be larger than half of the suppression time of Register packets. Otherwise, the configuration fails and a warning is displayed. In addition, the suppression time of Register packets multiplied by three plus the probing time of Register packets cannot be larger than 65,535. Otherwise, a warning will be displayed.

Configuring the KAT Interval on the RP

Command	**Ipv6 pim rp-register-kat** *seconds*
Parameter Description	*Seconds*: Indicates the time of the KAT timer. The unit is seconds. The value ranges from 1 to 65,535 and the default value is **210**.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Static RP First

Command	**ipv6 pim static-rp-preferred**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After this command is configured, the priority of the static RP is higher than that of the RP elected by using the BSR mechanism.

Configuration Example

Configuring Whether the C-RP Priority Is Considered for the Group-to-RP Mapping

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Set the address to 3000::5 and priority to 200 for Interface Loopback0 on Router A. (Omitted) ● Set the address to 4000: : 5 and priority to 56 for Interface Loopback1 on Router A. (Omitted) ● Set the C-BSR address to 5000: : 5 on Router B. (Omitted) ● Display the group corresponding to FF16::1. ● Configure the ignorance of C-RP priority on Router B.
	switch#configure terminal FS(config)# ipv6 pim ignore-rp-set-priority
Verification	Before the ignorance of the C-RP priority is configured, the following information is displayed:
	switch(config)#show ipv6 pim sparse-mode rp FF16::1 RP: 4000::5 Info source: 5000::5, via bootstrap PIMv2 Hash Value 126 RP 4000::5, via bootstrap, priority 56, hash value 892666309 RP 3000::5, via bootstrap, priority 200, hash value 1161101765
	After the ignorance of the C-RP priority is configured, the following information is displayed:
	switch(config)#show ipv6 pim sparse-mode rp FF16::1 RP: 3000::5 Info source: 5000::5, via bootstrap

Configuring the Reachability Detection of the RP Directly Connected to the Data Source

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Configure the reachability detection of the RP directly connected to the data source.
	FS(config)#ipv6 pim register-rp-reachability
Verification	Run the **show running-config** command to check the configuration. The following information is displayed:
	FS(config)#show running-config ! ! ! ipv6 pim register-rp-reachability

	ipv6 pim bsr-candidate Loopback 0 ! !

Restricting the (S,G) Address Range of Register Packets at the Data Source End

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Set Router A to filter packets by source address and receive packets only from the source address (1300::1/64).
	switch(config)#ipv6 pim accept-register list acl % access-list 101 not exist switch(config)#ipv6 access-list acl switch(config-ipv6-acl)#permit ipv6 1300::1/64 any switch(config-ipv6-acl)#exit
Verification	Before the (S,G) address range of Register packets at the data source end is restricted, run the **show ipv6 pim sparse-mode mroute** command to display multicast entries. The (S,G) entry and (S,G,RPT) entry exist.
	switch#show ipv6 pim sparse-mode mroute IPv6 Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 0 (*, ff16::1) RP: 4000::5 RPF nbr: :: RPF idx: None Upstream State: JOINED 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local 0 . . . i . . . . . . 1 . . . . . . Joined

	0 . . . . 1 . . . . Asserted 0 . . . . 1 . . . . FCR: (1100::2, ff16::1) RPF nbr: fe80::21a:a9ff:fe3a:6355 RPF idx: GigabitEthernet 0/2 SPT bit: 1 Upstream State: JOINED jt_timer expires in 36 seconds kat expires in 191 seconds 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Local
	After the (S,G) address range of Register packets at the data source end is restricted, run the **show ipv6 pim sparse-mode mroute** command to display multicast entries. The (S,G) entry and (S,G,RPT) entry exist.
	switch#show ipv6 pim sparse-mode mroute IPv6 Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 0 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 0 (*, ff16::1) RP: 4000::5 RPF nbr: :: RPF idx: None Upstream State: JOINED

```
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  i  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
FCR:


(1100::2, ff16::1, rpt)
RP: 4000::5
RPF nbr: ::
RPF idx: None
Upstream State: PRUNED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
```

Restricting the Rate for the DR at the Data Source End to Transmit Register Packets

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Check the number of PIM packets transmitted by Router B. ● Check the number of PIM packets transmitted by Router B one second later. ● Set the rate for Router B to transmit Register packets. ● Check the number of PIM packets transmitted by Router B one second later.
	`FS(config)#ipv6 pim register-rate-limit 1`
Verification	Before the rate limit is configured, check the number of PIM packets transmitted by the DR. The following information is displayed:

	`FS#show ipv6 pim sparse-mode track`
	`PIMv6 packet counters track`
	`Elapsed time since counters cleared: 17:14:54`

	received	sent
Valid PIMv6 packets:	5064	7727
Hello:	1329	4057
Join-Prune:	863	0
Register:	0	2636
Register-Stop:	975	0
Assert:	0	0
BSM:	0	1034
C-RP-ADV:	1897	0
PIMDM-Graft:	0	
PIMDM-Graft-Ack:	0	
PIMDM-State-Refresh:	0	
Unknown PIM Type:	0	

Errors:

Malformed packets: 0

Bad checksums: 0

Send errors: 5

Packets received with unknown PIM version: 0

Before the rate limit is configured, check the number of PIM packets transmitted by the DR one second later. The following information is displayed:

FS#show ipv6 pim sparse-mode track

PIMv6 packet counters track

Elapsed time since counters cleared: 17:14:55

	received	sent
Valid PIMv6 packets:	5064	7727
Hello:	1335	4063
Join-Prune:	866	0
Register:	0	2639
Register-Stop:	978	0
Assert:	0	0
BSM:	0	1035
C-RP-ADV:	1897	0

	PIMDM-Graft: 0 PIMDM-Graft-Ack: 0 PIMDM-State-Refresh: 0 Unknown PIM Type: 0 Errors: Malformed packets: 0 Bad checksums: 0 Send errors: 5 Packets received with unknown PIM version: 0
	After the rate limit is configured, check the number of PIM packets transmitted by the DR. The following information is displayed:
	FS#show ipv6 pim sparse-mode track PIMv6 packet counters track Elapsed time since counters cleared: 17:14:56 received sent Valid PIMv6 packets: 5064 7727 Hello: 1341 4069 Join-Prune: 869 0 Register: 0 2640 Register-Stop: 979 0 Assert: 0 0 BSM: 0 1036 C-RP-ADV: 1897 0 PIMDM-Graft: 0 PIMDM-Graft-Ack: 0 PIMDM-State-Refresh: 0 Unknown PIM Type: 0 Errors: Malformed packets: 0 Bad checksums: 0 Send errors: 5

	Packets received with unknown PIM version: 0

Configuring the Checksum Length of Register Packets

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Configure the checksum calculation of a Register packet based on the entire packet on Router A. ● Run the **show running-config** command to check the configuration.
	switch(config)#ipv6 pim register-checksum-wholepkt switch(config)#show running-config
Verification	Check the configuration on Router A. The configuration is displayed as follows:
	! ! ipv6 pim register-checksum-wholepkt ipv6 pim rp-candidate Loopback 0 priority 200 ipv6 pim rp-candidate Loopback 1 priority 56 ipv6 pim ssm default ! !

Configuring the Source Address of Register Packets

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Set the source address of Interface Loopback1 to 5500::5/64 on Router B. (Omitted) ● Set the source address of Register packets to the address of Interface Loopback2 on Router B. (Omitted) ● Run the **show running-config** command to check the configuration.
	FS(config)#ipv6 pim register-source Loopback 1
Verification	Check the configuration on Router B.
	! ! ipv6 pim register-source Loopback 1 ipv6 pim register-rate-limit 1 ipv6 pim bsr-candidate Loopback 0 ! !

Configuring the Suppression Time and Probing Time of Register Packets

Configuration	● Configure basic functions of PIM-SMv6. (Omitted)

Steps	● Set the suppression time to 20 seconds on Router B. ● Set the probing time to 2 seconds on Router B. ● Run the **show running-config** command to check the configuration.
	FS(config)#ipv6 pim register-suppression 20 FS(config)#ipv6 pim probe-interval 2 FS(config)# show ip pim sparse-mode track
Verification	Check the configuration on Router B.
	! ipv6 pim register-source Loopback 1 ipv6 pim register-rate-limit 1 ipv6 pim register-suppression 20 ipv6 pim probe-interval 2 ipv6 pim bsr-candidate Loopback 0 ! !

Configuring the TTL of Register Packets Received by the RP from the (S,G) Multicast Group Address

Configuration Steps	● Configure basic functions of PIM-SMv6. (Omitted) ● Se the TTL of Register packets received by Router A from the (S,G) multicast group address to 60 seconds. ● Run the **show ip pim sparse-mode mroute** command to check the number of Register packets.
	FS(config)#ip pim rp-register-kat 60
Verification	After the TTL is configured, check the TTL of Register packets from the (S,G) multicast group address on Router A. The TTL is not larger than 60 seconds.
	switch(config)#show ipv6 pim sparse-mode mroute IPv6 Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 0 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 0 (1100::2, ff16::1)

```
RPF nbr: fe80::21a:a9ff:fe3a:6355
RPF idx: GigabitEthernet 0/2
SPT bit: 0
Upstream State: NOT JOINED
kat expires in 60 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Outgoing
0   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .


(1100::2, ff16::1, rpt)
RP: 4000::5
RPF nbr: ::
RPF idx: None
```

Common Errors

- The basic functions of PIM-SMv6 are not configured or fail to be configured.
- The (S,G) address range of Register packets at the data source end is not restricted or fails to be configured on a C-RP or static RP.
- When the (S,G) address range of Register packets at the data source end is restricted, the referenced ACL is not configured or the source/group address range allowed by the ACL is configured incorrectly.
- The source/group address ranges allowed by C-RPs or static RPs are inconsistent.

7.4.5 Configuring the Transmission Interval of Join/Prune Packets

Configuration Effect

- Change the transmission interval of Join/Prune packets to form an RPT or SPT.

Notes

- The basic functions of PIM-SMv6 must be configured.

Configuration Steps

- Configure the transmission interval of Join/Prune packets.

Verification

Set the transmission interval of Join/Prune packets to 120 seconds on Router B. Run the **show ipv6 pim sparse-mode mroute** command to check the entry TTL.

Related Commands

Configuring the Transmission Interval of Join/Prune Packets

Command	**ipv6 pim jp-timer** *seconds*
Parameter Description	*Seconds*: Indicates the transmission interval of Join/Prune packets. The unit is seconds. The value ranges from 1 to 65,535 and the default value is **60**.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Transmission Interval of Join/Prune Packets on a Router

Configuration Steps	• Configure basic functions of PIM-SMv6. (Omitted) • Configure the transmission interval of Join/Prune packets on a router.
	FS(config)#ip pim jp-timer 120
Verification	Run the **show ipv6 pim sparse-mode mroute** command to check the entry. The transmission time of Join/Prune packets is not larger than 120.
	switch(config)#show ipv6 pim sparse-mode mroute IPv6 Multicast Routing Table (*,*,RP) Entries: 0 (*,G) Entries: 1 (S,G) Entries: 1 (S,G,rpt) Entries: 1 FCR Entries: 0 REG Entries: 0 (*, ff16::1) RP: 4000::5

```
RPF nbr: ::
RPF idx: None
Upstream State: JOINED
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  i  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Joined
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Asserted
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
1  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
FCR:


(1100::2, ff16::1)
RPF nbr: fe80::21a:a9ff:fe3a:6355
RPF idx: GigabitEthernet 0/2
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 116 seconds
kat expires in 59 seconds
   00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
```

Common Errors

- The basic functions of PIM-SMv6 are not configured or fail to be configured.

7.4.6 Configuring the Last-Hop Device to Switch from the RPT to the SPT

Configuration Effect

- Switch the last-hop device from the RPT to the SPT.

Notes

- The basic functions of PIM-SMv6 must be configured.

Configuration Steps

- Configure the last-hop device to switch from the RPT to the SPT.

Verification

Configure basic functions of PIM-SMv6, make the DR at the data source end transmit data streams to Group FF16::1, and make the receiver forcibly join the Group FF16::1 to form a RPT. The DR at the receive end forcibly performs the switching from the RPT to SPT. Check the configuration on the RP.

Related Commands

Enabling the SPT Switching Function

Command	**ipv6 pim spt-threshold** [**group-list** *ipv6_access-list*]
Parameter Description	**group-list** *ipv6_access-list*: References an IPv6 ACL to restrict the group address range that allows SPT switching. ipv6_access-list: A named ACL is supported.
Command Mode	Global configuration mode
Usage Guide	If **group-list** ipv6_access-list parameter is not carried in this command, all multicast groups are allowed to conduct SPT switching. If **no** is set in this command, **group-list** is carried, and the carried ACL is a configured ACL, the restriction of the ACL associated with **group-list** is cancelled and all groups are allowed to switch from the RPT to the SPT.

Configuration Example

Configuring the Last-Hop Device to Switch from the RPT to the SPT

Configuration Steps	• Configure basic functions of PIM-SMv6. (Omitted) • Make the DR at the data source end transmit code streams to Group FF16::1. • Make the DR at the receive end receive code streams from Group FF16::1. • Configure the last-hop device to switch from the RPT to the SPT on the DR at the receive end.
	switch(config)#ipv6 pim spt-threshold
Verification	Run the **show running-config** command to check the configuration.

```
switch(config)#show running-config
!
!
ip pim jp-timer 120
ip pim spt-threshold
ip pim rp-candidate Loopback 0
!
!
```

7.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears information about the dynamic RP.	**clear ipv6 pim sparse-mode bsr rp-set** *
Sets the packet statistics start time again and clears the PIMv6 packet counter.	**clear ipv6 pim sparse-mode track**

Displaying

Description	Command
Displays details about the BSR.	**show ipv6 pim sparse-mode bsr-router**
Displays the PIM-SM information about an interface.	**show ipv6 pim sparse-mode interface** [*interface-type interface-number*] [**detail**]
Displays the local MLD information about a PIM-SMv6 interface.	**show ipv6 pim sparse-mode local-members** [*interface-type interface-number*]
Displays the PIM-SMv6 routing information.	**show ipv6 pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]]
Displays the PIM-SMv6 neighbor information.	**show ipv6 pim sparse-mode neighbor** [**detail**]
Displays next hop-relevant information, including the next-hop interface ID, address, and metric.	**show ipv6 pim sparse-mode nexthop**
Displays all RPs configured on the local device and the groups served by the RPs.	**show ipv6 pim sparse-mode rp mapping**
Displays information about the RP that serves the group address.	**show ipv6 pim sparse-mode rp-hash** *ipv6-group-address*
Displays the number of PIM packets transmitted and received from the statistics start time to the current time.	**show ipv6 pim sparse-mode track**

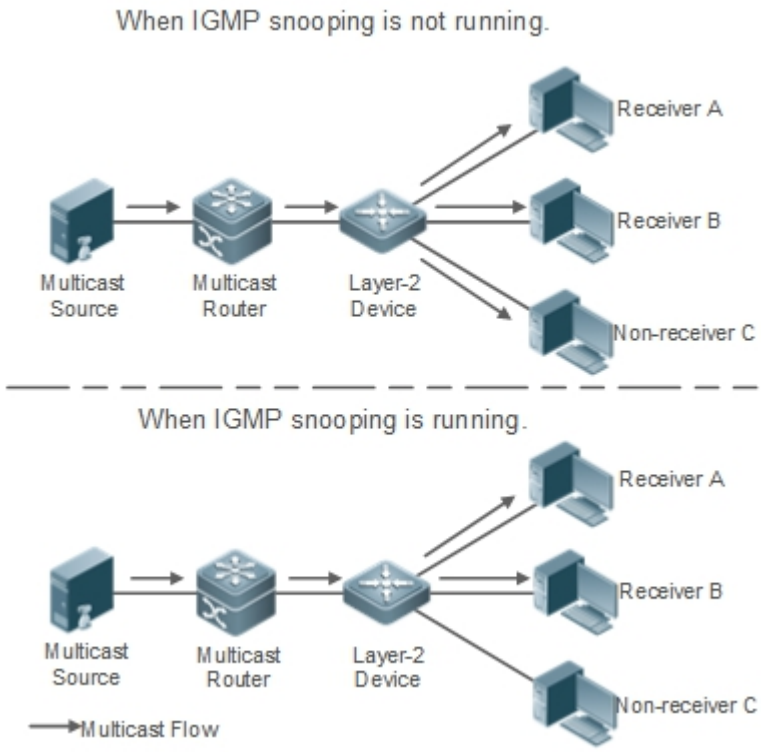
8 Configuring IGMP Snooping

8.1 Overview

Internet Group Management Protocol (IGMP) snooping is a mechanism of listening to IP multicast. It is used to manage and control the forwarding of IP multicast traffic within VLANs, realizing Layer-2 multicasting.

As shown in the following figure, when a Layer-2 device is not running IGMP snooping, IP multicast packets are broadcasted within the VLAN; when the Layer-2 device is running IGMP snooping, IP multicast packets are transmitted only to profile members.

Figure 2-1 Networking Topology of IP Multicast Forwarding within the VLAN Before and After IGMP Snooping Is Run on the Layer-2 Device



Protocols and Standards

- RFC4541: Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches

8.2 Applications

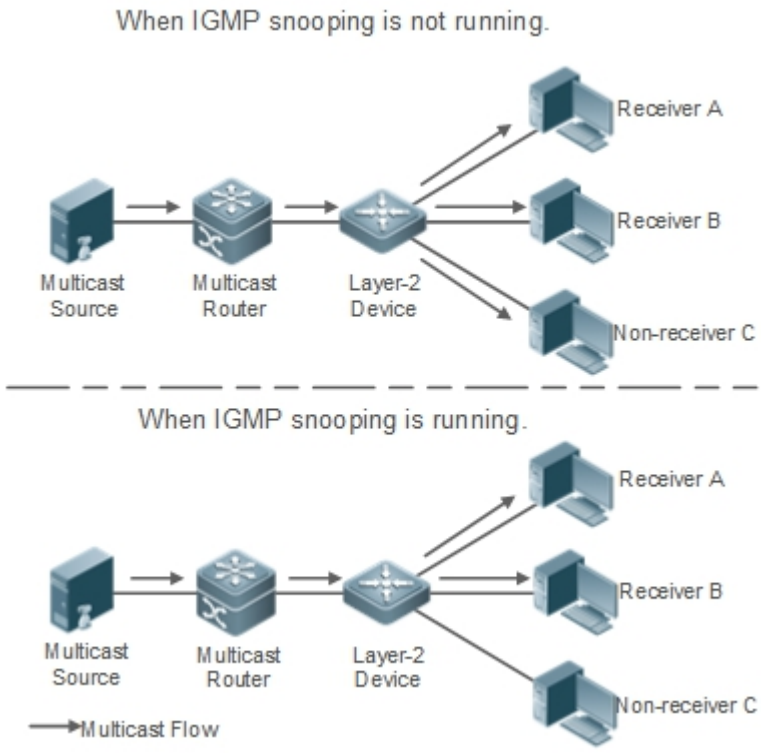
Application	Description
Layer-2 Multicast Control	Enables precise forwarding of Layer-2 multicast packets to avoid flooding at this layer.
Shared Multicast Services (Multicast VLAN)	Multiple users can share the multicast traffic of the same VLAN.
Premium Channels and Preview	Controls the range of multicast addresses that allow user demanding and allows preview for profiles who are inhibited from demanding.

8.2.1 Layer-2 Multicast Control

Scenario

As shown in the following figure, multicast packets are transmitted to users through a Layer-2 switch. When Layer-2 multicast control is not performed, namely, when IGMP snooping is not implemented, multicast packets are flooded to all the users including those who are not expected to receive these packets. After IGMP snooping is implemented, the multicast packets from an IP multicast profile will no longer be broadcast within the VLAN but transmitted to designated receivers.

Figure 2-2 Networking Topology of Implementing Layer-2 Multicast Control (Multicast VLAN)



Deployment

- Configure basic IGMP snooping functions.

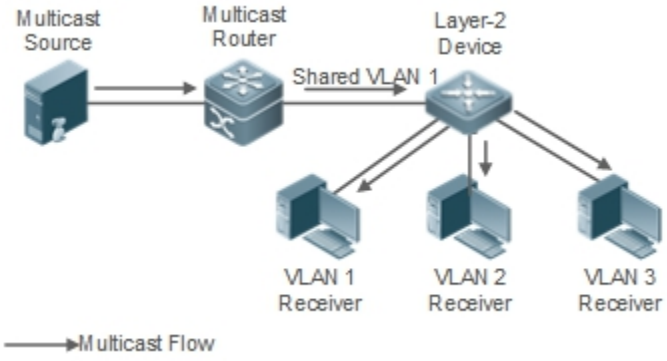
8.2.2 Shared Multicast Services (Multicast VLAN)

Scenario

In Shared VLAN Group Learning (SVGL) mode or IVGL-SVGL mode (IVGL: Independent VLAN Group Learning), a device running IGMP snooping can provide shared multicast services (or multicast VLAN services) to the VLAN users. Typically, this function is used to provide the same video-on-demand (VOD) services to multiple VLAN users.

The following figure shows the operation of a Layer-2 multicast device in SVGL mode of IGMP snooping. The multicast router sends a multicast packet to VLAN 1, and the Layer-2 multicast device automatically transfers the packet to VLAN 1, VLAN 2, and VLAN 3. In this way, the multicast services of VLAN 1 are shared by VLAN 2 and VLAN 3.

Figure 2-3 Networking Topology of Shared Multicast Services (Multicast VLAN)



If the Layer-2 multicast device operates in IVGL mode, the router must send a packet to each VLAN, which wastes bandwidth and burdens the Layer-2 multicast device.

Deployment

- Configure basic IGMP snooping functions (in SVGL mode or IVGL-SVG mode).

8.2.3 Premium Channels and Preview

Scenario

In VOD application, by limiting the range of the multicast addresses that a user host can access, unpaid users will not be able to watch the premium channels. Thereafter, the preview service is offered to unpaid users before they decide whether to pay for it.

The users can preview a premium channel for a certain period of time (for example 1 minute) after demanding it.

Deployment

- Configure basic IGMP snooping functions (in any working mode).
- Configure the range of multicast addresses that a user can access.
- Enable the preview function for VOD profiles that are denied access.

8.3 Features

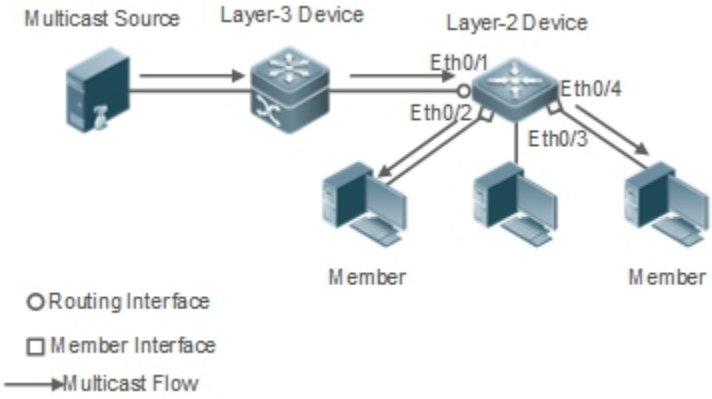
Basic Concepts

Multicast Router Ports and Member Ports

IGMP snooping is VLAN-based. The ports involved refer to the member ports within the VLAN.

The device running IGMP snooping identifies the ports within the VLAN as a multicast router port or member port so as to manage and control the forwarding of IP multicast traffic within the VLAN. As shown in the following figure, when IGMP snooping is run on a Layer-2 device, multicast traffic enters the multicast router port and exits from the member ports.

Figure 2-4 Networking Topology of Two IGMP Snooping Ports



- Multicast router port: The location of the multicast source is directed by the port on the Layer-2 multicast device which is connected to the multicast router (Layer-3 multicast device): By listening to IGMP packets, the Layer-2 multicast device can automatically detect the multicast router port and maintain the port dynamically. It also allows users to configure a static router port.
- Member port: The port is on a Layer-2 multicast device and is connected to member hosts. It directs the profile members. It is also called the Listener Port. By listening to IGMP packets, the Layer-2 multicast device can automatically detect the member port and maintain the port dynamically. It also allows users to configure a static member port.

Overview

Feature	Description
Listening to IGMP Packets	Discovers and identifies the router port and member port to establish and maintain the IGMP snooping forwarding entries. :
IGMP Snooping Working Modes	Provides independent or shared multicast services to the user VLAN.
Multicast Security Control	Controls the multicast service scope and load to prevent illegal multicast traffic.
Profile	Defines the range of multicast addresses that permit or deny user requests for reference of other functions.
Handling QinQ	Sets the forwarding mode of multicast packets on the QinQ interface.
IGMP Querier	On a network without a Layer-3 multicast device, the Layer-2 multicast device acts as an IGMP querier.

8.3.1 Listening to IGMP Packets

A device running IGMP snooping analyzes IGMP packets received, and finds and identifies the router port and member port using these packets, thereby creating and maintaining an IGMP snooping entry.

Working Principle

A device running IGMP snooping can identify and handle the following types of IGMP packets:

Query Packets

An IGMP querier periodically sends General Query packets. When the IGMP querier receives Leave packets, it sends Group-Specific Query packets.

When the device running IGMP snooping receives the Query packets, it performs the following operations within the VLAN:

- Forward the IGMP Query packets to all the ports (except the receiving port of these packets).
- If the receiving port is a dynamic router port, reset the aging timer. If the timer expires, the port will no longer be used as the dynamic router port.
- If the receiving port is not a dynamic router port, use it as a dynamic router port and enable the aging timer. If the timer expires, the port will no longer be used as the dynamic router port.
- For general queries, reset the aging timer for all the dynamic member ports. If the timer expires, the port will no longer be used as the dynamic member port for the general group. By default, the maximum response time carried by the IGMP query packets is used as the timeout time of the aging timer. If **ip igmp snooping query-max-response-time** is run, the time displayed is used as the timeout time of the aging timer.
- For designated query packets, reset the aging timer for all the dynamic member ports of the designated profile. If the timer expires, the port will no longer be used as the dynamic member port of the designated profile. By default, the maximum response time carried by the IGMP query packets is used as the timeout time of the aging timer. If **ip igmp snooping query-max-response-time** is run, the time displayed is used as the timeout time of the aging timer.
- If dynamic router port learning is disabled, IGMP snooping will not learn the dynamic router port.

Report Packets

When a member host receives a query, it responds to the query with a Report packet. If a host requests to join a profile, it will also send a report.

By default, IGMP Snooping is capable of processing IGMPv1 and IGMPv2 packets. For IGMPv3 Report packets, it processes profile information but does not process carried source information. IGMP Snooping v3 can be configured to process all information in IGMPv1, IGMPv2, and IGMPv3 packets.

When the device running IGMP snooping receives the Report packets, it performs the following operations within the VLAN:

- Forward the Report packets from all the router ports. After the **ip igmp snooping suppression enable** command is run in one IGMP query cycle, only the first report received by each profile will be forwarded.
- If the port on which Report packets are received is a dynamic member port, reset the aging timer. If the timer expires, the port will no longer be used as the dynamic member port of the designated profile.
- If the port on which Report packets are received is not a dynamic member port, use it as a dynamic member port and enable the aging timer. If the timer expires, the port will no longer be used as the dynamic member port of the designated profile.

Leave Packets

If a host requests to leave a profile, it will send a Leave packet.

When the device running IGMP snooping receives the Leave packets, it performs the following operations within the VLAN:

- Forward the leave packets from all the router ports.
- If the port on which leave packets are received is a dynamic member port and the Leave function is enabled, the port will be immediately deleted from the IGMP snooping forwarding entry of the designated profile and will no longer be used as the dynamic member port.
- If the port on which the leave packets are received is a dynamic member port and the Leave function is disabled, the port state should be maintained.

Related Configuration

Configuring a Static Router Port

Run the **ip igmp snooping vlan mrouter interface** command to configure a static router port.

Configuring a Static Member Port

Run the **ip igmp snooping vlan static interface** command to configure a static member port.

Enabling Report Suppression

Report suppression is disabled by default.

Run the **ip igmp snooping suppression enable** command to enable report suppression.

After report suppression is enabled, in one IGMP query cycle, only the first Report packet received by each profile will be forwarded. The source media access control (MAC) address of the forwarded report will be changed to the MAC address of the device.

Enabling Immediate Leave

Immediate leave is disabled by default.

Run the **ip igmp snooping fast-leave enable** command to enable immediate leave.

Enabling Dynamic Router Port Learning

Dynamic router port learning is enabled by default.

Run the **no ip igmp snooping mrouter learn pim-dvmrp** command to disable dynamic router port learning.

Run the **no ip igmp snooping vlan** *vid* **mrouter learn pim-dvmrp** command to disable dynamic router port learning for designated VLANs.

Configuring the Aging Time of a Dynamic Router Port

The default aging time is 300s.

When a dynamic router port receives a query packet, the aging timer of the port is enabled or reset; if the aging time is not configured, the maximum response time carried by the query packet is used as the aging time.

Run **ip igmp snooping dyn-mr-aging-time** to configure the aging time of the dynamic router port.

Configuring the Aging Time of a Dynamic Member Port

The default aging time is 260s.

When a dynamic member port receives a query packet, the aging timer of the port is enabled or reset, and the aging time is the maximum response time carried by the query packet.

When a dynamic member port receives a Report packet, the aging timer of the port is enabled or reset, and the aging time is the maximum response time of the dynamic member port.

Run **ip igmp snooping host-aging-time** to configure the aging time of the dynamic member port.

Configuring the Maximum Response Time of a Query Packet

The maximum response time of a query packet is not configured by default and the maximum response time carries by the query packet is used.

Run **ip igmp snooping query-max-response-time** to configure the maximum response time of a query packet.

8.3.2 IGMP Snooping Working Modes

A device running in the three modes (IVGL, SVGL, and IVGL-SVGL) of IGMP snooping can provide independent multicast services or shared multicast services to the user VLAN.

Working Principle

IVGL

In IVGL mode, a device running IGMP snooping can provide independent multicast services to each user VLAN.

Independent multicast services indicate that multicast traffic can be forwarded only within the VLAN it belongs to, and a user host can subscribe to the multicast traffic within the VLAN that the host belongs to.

SVGL

In SVGL mode, a device running IGMP snooping can provide shared multicast services to the user VLAN.

Shared multicast services can be provided only on shared VLANs and sub VLANs and SVGL multicast addresses are used. In a shared VLAN, the multicast traffic within the range of SVGL multicast addresses is forwarded to a sub VLAN, and the user hosts within the sub VLAN subscribe to such multicast traffic from the shared VLAN.

- In a shared VLAN and sub VLAN, shared multicast services will be provided to the multicast traffic within the range of SVGL multicast addresses. Other multicast traffic will be discarded.
- Other VLANs (except shared VLANs and sub VLANs) apply to independent multicast services.

When the user VLAN is set to a shared VLAN or sub VLAN, shared multicast services are provided; when a user VLAN is set to other VLANs, independent multicast services are provided.

IVGL-SVGL

IVGL-SVGL mode is also called the hybrid mode. In this mode, a device running IGMP snooping can provide both shared and independent multicast services to the user VLAN.

- In a shared VLAN and sub VLAN, multicast services will be provided to the multicast traffic within an SVGL profile. For other multicast traffic, independent multicast services will be provided.
- Other VLANs (except shared VLANs and sub VLANs) apply to independent multicast services.

When a user VLAN is configured as a shared VLAN or sub VLAN, both public multicast services and independent multicast services available. When a user VLAN is configured as a VLAN other than shared VLAN and sub VLAN, only the independent multicast services are available.

Related Configuration

Enabling IGMP Snooping and Selecting a Working Mode

IGMP snooping is disabled by default.

Run the **ip igmp snooping ivgl** command to enable IGMP snooping in IVGL mode.

Run the **ip igmp snooping svgl** command to enable IGMP snooping in SVGL mode.

Run the **ip igmp snooping ivgl-svgl** command to enable IGMP snooping in IVGL-SVGL mode.

A working mode must be designated when enabling IGMP snooping, namely, one of the preceding working modes must be selected.

Configuring Shared VLAN

The shared VLAN is VLAN 1 by default.

Run the **ip igmp snooping svgl vlan** command to designate a VLAN as the shared VLAN.

In SVGL mode and IVGL-SVGL mode, only one VLAN can be configured as the shared VLAN.

Configuring Sub VLAN

By default, a sub VLAN is any VLAN except the shared VLAN.

Run the **ip igmp snooping svgl subvlan** command to designate a VLAN as the sub VLAN.

In SVGL mode and IVGL-SVGL mode, the number of sub VLANs is not limited.

Configuring an SVGL Profile

No default setting.

Run the **ip igmp snooping svgl profile** *profile_num* command to configure the address range of an SVGL profile.

In SVGL mode and IVGL-SVGL mode, the SVGL profile range must be configured; otherwise, shared multicast services cannot be provided.

8.3.3 IGMP Security Control

A device running IGMP snooping can control the multicast service scope and load, and effectively prevents illegal multicast traffic.

Working Principle

Configuring the Profile Filtering for User Demanding

By configuring the profile list that a user can access, you can customize the multicast service scope to guarantee the interest of operators and prevent illegal multicast traffic.

To enable this function, you should use a profile to define the range of multicast addresses that a use is allowed to access.

- When the profile is applied on a VLAN, you can define the multicast addresses that a user is allowed to access within the VLAN.
- When the profile is applied on an interface, you can define the multicast addresses that a user is allowed to access under the port.

Multicast Preview

If the service provider wants to allow the users to preview some multicast video traffic that denies the users' access, and stop the multicast video traffic after the preview duration is reached, the user-based multicast preview function should be provided.

The multicast preview function is used together with multicast permission control. For example, in the application of videos, the administrator controls some premium channels by running the **ip igmp profile** command on a port or VLAN. In this way, unsubscribed users will not be able to watch these channels on demand. If users want to preview the channels before they decide whether to pay for watching or not, the multicast preview function can be enabled, allowing the premium channels to be previewed by unpaid users for a certain period of time (for example 1 minute).

Controlling the Maximum Number of Profiles Allowed for Concurrent Request

If there is too much multicast traffic requested at the same time, the device will be severely burdened. Configuring the maximum number of profiles allowed for concurrent request can guarantee the bandwidth.

- You can limit the number of profiles allowed for concurrent request globally.
- You can also limit the number of profiles allowed for concurrent request on a port.

Controlling the Entry of Multicast Traffic

By running the **ip igmp snooping source-check port** command to enable source port inspection, you can restrict the entry of multicast traffic to prevent illegal traffic.

- When source port inspection is enabled, only the multicast traffic entered from the router port is considered as legal; the traffic from other ports is considered as illegal and will be discarded.

- When source port inspection is disabled, the traffic entered from any port is considered as legal.

Configuring the Source IP Inspection for Multicast Traffic

By enabling source IP inspection, you can restrict the IP address of multicast traffic to prevent illegal traffic.

Source IP inspection includes the inspection of the source IP addresses of specific profiles and of default profiles.

- Inspection of the source IP addresses of default profiles (also called source-check default-server): Specifies the source IP addresses for all the multicast profiles within all VLANs. Only the multicast traffic whose source IP address is the same as the set one is considered as legal.
- Inspection of the source IP addresses of specific profiles (also called limit-ipmc): Specifies the source IP addresses for specific multicast profiles within specific VLANs. Among the multicast traffic received from the specific multicast profiles within the VLANs, only the one with the same source IP address as the set one is considered as legal and will be forwarded by the multicast device; other traffic will be discarded.

Related Configuration

Configuring the Profile Filtering

By default, profiles are not filtered and allow user access.

To filter multicast profiles, run the **ip igmp snooping filter** command in interface configuration mode or global configuration mode.

Enabling Preview

Preview is not enabled by default.

Run the **ip igmp snooping preview** command to enable preview and restrict the range of the profiles permitted for multicast preview.

Run the **ip igmp snooping preview interval** to set the multicast preview duration.

Configuring the Maximum Number of Profiles Allowed for Concurrent Request on a Port

By default, the number of profiles allowed for concurrent request is not limited.

Run the **ip igmp snooping max-groups** command to configure the maximum number of profiles allowed for concurrent request.

Configuring the Maximum Number of Multicast Profiles Allowed Globally

By default, the maximum number of multicast profiles allowed globally is 65,536.

Run the **ip igmp snooping l2-entry-limit** command to configure the maximum number of multicast profiles allowed globally.

Enabling Source Port Inspection

By default, source port inspection is not configured.

Run the **ip igmp snooping source-check port** command to enable source port inspection.

Enabling Source IP Inspection

By default, source IP inspection is disabled.

- Run the **ip igmp snooping source-check default-server** *address* command to enable source IP inspection and specify the default source IP address (applicable to any profile of any VLAN).
- (Optional) Run the **ip igmp snooping limit-ipmc vlan** *vid* **address** *group-address* **server** *source-address* command to specify a specific source IP address for a specific profile of specific VLAN (applicable to a specific profile of specific VLAN).

First, you must enable source IP inspection to specify default source address, and then a specific source address can be specified for a specific profile of specific VLAN. If a source address is specified for a specific profile of specific VLAN, the multicast traffic of the specific profile will perform inspection for the source address specified by this command. Other multicast traffic will perform inspection for default source addresses.

8.3.4 IGMP Profile

A multicast profile is used to define the range of multicast addresses that permit or deny user demanding request for reference of other functions.

Working Principle

The profile is used to define the range of multicast addresses.

When SVGL mode is enabled, an SVGL profile is used to define the range of SVGL multicast addresses.

When the multicast filter is configured on an interface, a profile is used to define the range of multicast addresses that permit or deny user request under the interface.

When a VLAN filter is configured, a profile is used to define the range of multicast addresses that permit or deny user request under within the VLAN.

When the preview function is enabled, a profile is used to define the range of multicast address allowed for preview.

Related Configuration

Configuring a Profile

Default configuration:

- Create a profile, which is **deny** by default.

Configuration steps:

- Run the **ip igmp profile** *profile-number* command to create a profile.
- Run the **range** *low-address high_address* command to define the range of multicast addresses. Multiple address ranges are configured for each profile.
- (Optional) Run the **permit** or **deny** command to permit or deny user request (**deny** by default). Only one **permit** or **deny** command can be configured for each profile.

8.3.5 IGMP QinQ

Working Principle

On a device with IGMP snooping enabled and dot1q-tunnel (QinQ) port configured, IGMP snooping will handle the IGMP packets received by the QinQ port using the following two approaches:

- Approach 1: Create a multicast entry on the VLAN where IGMP packets are located. The forwarding of IGMP packets on the VLAN where these packets are located is called transparent transmission. For example, presume that IGMP snooping is enabled for a device, Port A is designated as the QinQ port, the default VLAN of this port is VLAN 1, and it allows the passage of VLAN 1 and VLAN 10 packets. When a multicast Query packet is sent by VLAN 10 to Port A, IGMP snooping establishes a multicast entry for VLAN 10 and forwards the multicast Query packet to the router port of VLAN 10.
- Approach 2: Create a multicast entry on the default VLAN of the QinQ port. Encapsulate the multicast packet with the VLAN tag of the default VLAN where the QinQ port is located and forward the packet within the default VLAN. For example, presume that IGMP snooping is enabled for a device, Port A is designated as the QinQ port, the default VLAN of this port is VLAN 1, and it allows the passage of VLAN 1 and VLAN 10 packets. When a multicast Query packet is sent by VLAN 10 to Port A, IGMP snooping establishes a multicast entry for VLAN 1, encapsulates the multicast query packet with the tag of VLAN 1, and forward the packet to VLAN 1 router port.

Related Configuration

Configuring QinQ

By default, IGMP snooping works in the mode specified in Approach 2.

Run the **ip igmp snooping tunnel** command to implement Approach 1.

8.3.6 IGMP Querier

On a network with a Layer-3 multicast device, the Layer-3 multicast device acts as an IGMP querier. In this case, a Layer-2 device needs only to listen to IGMP packets to establish and maintain the forwarding entry, realizing Layer-2 multicast.

On a network without a Layer-3 multicast device, the Layer-2 multicast device must be configured with the IGMP querier function so that the device can listen to IGMP packets. In this case, a Layer-2 device needs to act as an IGMP querier as well as listen to IGMP packets to establish and maintain the forwarding entry to realize Layer-2 multicast.

Working Principle

A Layer-2 device acts as an IGMP querier to periodically send IGMP Query packets, listen to and maintain the IGMP Report packets replied by a user, and create a Layer-2 multicast forwarding entry. You can adjust relevant parameters of the Query packets sent by the IGMP querier through configuration.

When the device receives a Protocol-Independent Multicast (PIM) or Distance Vector Multicast Routing Protocol (DVMRP) packet, it considers that a multicast router, which will act as an IGMP querier, exists on the network and disables the querier function. In this way, IGMP routing will not be affected.

When the device receives the IGMP Query packets from other devices, it will compete with other devices for the IGMP querier.

Enabling the Querier Function

You can enable the querier for a specific VLAN or all VLANs.

Only when the global querier function is enabled can the queriers for specific VLANs take effect.

Specifying the IGMP Version for a Querier

The version of IGMP used for sending Query packets can be configured as IGMPv1, IGMPv2, or IGMPv3.

Configuring the Source IP Address of a Querier

You can configure the source IP address of a query packet sent by the querier based on VLANs.

When the source IP address of the querier is not configured, the querier will not take effect.

Configuring the Query Interval of a Querier

You can configure the intervals for sending global Query packets based on different queriers on different VLANs.

Configuring the Maximum Response Time of a Query Packet

You can configure the maximum response time carried by a Query packet that is sent by a querier. As IGMPv1 does not support the carrying of maximum response time by a Query packet, this configuration does not take effect when the querier is running IGMPv1. You can configure different maximum response time for queriers on different VLANs.

Configuring the Aging Time of a Querier

When other IGMP queriers exist on a network, the existing device will compete with other queriers. If the existing device fails to be elected and is in the non-querier state, the aging timer of a querier will be enabled. After the timer expires, other queriers on the network are considered as expired and the existing device will be resumed as the querier.

Related Configuration

Enabling the Querier Function

By default, the querier function of a device is disabled.

Run the **ip igmp snooping querier** command to enable the global querier function.

Run the **ip igmp snooping vlan** *num* **querier** command to enable the querier function for specific VLANs.

Specifying the IGMP Version for a Querier

By default, a querier runs IGMPv2.

Run the **ip igmp snooping querier version** command to configure the global querier version.

Run the **ip igmp snooping vlan querier version** command to specify the querier version for specific VLANs.

Configuring the Source IP Address of a Querier

By default, the source IP address of a querier is 0.

Run the ip igmp snooping querier address command to enable global source IP addresses of queriers.

Run the ip igmp snooping vlan querier address command to specify the source IP addresses of the queriers on specific VLANs.

Configuring the Query Interval of a Querier

By default, the query interval of a querier is 60s.

Run the **ip igmp snooping querier query-interval** command to enable the global query interval of queriers.

Run **ip igmp snooping vlan querier query-interval** to specify the global query interval of the queriers on specific VLANs.

Configuring the Maximum Response Time of a Query Packet

By default, the maximum response time of a query packet is 10s.

Run the **ip igmp snooping querier max-response-time** command to configure the maximum response time of the query packets sent by global queriers.

Run the **ip igmp snooping vlan querier max-response-time** command to specify the maximum response time of the query packets sent by the queriers on specific VLANs.

Configuring the Aging Time of a Querier

By default, the aging time of a querier is 125s.

Run the **ip igmp snooping querier max-response-time** command to configure the aging time of global queriers.

Run the **ip igmp snooping vlan querier max-response-time** command to configure the aging time of queriers on specific VLANs.

8.4 Configuration

Configuration	Description and Command	
Configuring Basic IGMP Snooping Functions (IVGL Mode)	Any of IVGL mode, SVGL mode, and IVGL-SVGL mode must be selected.	
	ip igmp snooping ivgl	Enables global IGMP snooping in IVGL mode.
	no ip igmp snooping vlan *num*	Disables IGMP snooping for a VLAN.
Configuring Basic IGMP Snooping Functions (SVGL Mode)	Any of IVGL mode, SVGL mode, and IVGL-SVGL mode must be selected. It is used to enable IGMP snooping in SVGL mode.	
	ip igmp snooping svgl	Enables global IGMP snooping in IVGL mode.
	no ip igmp snooping vlan *num*	Disables IGMP snooping for a VLAN.
	ip igmp snooping svgl profile *profile_num*	Configures the SVGL profile.
	ip igmp snooping svgl vlan	Specifies the SVGL shared VLAN.
	ip igmp snooping svgl subvlan	Specifies the SVGL sub VLAN.
Configuring Basic IGMP Snooping Functions (IVGL-SVGL Mode)	Any of IVGL mode, SVGL mode, and IVGL-SVGL mode must be selected. It is used to enable IGMP snooping in IVGL-SVGL mode.	
	ip igmp snooping ivgl-svgl	Enables global IGMP snooping in IVGL-SVGL mode.
	no ip igmp snooping vlan *num*	Disables IGMP snooping for a VLAN.
	ip igmp snooping svgl profile *profile_num*	Configures the SVGL profile.
	ip igmp snooping svgl vlan	Specifies the SVGL shared VLAN.
	ip igmp snooping svgl subvlan	Specifies the SVGL sub VLAN.
Configuring the Packet Processing	(Optional) It is used to adjust relevant configurations for processing protocol packets.	
	ip igmp snooping vlan *vlan-id* **mrouter interface** *interface-id*	Configures a static router port.
	p igmp snooping vlan *vid* **static** *group-address* **interface** *interface-type interface-number*	Configures a static member port.
	ip igmp snooping vlan *vlan-id* **mrouter learn pim-dvmrp**	Enables dynamic router port learning.
	ip igmp snooping dyn-mr-aging-time *time*	Configures the aging time of a dynamic router port.

	ip igmp snooping host-aging-time *time*	Configures the aging time of a dynamic member port.
	ip igmp snooping fast-leave enable	Enables the immediate-leave function for a dynamic member port.
	ip igmp snooping query-max-response-time *time*	Configures the maximum response time of an IGMP query packet.
	ip igmp snooping suppression enable	Enables IGMP Report packet suppression.
Configuring IGMP Security Control	(Optional) It used to guarantee the security when a user requests a multicast profile.	
	ip igmp snooping filter *profile-number*	Configures the profile filtering for user access.
	ip igmp snooping vlan *num* **filter** *profile-number*	Configures the per-VLAN profile filtering for user access.
	ip igmp snooping l2-entry-limit *number*	Configures the maximum number of profiles globally for user access.
	ip igmp snooping max-groups *number*	Configures the maximum number of dynamic profiles for user access.
	ip igmp snooping source-check port	Enables source IP inspection, which ensures the multicast traffic from the router port is legal.
	ip igmp snooping source-check default-server *address*	Enables source IP inspection. The multicast traffic whose source IP address matches the specified source IP address is considered as legal traffic.
	ip igmp snooping limit-ipmc vlan *vid* **address** *group-address* **server** *source-address*	Specifies a VLAN. In the multicast traffic of multicast addresses, the one whose source IP address matches the specified source IP address is considered as legal traffic.
	ip igmp snooping preview *profile-number*	Enables the preview function for a specified profile.
	ip igmp snooping preview interval *num*	Configures the preview duration.
Configuring an IGMP Profile	(Optional) It is used to define the range of multicast addresses that permits or denies the access of a user host.	
	ip igmp profile *profile-number*	Creates a profile.
	range *low-address high_address*	Configures the profile range.
	permit	Permits the access of a user host.
	deny	Denies the access of a user host.
Configuring IGMP QinQ	(Optional) It is used to configure QinQ interface to forward multicast packets using the VLAN identifier (VID) carried by packets.	
	ip igmp snooping tunnel	Configures QinQ to transmit IGMP packets transparently.
Configuring an IGMP Querier	(Optional) It is used to enable IGMP querier function on a network without a Layer-3 multicast device.	

	ip igmp snooping querier	Enables global querier function.
	ip igmp snooping vlan *num* **querier**	Enables the querier for a VLAN.
	ip igmp snooping querier version *num*	Specifies the IGMP version for queriers globally.
	ip igmp snooping vlan *num* **querier version** *num*	Specifies the IGMP version for a querier of a VLAN.
	ip igmp snooping querier address a.b.c.d	Configures the source IP address of queriers globally.
	ip igmp snooping vlan *num* **querier address** a.b.c.d	Configures the source IP address for a querier of a VLAN.
	ip igmp snooping querier query-interval *num*	Configures the query interval of queriers globally.
	ip igmp snooping vlan *num* **querier query-interval** *num*	Configures the query interval for a querier of a VLAN.
	ip igmp snooping querier max-response-time *num*	Configures the maximum response time for query packets globally.
	ip igmp snooping vlan *num* **querier max-response-time** *num*	Configures the maximum response time of query packets for a VLAN.
	ip igmp snooping querier timer expiry *num*	Configures the aging timer for queriers globally.
	ip igmp snooping vlan *num* **querier timer expiry** *num*	Configures the aging timer for a querier of a VLAN.

8.4.1 Configuring Basic IGMP Snooping Functions (IVGL Mode)

Configuration Effect

- Enable IGMP snooping to realize Layer-2 multicast.
- Provide independent multicast services to each VLAN.

Notes

- IP multicast cannot be realized in SVGL mode. If IP multicast must be used, select the IVGL mode.

Configuration Steps

Enabling Global IGMP Snooping in IVGL Mode

Mandatory.

After IGMP snooping is enabled globally, this function will be enabled for all VLANs.

If not specified, it is advised to run global IGMP snooping on all the devices connected user hosts.

Disabling IGMP Snooping for a VLAN

(Optional) You can use this function if you wish to disable IGMP snooping on specified VLANs.

Only when global IGMP snooping is enabled can it be disabled on specified VLANs.

In IVGL mode, each VLAN can enjoy independent multicast services. Disabling any VLAN multicast services will not interfere in the services provided to the others.

Verification

- Run the **show ip igmp snooping gda-table** command to display the IGMP snooping forwarding table and verify that the member ports include only those connecting member hosts.

- Run the **show ip igmp snooping** command to display the basic IGMP snooping information and verify that IGMP snooping is working in IVGL mode.

Related Commands

Enabling Global IGMP Snooping in IVGL Mode

Command	**ip igmp snooping ivgl**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After this command is executed, IGMP snooping will be run on all VLANs. By default, IGMP snooping is disabled.

Disabling IGMP Snooping for a VLAN

Command	**no ip igmp snooping vlan** *num*
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Only when global IGMP snooping is enabled can it be disabled on specified VLANs. In IVGL mode, you can disable IGMP snooping on any VLAN.

Displaying the IGMP Snooping Entry

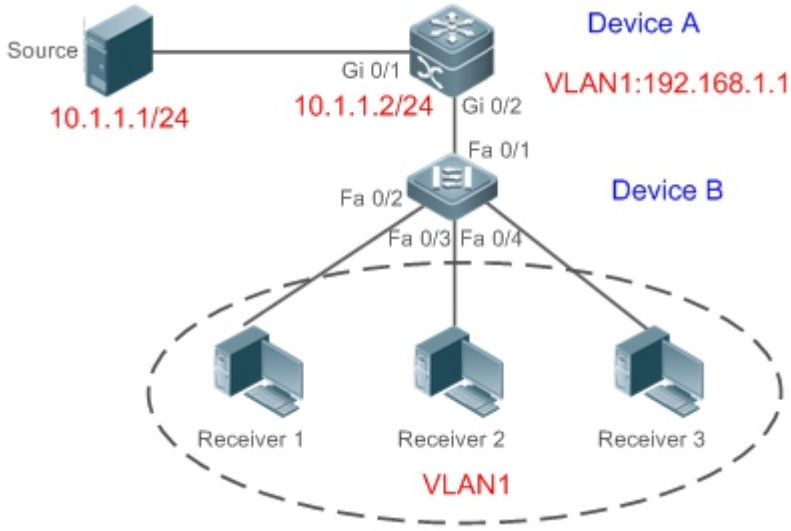
Command	**show ip igmp snooping gda-table**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	This command is used to verify that the ports include only those connecting member hosts.

Displaying the IGMP Snooping Working Mode

Command	**show ip igmp snooping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If a device is running in IVGL mode, the following information is displayed: IGMP Snooping running mode: IVGL

Configuration Example

Providing Layer-2 Multicast Services for the Subnet Hosts

Scenario **Figure 2- 5**	
	A is the multicast router and is connected directly to the multicast source. B is the Layer-2 device and is connected directly to the user host. Receiver 1, Receiver 2, and Receiver 3 belong to VLAN 1.
Configuration Steps	● Configure the IP address and VLAN. ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select IVGL mode.
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit
B	B# configure terminal B(config)# ip igmp snooping ivgl
Verification	Send packets from the source (10.1.1.1) to G (229.1.1.1) to add Receiver 1 to G. ● Confirm that the packets (10.1.1.1 and 229.1.1.1) are received by Receiver 1. ● Display the IGMP snooping forwarding entry on B and ensure that the port (10.1.1.1, 229.1.1.1, 1) includes only Fa0/2. ● Check whether the IGMP snooping working mode is IVGL.
B	B# show ip igmp snooping gda-table Multicast Switching Cache Table D: DYNAMIC S: STATIC M: MROUTE (*,224.1.1.1, 1): VLAN(1) 2 OPORTS:

```
    FastEthernet 0/1(M)
    FastEthernet 0/2(D)

B# show ip igmp snooping
IGMP Snooping running mode: IVGL
IGMP Snooping L2-entry-limit: 65536
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Disable
IGMP Report suppress: Disable
IGMP Global Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)

vlan 1
-------------
IGMP Snooping state: Enable
Multicast router learning mode: pim-dvmrp
IGMP Fast-Leave: Disabled
IGMP VLAN querier: Disable
IGMP VLAN Mode: STATIC
```

Common Errors

- The working mode of IGMP snooping is improper.

8.4.2 Configuring Basic IGMP Snooping Functions (SVGL Mode)

Configuration Effect

- Enable IGMP snooping and select SVGL mode to realize Layer-2 multicast.
- Share the VLAN multicast services.

Configuration Steps

Enabling Global IGMP Snooping in SVGL Mode

Mandatory.

Enable global IGMP snooping in SVGL mode.

Configure the range of associated SVGL profiles.

Specifying the SVGL Shared VLAN

(Optional) By default, VLAN 1 is used as the shared VLAN. You can adjust this configuration for other options.

Specifying the SVGL Sub VLAN

(Optional) By default, all the VLANs are used as the sub VLANs of SVGL and can share the multicast services of the shared VLAN. You can adjust this configuration for other options.

Verification

- Run the show ip igmp snooping command to display the basic IGMP snooping information and verify that IGMP snooping is working in SVGL mode.
- Run the **show ip igmp snooping gda-table** command to check whether inter-VLAN multicast entries are properly formed.

Related Commands

Enabling Global IGMP Snooping in SVGL Mode

Command	**ip igmp snooping svgl**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	By default, IGMP snooping is disabled. After the SVGL mode is selected, the range of profiles within SVGL multicast addresses needs to be associated.

Configuring the SVGL profile

Command	**ip igmp snooping svgl profile** *profile_num*
Parameter Description	*profile_num:* Configures SVGL to associate a profile.
Command Mode	Global configuration mode
Usage Guide	By default, no profile is associated with SVGL.

Specifying the SVGL Shared VLAN

Command	**ip igmp snooping svgl vlan** *vid*
Parameter Description	*vid*: Indicates a VLAN.
Command Mode	Interface configuration mode
Usage Guide	By default, VLAN 1 is used as the shared VLAN.

Specifying the SVGL Sub VLAN

Command	**ip igmp snooping svgl subvlan** *vid-range*
Parameter Description	*vid-range*: Indicates VLAN ID or the range of VLAN IDs.
Command Mode	Interface configuration mode
Usage Guide	By default, all the VLANs except the shared VLAN are used as sub VLANs.

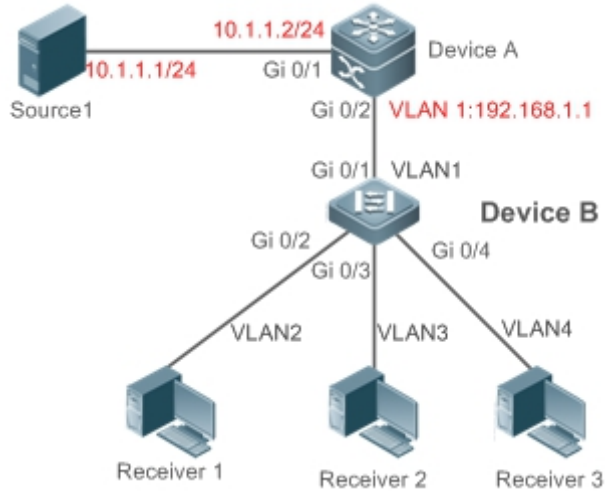
Displaying the IGMP Snooping Working Mode

Command	**show ip igmp snooping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If a device is running in SVGL mode, the following information is displayed:

	IGMP Snooping running mode: SVGL

Configuration Example

Enabling SVGL on the Access Device

Scenario **Figure 2- 6**	
	A is the multicast router and is connected directly to the multicast source. B is the Layer-2 device and is connected directly to the user host. Receiver 1 is connected to VLAN 2, Receiver 2 is connected to VLAN 3, and Receiver 3 is connected to VLAN 4.
Configuration Steps	● Configure the IP address and VLAN. (Omitted) ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select SVGL mode. ● Configure the range of associated SVGL multicast addresses on B.
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit
B	B# configure terminal B(config)#ip igmp profile 1 B(config-profile)#permit B(config-profile)#range 224.1.1.1 238.1.1.1 B(config-profile)#exit B(config)#ip igmp snooping svgl B(config)#ip igmp snooping svgl profile 1
Verification	Send packets from the source (10.1.1.1) to G (229.1.1.1) and add Receiver 1, Receiver 2 and Receiver 3 to G. ● Confirm that the packets (10.1.1.1 and 224.1.1.1) are received by Receiver 1, Receiver 2, and Receiver 3. ● Display the IGMP snooping forwarding entry on B and ensure that the ports (*, 224.1.1.1, 1) include Gi0/2, Gi0/3, and Gi0/4. ● Check whether the IGMP snooping working mode is SVGL.

B	B# show ip igmp snooping gda-table Multicast Switching Cache Table D: DYNAMIC S: STATIC M: MROUTE (*,224.1.1.1, 1): VLAN(2) 1 OPORTS: GigabitEthernet 0/2(D) VLAN(3) 1 OPORTS: GigabitEthernet 0/3(D) VLAN(4) 1 OPORTS: GigabitEthernet 0/4(D) B# show ip igmp snooping IGMP Snooping running mode: SVGL IGMP Snooping L2-entry-limit: 65536 SVGL vlan: 1 SVGL profile number: 1 Source port check: Disable Source ip check: Disable IGMP Fast-Leave: Disable IGMP Report suppress: Disable IGMP Globle Querier: Disable IGMP Preview: Disable IGMP Tunnel: Disable IGMP Preview group aging time : 60(Seconds) Dynamic Mroute Aging Time : 300(Seconds) Dynamic Host Aging Time : 260(Seconds)

Common Errors

- The SVGL profile is not configured.
- The sent multicast traffic is not within the SVGL profile.

8.4.3 Configuring Basic IGMP Snooping Functions (IVGL-SVGL Mode)

Configuration Effect

- Enable IGMP snooping and select IVGL-SVGL mode to realize Layer-2 multicast.
- The SVGL profiles can share the multicast services.
- The non-SVGL profiles run in IVGL mode.

Configuration Steps

Enabling Global IGMP Snooping in IVGL-SVGL Mode

Mandatory.

Enable global IGMP snooping in IVGL-SVGL mode.

Configure the range of associated SVGL profiles.

Specifying the SVGL Shared VLAN

(Optional) By default, VLAN 1 is used as the shared VLAN. You can adjust this configuration for other options.

Specifying the SVGL Sub VLAN

(Optional) By default, all the VLANs are used as the sub VLANs of SVGL and can share the multicast services of the shared VLAN. You can adjust this configuration for other options.

Verification

- Run the **show ip igmp snooping** command to display the basic IGMP snooping information and verify that IGMP snooping is working in IVGL-SVGL mode.
- Run the **show ip igmp snooping gda-table** command to check whether inter-VLAN multicast entries are properly formed for the SVGL profiles.
- Run the **show ip igmp snooping gda-table** command to check whether intra-VLAN multicast entries are properly formed for the SVGL profiles.

Related Commands

Enabling Global IGMP Snooping in IVGL-SVGL Mode

Command	**ip igmp snooping ivgl-svgl**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	By default, IGMP snooping is disabled. After the IVGL-SVGL mode is selected, the SVGL profiles needs to be associated.

Configuring the SVGL Profile

Command	**ip igmp snooping svgl profile** *profile_num*
Parameter Description	*profile_num:* Configures SVGL to associate a profile.
Command Mode	Global configuration mode
Usage Guide	By default, no profile is associated with SVGL.

Specifying the SVGL Shared VLAN

Command	**ip igmp snooping svgl vlan** *vid*
Parameter Description	*vid*: Indicates a VLAN.
Command Mode	Interface configuration mode
Usage Guide	By default, VLAN 1 is used as the shared VLAN.

Specifying the SVGL Sub VLAN

Command	**ip igmp snooping svgl subvlan** *vid-range*
Parameter Description	*vid-range*: Indicates VLAN ID or the range of VLAN IDs.
Command Mode	Interface configuration mode
Usage Guide	By default, all the VLANs except the shared VLAN are used as sub VLANs.

Displaying the IGMP Snooping Working Mode

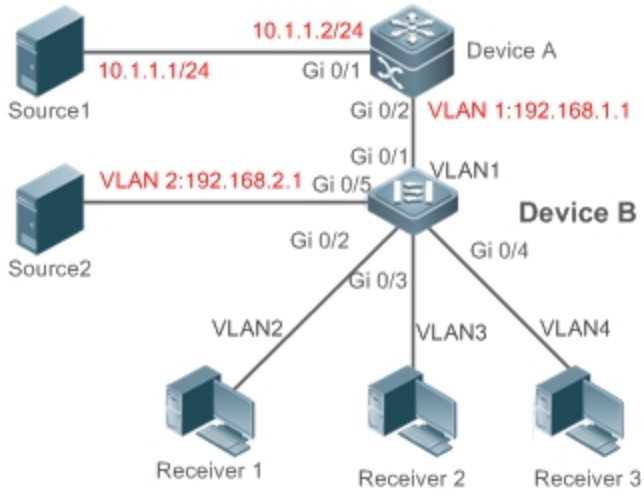
Command	**show ip igmp snooping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If a device is running in SVGL mode, the following information is displayed: IGMP Snooping running mode: SVGL

Displaying the IGMP Snooping Working Mode

Command	**show ip igmp snooping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If a device is running in IVGL-SVGL mode, the following information is displayed: IGMP Snooping running mode: IVGL-SVGL

Configuration Example

Enabling IVGL-SVGL on the Access Device

Scenario **Figure 2- 7**	
	A is the multicast router and is connected directly to multicast Source 1. B is a Layer-2 device and is connected directly to the user host and multicast Source 2. Receiver 1 is connected to VLAN 2, Receiver 2 is connected to VLAN 3, and Receiver 3 is connected to VLAN 4.
Configuration Steps	● Configure the IP address and VLAN. ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select IVGL-SVGL mode. ● Configure the range of associated SVGL multicast addresses on B.
A	A# configure terminal

	A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit
B	B# configure terminal B(config)#ip igmp profile 1 B(config-profile)#permit B(config-profile)#range 224.1.1.1 238.1.1.1 B(config-profile)#exit B(config)#ip igmp snooping ivgl-svgl B(config)#ip igmp snooping svgl profile 1
Verification	Send packets from Source 1 (10.1.1.1) to G (224.1.1.1) and add Receiver 1, Receiver 2 and Receiver 3 to G. Send packets from Source 2 (192.168.2.1) to the destination (239.1.1.1) and add Receiver 1 239.1.1.1. ● Confirm that the packets (10.1.1.1 and 224.1.1.1) are received by Receiver 1, Receiver 2, and Receiver 3. ● Check that packets (192.168.2.1 and 239.1.1.1) can be received by Receiver 1. ● Display the IGMP snooping forwarding entry on B and ensure that the ports (*, 224.1.1.1, 1) include Gi0/2, Gi0/3, and Gi0/4, and the port (*, 239.1.1.1, 1) is Gi0/2. ● Check whether the IGMP snooping working mode is IVGL-SVGL.
B	B# show ip igmp snooping gda-table Multicast Switching Cache Table D: DYNAMIC S: STATIC M: MROUTE (*,224.1.1.1, 1): VLAN(2) 1 OPORTS: GigabitEthernet 0/2(D) VLAN(3) 1 OPORTS: GigabitEthernet 0/3(D) VLAN(4) 1 OPORTS: GigabitEthernet 0/4(D) (*,239.1.1.1, 2): VLAN(2) 1 OPORTS: GigabitEthernet 0/2(D) B# show ip igmp snooping IGMP Snooping running mode: IVGL-SVGL IGMP Snooping L2-entry-limit: 65536 SVGL vlan: 1

	SVGL profile number: 0 Source port check: Disable Source ip check: Disable IGMP Fast-Leave: Disable IGMP Report suppress: Disable IGMP Globle Querier: Disable IGMP Preview: Disable IGMP Tunnel: Disable IGMP Preview group aging time : 60(Seconds) Dynamic Mroute Aging Time : 300(Seconds) Dynamic Host Aging Time : 260(Seconds)

Common Errors

- The SVGL profile is not configured.
- The sent multicast traffic is not within the SVGL profile.
- The IVGL multicast traffic cannot be forwarded within the SVGL profile.

8.4.5 Configuring the Packet Processing

Configuration Effect

- Configure specified ports as the static router ports to receive the multicast traffic from all profiles.
- Configure specified ports as the static member ports to receive the multicast traffic from specified profiles
- Enable Report packets suppression to forward only the first Report packet from a specified VLAN or profile to the router port within a query interval, and the following Report packets will not be forwarded to the router port, thereby reducing the quantity of packets on the network.
- Configure the immediate-leave function to delete a port from the entry of member ports when a leave packet is received by the port.
- Disable dynamic router port learning to disable the learning of any router port.
- Based on network load and configuration of a multicast device, you can adjust the aging time of a router port and member port as well as the maximum response time of a query packet.

Notes

- Only when basic IGMP snooping is configured can relevant configurations take effect.

Configuration Steps

↘ **Configuring a Static Router Port**

- Optional.
- You can perform this configuration if you want to specify a static port to receive all the multicast traffic within the VLAN.

↘ **Configuring a Static Member Port**

- Optional.
- You can perform this configuration if you want to specify a static port to receive specific multicast traffic within the VLAN.

↘ **Enabling Report Packet Suppression**

- Optional.
- When there are numerous receivers to receive the packets from the same multicast profile, you can enable Report packets suppression to suppress the number of Report packets to be sent.

↘ **Enabling the Immediate-Leave Function**

- Optional.
- When there is only one receiver on a port, you can enable Leave to speed up the convergence of protocol upon leave.

Disabling Dynamic Router Port Learning

- Optional.
- This function is used when multicast traffic needs to be forwarded only within the Layer-2 topology but not to a Layer-3 router.

Configuring the Aging Time of a Dynamic Router Port

- Optional.
- You can configure the aging time based on network load.

Configuring the Aging Time of a Dynamic Member Port

- Optional.
- You can configure the aging time based on the interval for sending IGMP query packets by the connected multicast router. Typically, the aging time is calculated as follows: Interval for sending IGMP query packets x 2 + Maximum response time of IGMP packets

Configuring the Maximum Response Time of a Query Packet

- Optional.
- You can configure the aging time based on network load.

Verification

- Run the **show ip igmp snooping mrouter** command to check whether the configured static router port has an "S" in the displayed configuration information.
- Run the **show ip igmp snooping gda** command to check whether the configured static member port is marked with an S.
- Run the **show ip igmp snooping** command to check whether Report packets suppression, immediate leave, router port learning, router port aging time, member port aging time, and the maximum response time of the Query packet take effect.

Related Commands

Configuring a Static Router Port

Command	**ip igmp snooping vlan** *vid* **mrouter interface** *interface-type interface-number*
Parameter Description	*vid*: Indicates a VLAN. The value ranges from 1 to 4,094. *interface-type interface-number*: Indicates an interface name.
Command Mode	Global configuration mode
Usage Guide	In SVGL mode, if a sub VLAN is not configured, only the configurations for the static router port within the shared VLAN can take effect, and the others can be configured but cannot take effect. If a sub VLAN is configured, only the configurations for the static router port within the shared VLAN or a non-sub VLAN can take effect, and the others can be configured but cannot take effect. In IVGL-SVGL mode, if a sub VLAN is not configured, the configurations for the static router ports within all the VLANs can take effect; if a sub VLAN is configured, only the configurations for the static router port within the shared VLAN or a non-sub VLAN can take effect, and the others can be configured but cannot take effect. In IVGL mode, the configurations for the static router ports within all the VLANs can take effect.

Configuring a Static Member Port

Command	**ip igmp snooping vlan** *vid* **static** *group-address* **interface** *interface-type interface-number*
Parameter Description	*vid*: Indicates a VLAN. The value ranges from 1 to 4,094. *group-address*: Indicates a profile address. *interface-type interface-number*: Indicates an interface name.
Command Mode	Global configuration mode

Usage Guide	By default, no static member port is configured.

Enabling Report Packet Suppression

Command	**ip igmp snooping suppression enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	When Report packets suppression is enabled, only the first Report packet from a specified VLAN or profile is forwarded to the router port within a Query interval, and the following Report packets will not be forwarded to the router port, thereby reducing the quantity of packets on the network. Only the IGMPv1 and IGMPv2 Report packets can be suppressed, and the IGMPv3 Report packets cannot be suppressed.

Enabling the Immediate-Leave Function

Command	**ip igmp snooping fast-leave enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	When this function is enabled, a port will be deleted from the entry of the member port when the port receives a leave packet. After that, the packets will no longer be forwarded to this port when it receives the query packets of specified profiles. Leave packets include the IGMPv2 Leave packets as well as the IGMPv3 Report packets that include types but carry no source address. The immediate-leave function applies only to the scenario where only one host is connected to a device port. It is used to conserve bandwidth and resources.

Enabling Dynamic Router Port Learning

Command	**ip igmp snooping** [**vlan** *vid*] **mrouter learn pim-dvmrp**
Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default.
Command Mode	Global configuration mode
Usage Guide	A router port is the port that is connected directly to a multicast device running IGMP snooping and a multicast neighbor device running multicast routing protocol. By default, dynamic router port learning is enabled and the device automatically listens to IGMP Query packets, DVMRP packets, and PIM Hello packets.

Configuring the Aging Time of a Dynamic Router Port

Command	**ip igmp snooping dyn-mr-aging-time** *seconds*
Parameter Description	*seconds*: Indicates the aging time of a dynamic router port in the unit of seconds. The value ranges from 1 to 3,600.
Command Mode	Global configuration mode
Usage Guide	If a dynamic router port does not receive an IGMP general query packet or a PIM Hello packet before the aging timer expires, the device will delete this port from the router port entry.

	When dynamic router port learning is enabled, you can run this command to adjust the aging time of the dynamic router port. If the aging time is too short, the multicast device may frequently add or delete a router port.

Configuring the Aging Time of a Dynamic Member Port

Command	**ip igmp snooping host-aging-time** *seconds*
Parameter Description	*seconds*: Indicates the aging time.
Command Mode	Global configuration mode
Usage Guide	The aging time of a dynamic member port indicates the time when a device port receives the IGMP join packet sent from host for subscribing to an IP multicast profile. When the IGMP join packet is received, the aging time of the dynamic member port will be reset. The value of the timer time is host-aging-time. If the timer expires, the multicast device deems that no user host for receiving the multicast packet exists under the port, and will delete the port from the entry of IGMP snooping member port. After the aging time is configured, the aging time of following received IGMP join packets will be host-aging-time. This configuration takes effect after the next IGMP join packet is received, and the timer of the port in use will not be refreshed.

Configuring the Maximum Response Time of a Query Packet

Command	**ip igmp snooping query-max-response-time** *seconds*
Parameter Description	*seconds*: Indicates the maximum response time.
Command Mode	Global configuration mode
Usage Guide	When an IGMP general Query packet is received, the multicast device will reset the aging time of all the dynamic member ports, which is query-max-response-time. If the timer expires, the multicast device deems that no user host for receiving the multicast packet exists under the port, and will delete the port from the entry of IGMP snooping member port. When an IGMP profile-specific Query packet is received, the multicast device will reset the aging time of all the dynamic member ports of the specific profile, which is query-max-response-time. If the timer expires, the multicast device deems that no user host for receiving the multicast packet exists under the port, and will delete the port from the entry of IGMP snooping member port. This configuration takes effect after the next Query packet is received, and the timer in use will not be refreshed. The timer of an IGMPv3 profile-specific Query packet is not refreshed.

Displaying Router Ports

Command	**show ip igmp snooping mroute**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode

Usage Guide	If the router port is successfully configured, an "S" will be displayed in the port information. FS(config)#show ip igmp snooping mrouter Multicast Switching Mroute Port D: DYNAMIC S: STATIC (*, *, 1): VLAN(1) 1 MROUTES: GigabitEthernet 0/1(S)

Displaying the Information of Dynamic Router Port Learning

Command	**show ip igmp snooping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	Run the **show ip igmp snooping** command to display the aging time and learning status of the dynamic router port. Dynamic Mroute Aging Time : 300(Seconds) Multicast router learning mode: pim-dvmrp

Displaying the Information of a Member Port

Command	**show ip igmp snooping gda-table**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If the member port is successfully configured, an "S" will be displayed in the port information. FS(config)#show ip igmp snooping gda-table Multicast Switching Cache Table D: DYNAMIC S: STATIC M: MROUTE (*, 224.1.1.1, 1): VLAN(1) 1 OPORTS: GigabitEthernet 0/1(S

Displaying Other Parameters

Command	**show ip igmp snooping**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode

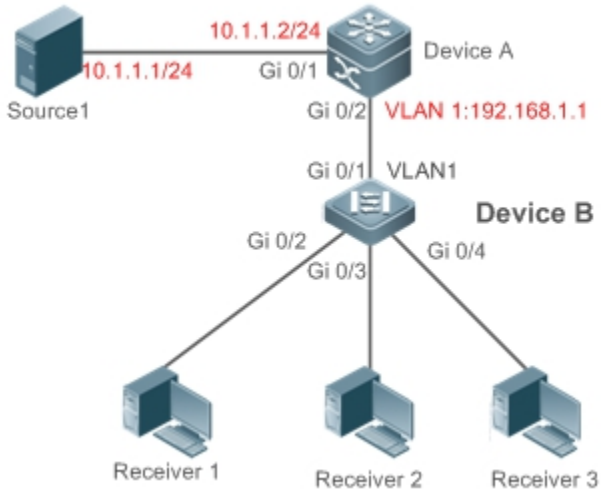
Usage Guide	Run the **show ip igmp snooping** command to display the aging time of the router port, aging time of the dynamic member port, response time of the query packet, and Report packets suppression, and immediate leave. IGMP Fast-Leave: Enable IGMP Report suppress: Enable Query Max Response Time: 20(Seconds) Dynamic Mroute Aging Time : 300(Seconds) Dynamic Host Aging Time : 260(Seconds)

Configuration Example

Configuring a Static Router Port and Static Member Port

Configuration Steps	● Configure basic IGMP snooping functions. ● Configure a static router port and static member port.
	FS# configure terminal FS(config)# ip igmp snooping vlan 1 mrouter interface GigabitEthernet 0/0 FS(config)# ip igmp snooping vlan 1 static 224.1.1.1 interface GigabitEthernet 0/0 FS(config)# end
Verification	Run the **show ip igmp snooping mrouter** and **show ip igmp snooping gda-table** commands to check whether the configuration takes effect.
	FS#show ip igmp snooping mrouter Multicast Switching Mroute Port D: DYNAMIC S: STATIC (*, *, 1): VLAN(1) 1 MROUTES: GigabitEthernet 0/0(S) FS#show ip igmp snooping gda-table Multicast Switching Cache Table D: DYNAMIC S: STATIC M: MROUTE (*, 224.1.1.1, 1): VLAN(1) 1 OPORTS: GigabitEthernet 0/0(SM)

Enabling Report Packet Suppression

Scenario **Figure 2- 8**	
	A is the multicast router and is connected directly to multicast Source 1. B is a Layer-2 device and is connected directly to the user host and multicast Source 2. Receiver 1, Receiver 2, and Receiver 3 are connected to VLAN 1.
Configuration Steps	● Configure the IP address and VLAN. (Omitted) ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select IVGL mode. ● Enable Report packets suppression on B.
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit
B	B# configure terminal B(config)#ip igmp snooping ivgl B(config)# ip igmp snooping suppression enable
Verification	Check whether Receiver 1 and Receiver 2 are added to profile 239.1.1.1, and only the IGMP Report packets of profile 239.1.1.1 are forwarded from interface Gi0/1 of B.
B	B# show ip igmp snooping IGMP Snooping running mode: IVGL IGMP Snooping L2-entry-limit: 65536 Source port check: Disable Source ip check: Disable IGMP Fast-Leave: Disable IGMP Report suppress: Enable

	IGMP Globle Querier: Disable IGMP Preview: Disable IGMP Tunnel: Disable IGMP Snooping version: 2IGMP Preview group aging time : 60(Seconds) Dynamic Mroute Aging Time : 300(Seconds) Dynamic Host Aging Time : 260(Seconds)

Configuring Other Parameters

Configuration Steps	• Configure basic IGMP snooping functions. • Enable Immediate-leave function. • Disable router port learning. • Configure the aging time of a router port. • Configuring the aging time of a member port. • Configure the response time of a Query packet.
	FS# configure terminal FS(config)# ip igmp snooping fast-leave enable FS(config)# no ip igmp snooping mrouter learn pim-dvmrp FS(config)#ip igmp snooping dyn-mr-aging-time 200 FS(config)#ip igmp snooping host-aging-time 100 FS(config)#ip igmp snooping query-max-response-time 60 FS(config)# end
Verification	Run the **show ip igmp snooping** command to check whether the configuration is successful.
	FS#show ip igmp snooping IGMP Snooping running mode: IVGL IGMP Snooping L2-entry-limit: 65536 Source port check: Disable Source ip check: Disable IGMP Fast-Leave: Enable IGMP Report suppress: Enable IGMP Globle Querier: Disable IGMP Preview: Disable IGMP Tunnel: Disable IGMP Snooping version: 2Query Max Response Time: 60(Seconds) IGMP Preview group aging time : 60(Seconds) Dynamic Mroute Aging Time : 200(Seconds) Dynamic Host Aging Time : 100(Seconds)

Common Errors

- Basic IGMP snooping functions are not configured or the configuration is not successful.

8.4.6 Configuring IGMP Security Control

Configuration Effect

- Configure the range of multicast addresses that a user can access.
- Configure to allow a user from an unauthorized profile to preview a multicast channel.
- Configure the number of multicast addresses that a user can access.

- Configure to limit a user to receive only the multicast traffic from a router port to prevent illegal multicast traffic sent by the end user.
- Configure to limit a user to receive only the multicast traffic from designated source IP addresses to prevent illegal multicast traffic.

Notes

- Basic IGMP snooping functions must be configured.

Configuration Steps

Configuring the Profile Filtering

- Optional.
- If you want to limit the profile packets to be received by a port, you can configure the profile filtering on the port.
- If you want to limit the multicast packets to be received by a VLAN, you can configure the per-VLAN profile filtering.

Enabling Multicast Preview

- Optional.
- You can enable multicast preview for a user from an unauthorized profile.

Configuring the Maximum Number of Profiles

- Optional.
- If you want to limit the number of multicast profiles that a port is allowed to receive, you can configure the maximum number of multicast profiles allowed for this port.
- If you want to limit the number of multicast profiles that global ports are allowed to receive, you can configure the maximum number of multicast profiles allowed for these ports.

Configuring Source Port Inspection

- Optional.
- You can perform this configuration if you want to allow a port to receive only the multicast traffic from the router port.

Configuring Source IP Inspection

- Optional.
- You can perform this configuration to specify the source IP address for all the multicast profiles of all VLANs. Only the multicast traffic whose source IP address is the same as the set one is considered as legal.
- You can also specify the source IP addresses for specific multicast profiles within specific VLANs. Among the multicast traffic received from the specific multicast profiles within the VLANs, only the one with the same source IP address as the set one is considered as legal and will be forwarded by the multicast device; other traffic will be discarded.

Verification

- Run the **show ip igmp snooping interfaces** command to display the profile filtering and the maximum number of multicast profiles for a port.
- Run the **show ip igmp snooping vlan** command to display the per-VLAN profile filtering.
- Run the **show ip igmp snooping** command to check whether the maximum number of global multicast profiles, preview function, source port inspection, and source IP address inspection take effect.

Related Commands

Configuring the Profile Filtering

Command	**ip igmp snooping filter** *profile-number*
Parameter Description	*profile-number*: Indicates a profile number.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Per-VLAN Profile Filtering

Command	**ip igmp snooping vlan** *vid* **filter** *profile-number*
Parameter Description	*vid*: Indicates a VLAN. The value ranges from 1 to 4,094. *profile-number*: Indicates a profile number.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Maximum Number of Profiles on a Port

Command	**ip igmp snooping max-groups** *number*
Parameter Description	*number*: Indicates the maximum number of multicast profiles.
Command Mode	Interface configuration mode
Usage Guide	This value indicates only the number of dynamic multicast profiles, and the number of static profiles is not included. The counter of multicast profiles is based on the VLAN that the port belongs to. For example, if a port belongs to three VLANs, and all three of them receive a request packet from multicast profile 224.1.1.1 simultaneously, then the counter of multicast profiles will be 3 but not 1.

Configuring the Maximum Number of Global Profiles

Command	**ip igmp snooping l2-entry-limit** *number*
Parameter Description	*number*: Indicates the maximum number of multicast profiles.
Command Mode	Global configuration mode
Usage Guide	This value includes the number of both dynamic profiles as well as static profiles.

Configuring Source Port Inspection

Command	**ip igmp snooping source-check port**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After source port inspection is enabled, the multicast traffic received by a device will be discarded if no router port is detected in the network environment.

Configuring Source IP Inspection

Command	**ip igmp snooping source-check default-server** *source-address*
Parameter Description	*source-address*: Indicates the source IP address.
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling Source IP Inspection for a Specific Profile

Command	**ip igmp snooping limit-ipmc vlan** *vid* **address** *group-address* **server** *source-address*
Parameter Description	*vid* vlan id *group-address*: Indicates a profile address. *source-address*: Indicates the source IP address.
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling Preview

Command	**ip igmp snooping preview** *profile-number*
Parameter Description	*profile number:* Indicates the range of multicast addresses allowed for preview. The value ranges from 1 to 1,024.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Preview Duration

Command	**ip igmp snooping preview interval** *num*
Parameter Description	*num:* Specifies the preview duration which ranges from 1s to 300s (60s by default).
Command Mode	Global configuration mode
Usage Guide	This configuration allows unauthorized users to receive multicast traffic within the preview duration. After the duration is met, the preview will be stopped; the preview can be resumed in 300s.

Displaying the Per-Port Profile Filtering

Command	**show ip igmp snooping interface**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If the function is configured, the profile will be displayed, for example: FS#show ip igmp snooping interfaces gigabitEthernet 0/1 Interface Filter profile number max-group ---------- ---------- --------- GigabitEthernet 0/1 1

Displaying the Per-VLAN Profile Filtering

Command	**show ip igmp snooping vlan**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If the function is configured, the profile will be displayed, for example:

	IGMP VLAN filter: 1

Displaying the Maximum Number of Interface Profiles

Command	**show ip igmp snooping interface**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If the maximum number of multicast addresses for a port is configures, the value will be displayed, for example: FS#show ip igmp snooping interfaces gigabitEthernet 0/1 Interface Filter profile number max-group ---------- ---------- --------- GigabitEthernet 0/1 1 200

Displaying the Maximum Number of Global Profiles

Command	**show ip igmp snooping vlan**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If the function is configured, the profile will be displayed, for example: IGMP Snooping L2-entry-limit: 65536

Displaying the Information of Source Port Inspection

Command	**show ip igmp snooping vlan**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If source port inspection is enabled, the following information will be displayed: Source port check: Enable

Displaying the Information of Source IP Inspection

Command	**show ip igmp snooping vlan**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If source IP address inspection is enabled, the following information will be displayed: Source ip check: Enable

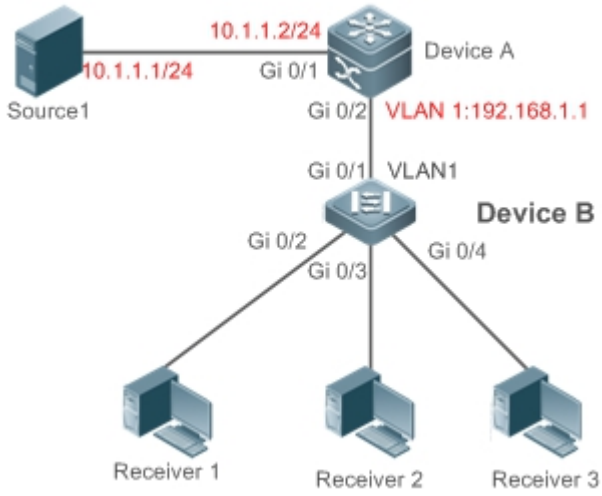
Displaying the Information of the Preview Function

Command	**show ip igmp snooping**
Parameter Description	N/A

Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If the range of multicast addresses for a port is configured, preview will be enabled, for example: IGMP Preview: Enable IGMP Preview group aging time : 60(Seconds)

Configuration Example

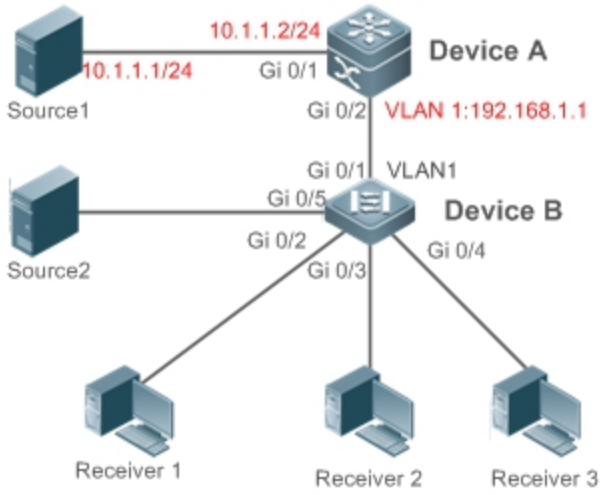
Configuring the Profile Filtering and the Maximum Number of Demanded Profiles

Scenario **Figure 2- 9**	
	A is the multicast router and is connected directly to multicast Source 1. B is a Layer-2 device and is connected directly to the user host and multicast Source 2. Receiver 1, Receiver 2, and Receiver 3 are connected to VLAN 1. By configuring VLAN 1, you can configure to allow the users within VLAN 1 to receive only the profiles whose addresses range from 225.1.1.1 to 225.1.255.255. You can configure Receiver 1 to receive only the profiles whose addresses range from 225.1.1.1 to 225.1.1.255, Receiver 2 to receive only the profiles whose addresses range from 225.1.2.1 to 255.1.2.255, and Receiver 3 to receive only the profiles whose addresses range from 225.1.3.1 to 225.1.3.255. At most 10 profiles can be added to a port and at most 100 profiles can be added globally.
Configuration Steps	● Configure the IP address and VLAN. (Omitted) ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select IVGL mode. ● Configure the range and maximum number of multicast addresses on B.
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit

B	B# configure terminal B(config)#ip igmp snooping ivgl B(config)#ip igmp profile 1 B(config-profile)#permit B(config-profile)#rang B(config-profile)#range 225.1.1.1 225.1.255.255 B(config-profile)#exit B(config)#ip igmp profile 2 B(config-profile)#permit B(config-profile)#range 225.1.1.1 225.1.1.255 B(config-profile)#exit B(config)#ip igmp profile 3 B(config-profile)#permit B(config-profile)#range 225.1.2.1 225.1.2.255 B(config-profile)#exit B(config)#ip igmp profile 4 B(config-profile)#permit B(config-profile)#range B(config-profile)#range 225.1.3.1 225.1.3.255 B(config-profile)#exit B(config)#ip igmp snooping l2-entry-limit 100 B(config)#ip igmp snooping vlan 1 filter 1 B(config)#int gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/0)#ip igmp snooping filter 2 FS(config-if-GigabitEthernet 0/0)#ip igmp snooping max-groups 10 B(config)#int gigabitEthernet 0/3 FS(config-if-GigabitEthernet 0/0)#ip igmp snooping filter 3 FS(config-if-GigabitEthernet 0/0)#ip igmp snooping max-groups 10 B(config)#int gigabitEthernet 0/4 FS(config-if-GigabitEthernet 0/0)#ip igmp snooping filter 4 FS(config-if-GigabitEthernet 0/0)#ip igmp snooping max-groups 10
Verification	● Run the **show ip igmp snooping interfaces** command to display the profile filtering and the maximum number of multicast profiles for a port. ● Run the **show ip igmp snooping** command to display the maximum number of global multicast groups.
B	B#show ip igmp snooping interfaces Interface Filter profile number max-group ---------- ---------- --------- GigabitEthernet 0/2 2 10 GigabitEthernet 0/3 3 10 GigabitEthernet 0/4 4 10 B#show ip igmp snooping IGMP Snooping running mode: IVGL IGMP Snooping L2-entry-limit: 100

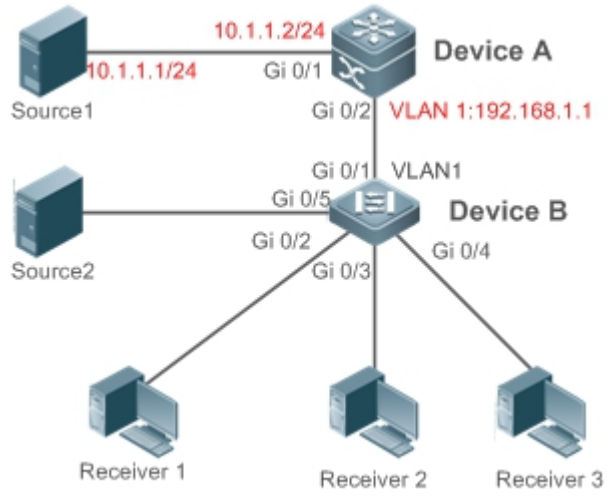
	Source port check: Disable Source ip check: Disable IGMP Fast-Leave: Disable IGMP Report suppress: Disable IGMP Globle Querier: Disable IGMP Preview: Disable IGMP Tunnel: Disable IGMP Preview group aging time : 60(Seconds) Dynamic Mroute Aging Time : 300(Seconds) Dynamic Host Aging Time : 260(Seconds)

Configuring Source Port Inspection

Scenario **Figure 2- 10**	
	A is the multicast router and is connected directly to multicast Source 1. B is a Layer-2 device and is connected directly to the user host and multicast Source 2. Receiver 1, Receiver 2, and Receiver 3 are connected to VLAN 1. Source 1 sends the multicast address traffic from profile 224.1.1.1, and Source 2 sends the multicast address traffic from profile 225.1.1.1. Receiver 1 can request profiles 224.1.1.1 and 225.1.1.1 respectively. Source port inspection is enabled.
Configuration Steps	● Configure the IP address and VLAN. ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select IVGL mode. ● Enable source port inspection on B.
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1

	A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit
B	B# configure terminal B(config)#ip igmp snooping ivgl B(config)#ip igmp snooping source-check port
Verification	● Run the **show ip igmp snooping mroute** command to check whether Gi0/1 is learned as a router port. ● Check whether Receiver 1 can request the multicast traffic of profile 224.1.1 and cannot request that of profile 225.1.1.1.
B	Multicast Switching Mroute Port D: DYNAMIC S: STATIC (*, *, 1): VLAN(1) 1 MROUTES: GigabitEthernet 0/1(S) B#show ip igmp snooping IGMP Snooping L2-entry-limit: 100 Source port check: Enable Source ip check: Disable

Configuring Source IP Inspection

Scenario **Figure 2- 11**	
	A is the multicast router and is connected directly to multicast Source 1. B is a Layer-2 device and is connected directly to the user host and multicast Source 2. Receiver 1, Receiver 2, and Receiver 3 are connected to VLAN 1. Source 1 sends the multicast address traffic from profiles 10.1.1.1 and 224.1.1.1, Source 2 sends the multicast address traffic from profiles 192.168.1.3 and 225.1.1.1, and Source 3 sends the multicast address traffic from profiles 192.168.1.3 and 226.1.1.1. Receiver 1 can request profiles 224.1.1.1, 225.1.1.1, and 226.1.1.1 respectively. The default IP address for source IP inspection is 10.1.1.1. Configure limit-ipmc and the multicast traffic of profile 225.1.1.1, and set the legal source address as 192.168.1.3.

Configuration Steps	● Configure the IP address and VLAN. ● Enable multicast routing on A and enable the multicast routing protocol on Layer-3 interface (Gi0/1 and VLAN 1). ● Enable IGMP snooping on B and select IVGL mode. ● Enable source port inspection on B.
A	A# configure terminal A(config)# ip multicast-routing A(config)# interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)# interface vlan 1 A(config-if-VLAN 1)# ip pim sparse-mode A(config-if-VLAN 1)# exit
B	B# configure terminal B(config)#ip igmp snooping ivgl B(config)# ip igmp snooping source-check default-server 10.1.1.1 B(config)# ip igmp snooping limit-ipmc vlan 1 address 225.1.1.1 server 192.168.1.3
Verification	● Run the **show ip igmp snooping** command to check whether source IP inspection is enabled. ● Check whether Receiver 1 can request the multicast traffic of profile 224.1.1 and 225.1.1.1 and cannot request that of profile 226.1.1.1.
B	B#show ip igmp snooping IGMP Snooping running mode: IVGL IGMP Snooping L2-entry-limit: 65536 Source port check: Disable Source ip check: Enable IGMP Fast-Leave: Disable IGMP Report suppress: Disable IGMP Globle Querier: Disable IGMP Preview: Disable IGMP Tunnel: Disable IGMP Preview group aging time : 60(Seconds) Dynamic Mroute Aging Time : 300(Seconds) Dynamic Host Aging Time : 260(Seconds)

Common Errors

- Basic IGMP snooping functions are not configured or the configuration is not successful.
- The multicast router port is not learned, leading to failure to receive the multicast traffic.
- The IP address for source IP inspection is inconsistent with the multicast IP address, leading to failure to receive the multicast traffic.

8.4.7 Configuring an IGMP Profile

Configuration Effect

- Create an IGMP filtering profile.

Configuration Steps

Creating a Profile

- (Optional) Create an IGMP filtering profile.

Configuring the Profile Range

- (Optional) Configure the range of multicast profile addresses.

Configuring the Profile Filtering

- (Optional) Configure the filtering mode of profile to **permit** or **deny**.

Verification

- Run the **show running-config** command to check whether the preceding configurations take effect.

Related Commands

Creating a Profile

Command	**ip igmp profile** *profile-number*
Parameter Description	*profile-number*: Indicates the number of a profile.
Command Mode	Global configuration mode
Usage Guide	

Configuring the Profile Range

Command	**range** *low-ip-address* [*high-ip-address*]
Parameter Description	*low-ip-address*: Specifies the start address. *low-ip-address*: Specifies the end address. Only one address is configured by default.
Command Mode	Profile configuration mode
Usage Guide	You can configure multiple addresses. If the IP addresses of different ranges are consecutive, the addresses will be combined.

Configuring the Profile Filtering

Command	**deny**
Parameter Description	N/A
Command Mode	Profile configuration mode
Usage Guide	If the filtering mode of profile is set to **deny** while the range of multicast profiles is not specified, no profile is to be denied, which means to permit all profiles.

Configuring the Profile Filtering

Command	**permit**
Parameter Description	N/A
Command Mode	Profile configuration mode
Usage Guide	If the filtering mode of profile is set to **permit** while the range of multicast profiles is not specified, no profile is to be permitted, which means to deny all profiles.

Configuration Example

Creating a Filtering Profile

Configuration Steps	● Create a filtering profile.
	B(config)#ip igmp profile 1 B(config-profile)#permit B(config-profile)#range B(config-profile)#range 224.1.1.1 235.1.1.1 B(config-profile)#
Verification	Run the **show running-config** command to check whether the configuration is successful.

	ip igmp profile 1
	  permit
	  range 224.1.1.1 235.1.1.1
	!

Common Errors

- Basic IGMP snooping functions are not configured or the configuration is not successful.
- The mode of profile is set to **permit** while the range of multicast profiles is not specified, leading to the denial of all profiles.

8.4.8 Configuring IGMP QinQ

Configuration Effect

- Create a multicast entry on the VLAN where IGMP packets are located. Forward IGMP packets on the VLAN where these packets are located, realizing transparent transmission.

Notes

- Basic IGMP snooping functions must be configured.

Configuration Steps

Configuring QinQ Transparent Transmission

- If the QinQ interface needs to forward multicast packets on the VLANs where the VIDs of the packets specify, enable QinQ to realize transparent transmission.

Verification

- Run the **show ip igmp snooping** command to check whether the configuration takes effect.

Related Commands

Configuring QinQ Transparent Transmission

Command	**ip igmp snooping tunnel**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Enable QinQ to realize transparent transmission of IGMP packets.

Displaying QinQ Configuration

Command	**show ip igmp snooping**
Parameter Description	N/A

Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If QinQ is enabled, the following content is displayed. IGMP Tunnel: Enable

Configuration Example

Configuring QinQ Transparent Transmission

Configuration Steps	● Configure basic IGMP snooping functions. ● Configure QinQ transparent transmission.
	FS# configure terminal FS(config)# ip igmp snooping tunnel FS(config)# FS(config)# end
Verification	Run the **show ip igmp snooping** command to check whether the configuration is successful.
	IGMP Tunnel: Enable

Common Errors

- Basic IGMP snooping functions are not configured or the configuration is not successful.

8.4.9 Configuring an IGMP Querier

Configuration Effect

- Configure the device as an IGMP querier, which will send IGMP Query packets periodically and collect user demanding information.

Notes

- Basic IGMP snooping functions must be configured.

Configuration Steps

Enabling the Querier Function

- (Optional) Enable IGMP querier function globally or for a specified VLAN.
- (Optional) Disable the IGMP querier function for a specified VLAN.

Configuring the Source IP Address of a Querier

- (Optional) You can configure the source IP address of a Query packet sent by the querier based on VLANs.
- After a querier is enabled, a source IP address must be specified for the querier; otherwise, the configuration will not take effect.

Configuring the Maximum Response Time of a Query Packet

- (Optional) Adjust the maximum response time carried by an IGMP Query packet. As IGMPv1 does not support the carrying of maximum response time by a Query packet, this configuration does not take effect when the querier is running IGMPv1.

Configuring the Query Interval of a Querier

- (Optional) Adjust the interval of the IGMP querier for sending query packets.

Configuring the Aging Timer of a Querier

- (Optional) Configure the aging timer of other IGMP queriers on the network.

Specifying the IGMP Version for a Querier

- (Optional) Specify the IGMP version for a querier (IGMPv2 by default).

Verification

- Run the **show ip igmp snooping querier detail** command to check whether the configuration takes effect.

Related Commands

Enabling the IGMP Querier Function

Command	**ip igmp snooping** [**vlan** *vid*] **querier**

Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default.
Command Mode	Global configuration mode
Usage Guide	IGMP querier for a specified VLAN will take effect only after global IGMP querier is enabled. If global IGMP querier is disabled, IGMP querier for all the VLANs will be disabled.

Configuring the Source IP Address of a Querier

Command	**ip igmp snooping** [**vlan** *vid*] **querier address** *a.b.c.d*
Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default. *a.b.c.d*: Indicates the source IP address.
Command Mode	Global configuration mode
Usage Guide	After a querier is enabled, a source IP address must be specified for the querier; otherwise, the configuration will not take effect. If the source IP address is specified by a VLAN, the address will be used preferentially.

Configuring the Maximum Response Time of a Querier

Command	**ip igmp snooping** [**vlan** *vid*] **querier max-response-time** *seconds*
Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default. *seconds*: Indicates the maximum response time. in the unit of seconds. The value ranges from 1 to 25.
Command Mode	Global configuration mode
Usage Guide	If the query interval is specified by a VLAN, the value will be used preferentially.

Configuring the Query Interval of a Querier

Command	**ip igmp snooping** [**vlan** *vid*] **querier address** *a.b.c.d*
Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default. *seconds*: Indicates the query interval in the unit of seconds. The value ranges from 1 to 18,000.
Command Mode	Global configuration mode
Usage Guide	If the query interval is specified by a VLAN, the value will be used preferentially.

Configuring the Aging Timer of a Querier

Command	**ip igmp snooping** [**vlan** *vid*] **querier timer expiry** *seconds*
Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default. *seconds*: Indicates the timeout time in the unit of seconds. The value ranges from 60 to 300.
Command Mode	Global configuration mode
Usage Guide	A device may fail to be elected as the querier even when its querier function is enabled. If a device that fails to be elected does not receive the Query packet sent by the querier in the aging time, the querier in use is considered as expired, and a new round of election will be raised. If the aging time is specified by a VLAN, the value will be used preferentially.

Specifying the IGMP Version for a Querier

Command	**ip igmp snooping** [**vlan** *vid*] **querier version 1**
Parameter Description	**vlan** *vid*: Specifies a VLAN. This configuration applies to all VLANs by default.
Command Mode	Global configuration mode
Usage Guide	A querier can be run in IGMPv1 and IGMPv2 (IGMPv2 by default). You can also run a command to configure the version to IGMPv1. If the IGMP version for a querier is specified by a VLAN, the version will be used preferentially.

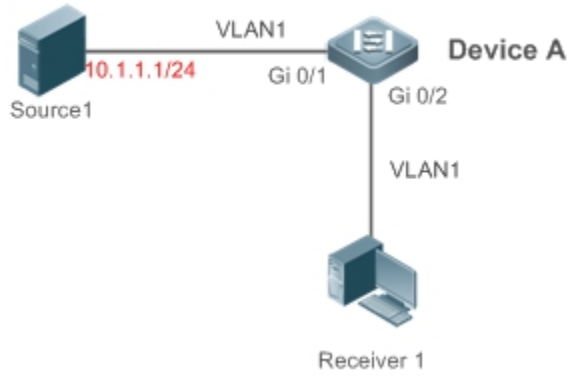
Displaying the IGMP Querier Configuration

Command	**show ip igmp snooping querier detail**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	If QinQ is enabled, the following content is displayed.

```
FS(config)#show ip igmp snooping querier detail
Vlan        IP Address                IGMP Version                Port
-------------------------------------------------------


Global IGMP switch querier status
-------------------------------------------------------
admin state                          : Enable
admin version                        : 2
source IP address                    : 1.1.1.1
query-interval (sec)                 : 60
max-response-time (sec)              : 10
querier-timeout (sec)                : 125

Vlan 1:     IGMP switch querier status
-------------------------------------------------------
admin state                          : Disable
admin version                        : 2
source IP address                    : 1.1.1.1
query-interval (sec)                 : 60
max-response-time (sec)              : 10
querier-timeout (sec)                : 125
operational state                    : Disable
operational version                  : 2
```

Configuration Example

Enabling the IGMP Querier Function

Scenario **Figure 2-12**	
	In the scenario without Layer-3 multicast equipment, the multicast traffic can be forwarded only on the Layer-2 network. A acts as a Layer-2 device to connect to the multicast source and receiver.
Configuration Steps	● Enable global IGMP snooping on A in IVGL mode. ● Enable IGMP querier for VLAN 1 on A.
A	A(config)#ip igmp snooping ivgl A(config)#ip igmp snooping querier A(config)#ip igmp snooping querier address 10.1.1.1 A(config)#ip igmp snooping vlan 1 querier
Verification	Run the **show ip igmp snooping querier** command to check whether the querier of VLAN 1 takes effect.
A	A(config)#show ip igmp snooping querier Vlan IP Address IGMP Version Port ---------- 1 10.1.1.1 2 switch A(config)#show ip igmp snooping querier vlan 1 Vlan 1: IGMP switch querier status ---------- elected querier is 10.1.1.1 (this switch querier) ---------- admin state : Enable admin version : 2 source IP address : 10.1.1.1 query-interval (sec) : 60 max-response-time (sec) : 10 querier-timeout (sec) : 125 operational state : Querier operational version : 2

Common Errors

- The source IP address is not configured for the querier and the querier does not take effect.

8.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears the statistics on IGMP snooping.	**clear ip igmp snooping statistics**
Clears the dynamic router ports and member ports.	**clear ip igmp snooping gda-table**

Displaying

Description	Command
Displays basic IGMP snooping configurations.	**show ip igmp snooping** [**vlan** *vlan-id*]
Displays the statistics on IGMP snooping.	**show ip igmp snooping statistics** [**vlan** *vlan-id*]
Displays the router ports.	**show ip igmp snooping mrouter**
Displays the IGMP snooping entries.	**show ip igmp snooping gda-table**
Displays the profile.	**show ip igmp profile** [*profile-number*]
Displays the IGMP snooping configurations on an interface.	**show ip igmp snooping interface** *interface-name*
Displays the IGMP querier.	**show ip igmp snooping querier** [**detail**]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs all IGMP Snooping functions.	**debug igmp-snp**
Debugs the IGMP snooping events.	**debug igmp-snp event**
Debugs the IGMP snooping packets.	**debug igmp-snp packet**
Debugs the communications between IGMP snooping and MSF.	**debug igmp-snp msf**
Debugs the IGMP snooping alarms.	**debug igmp-snp warning**

9 Configuring MLD Snooping

9.1 Overview

Multicast Listener Discovery (MLD) Snooping is used to control and manage the forwarding behaviors of IPv6 multicast packets at Layer 2.

The device running MLD Snooping analyzes MLD packets received by a port to create a mapping between the port and the MAC multicast address and forwards IPv6 multicast data at Layer 2 based on the mapping. When MLD Snooping is disabled, IPv6 multicast data packets are broadcasted at Layer 2. When MLD Snooping is enabled, multicast data packets of a known IPv6 multicast group are forwarded to a specified receiver at Layer 2 instead of being broadcasted at Layer 2.

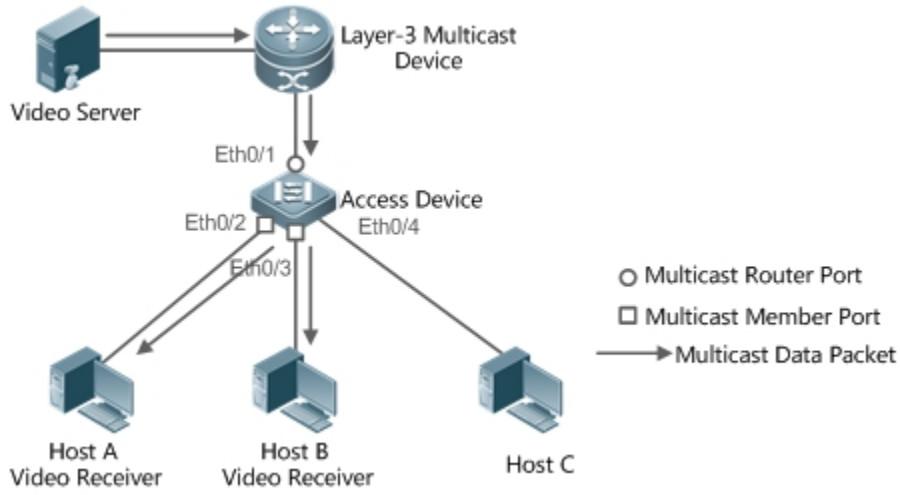
Protocols and Standards

RFC4541: Considerations for Internet Group Management Protocol (IGMP)and Multicast Listener Discovery (MLD) Snooping Switches

9.1.1 Two Types of MLD Snooping Ports

As shown in Figure 9- 1, the Layer-3 multicast device is connected to the multicast source. MLD Snooping is enabled on the access device. Host A and Host B are receivers (that is, members of the IPv6 multicast group).

Figure 9- 1 Two Types of MLD Snooping Ports



- Multicast router port: Indicates the port on the access device for connecting to the Layer-3 multicast device, for example, Port Eth0/1 of the access device.
- Member port: Is short for IPv6 multicast group member port, also called listener port, and indicates the port on the access device for connecting to an IPv6 multicast group member, for example, Port Eth0/2 and Port Eth0/3 on the access device.

9.1.2 Work Mode of MLD Snooping

- DISABLE mode: MLD Snooping does not take effect in this mode. That is, the Layer-2 multicast device does not "snoop" MLD packets between the host and the router, and multicast streams are broadcasted within VLANs.

- Independent VLAN Group Learn (IVGL) mode: In this mode, multicast streams between VLANs are mutually independent. A host can request only the multicast router port in the same VLAN as the host to receive multicast packets, and can forward the received multicast data packets of any VLAN only to the member port and multicast router port in the same VLAN as the host.

- Shared VLAN Group Learn (SVGL) mode: In this mode, hosts of VLANs share the same multicast stream. A host in one VLAN can request multicast streams of another VLAN. When a shared VLAN is specified, only the multicast data streams of this VLAN can be forwarded to hosts of other VLANs. Multicast data streams of a shared VLAN, can be forwarded to the member ports of this multicast address, even though some member ports do not belong to the shared VLAN. In SVGL mode, MLD profiles must be used to allocate a batch of multicast address ranges to SVGL. Within the multicast address ranges, member ports in the multicast forwarding entries support trans-VLAN packet forwarding. By default, all the group ranges are not within the SVGL application ranges, and all the multicast packets are discarded.

- IVGL-SVGL mode: In this mode, IVGL and SVGL coexist. You can use MLD profiles to allocate a batch of multicast address ranges to SVGL. Within the multicast address ranges, member ports in the multicast forwarding entries support trans-VLAN packet forwarding. Member ports in the multicast forward entries corresponding to other multicast address ranges must belong to the same VLAN.

9.1.3 Working Principle of MLD Snooping

The device running MLD Snooping processes different MLD packets as follows:

MLD QUERY

The Layer-3 multicast device regularly sends an MLD General Query packet to all hosts and routers (with the address of FF02::1) in the local network segment, to query the IPv6 multicast group members in this network segment. When receiving the MLD General Query packet, the device running MLD Snooping forwards the packet all ports in the VLAN except the one receiving the packet, and processes the port receiving the packet as follows:

- If the port is already in the router multicast port list, its aging timer is reset.

- If the port is not contained in the router multicast port list, the port is added to the router multicast port list and its aging timer is started.

- Each time the Layer-2 multicast device receives an MLD General Query packet, it starts the aging timer for each member port, and updates the timer time to the configured maximum response time of MLD query packet. When the aging timer time of a port is reduced to 0, it is deemed that no member receives multicast streams through this port, and therefore, the Layer-2 multicast device deletes the port from the MLD Snooping forwarding table.

- Each time the Layer 2 multicast device receives a MLD Group-Specific Query packet, it starts the aging timer for each member port in the specific group, and updates the timer time to the configured maximum response time of MLD query packet. When the aging timer time of a port is reduced to 0, it is deemed that no member receives multicast streams through this port, and therefore, the Layer-2 multicast device deletes the port from the MLD Snooping forwarding table.

- When the Layer-2 multicast device receives a MLD Group-Specific Query packet, it no longer updates the preceding two types of timers.

MLD REPORT

In either of the following cases, the host sends an MLD Membership Report packet to the MLD querier.

- After receiving an MLD query (General Query or Group-Specific Query) packet, an IPv6 multicast group member host responds with an MLD Membership Report packet.

- If a host needs to join an IPv6 multicast group, it actively sends an MLD Membership Report packet to MLD querier to request to join this IPv6 multicast group.

When receiving an MLD Membership Report packet, the device running MLD Snooping forwards it to all multicast router ports in the VLAN, retrieves, from the packet, the address of the IPv6 multicast group that the host needs to join, and processes the port receiving the packet as follows:

- If there is no forwarding entry corresponding to the IPv6 multicast group, the forwarding entry is created, the port is added to the egress port list as a dynamic member port, and its aging timer is started.
- If there is a forwarding entry corresponding to the IPv6 multicast group but the port is not contained in the egress port list, the port is added to the egress port list as a dynamic member port, and its aging timer is started.
- If there is a forwarding entry corresponding to the IPv6 multicast group and dynamic member port is contained in the egress port list, its aging timer is reset.

MLD LEAVE

When a host leaves an IPv6 multicast group, it sends an MLD Leave packet (with the address of FF02::2) to notify the multicast router that it has left the IPv6 multicast group. When receiving an MLD Leave packet from a member port, the device running MLD Snooping directly forwards it to the multicast router port. If the fast leave function is enabled, the device directly deletes the port from the forwarding port list of the relevant multicast group.

9.1.4 Source Port Check

The source port check function of MLD Snooping improves the network security.

This function strictly limits the ingress ports of MLD multicast streams. When this function is disabled, multicast streams from any port are valid and the Layer-2 multicast device forwards them to registered member ports according to the forwarding list of MLD Snooping. When this function is enabled, multicast streams only from the multicast router ports are valid and the Layer-2 multicast device forwards them to registered ports. Multicast data streams from non-multicast router ports are invalid and discarded.

9.2 Applications

Application	Description
MLD Snooping SVGL Trans-VLAN Multicast On demand	MLD Snooping works in SVGL mode
Source Port Filtering	Multicast streams only from multicast router ports are received.

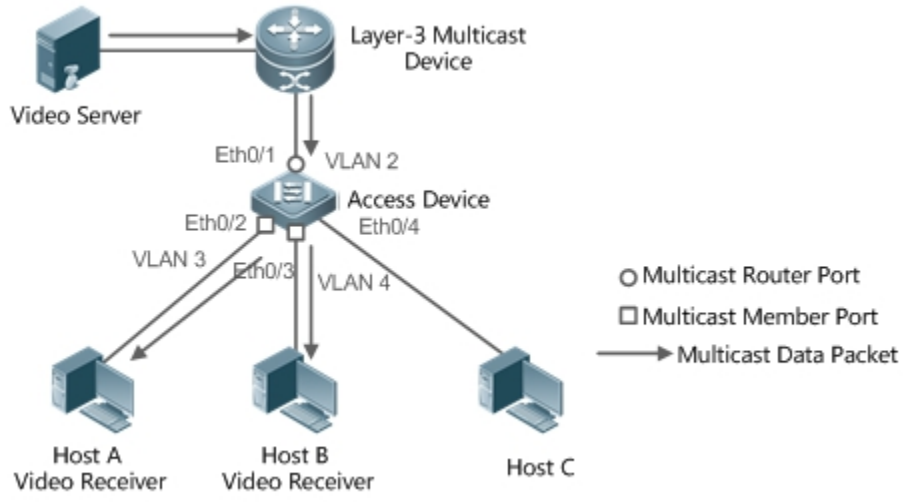
9.2.1MLD Snooping SVGL Trans-VLAN Multicast On demand

Scenario

As shown in Figure 9- 2, Host A of VLAN 3 and Host B of VLAN 4 order a video. The video streams are in VLAN 2.

- Enable the SVGL mode on the access device and set a shared VLAN 2.

Figure 9- 2



Remarks	VLAN 2 is a shared VLAN. VLAN 3 and VLAN 4 are the VLANs through which the video on-demand service is output.

Deployment

- Enable the Layer-3 multicast protocol on the Layer-3 multicast device.
- Enable the SVGL mode on the Layer-2 device.

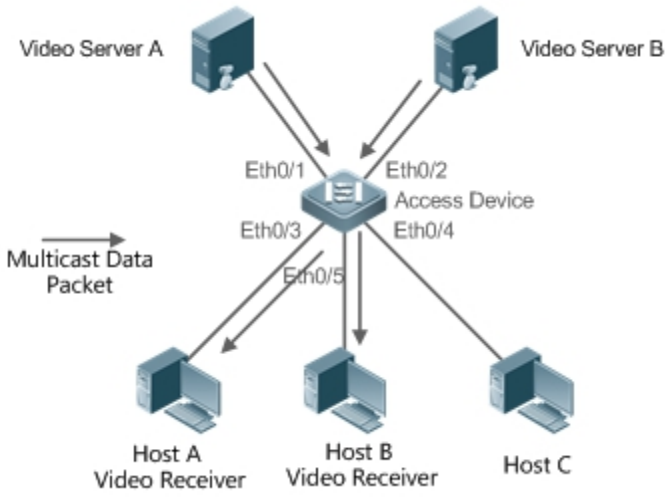
9.2.1 Source Port Filtering

Scenario

As shown in Figure 9- 3, when the source port check function is configured, video streams can be received only from the source multicast router port. Multicast video streams from other ports are invalid and discarded. Note that when the source port check function is configured, there shall be at least one multicast router port. Otherwise, packet filtering is not performed on the multicast router port even though the source port filtering is enabled. When the source port check function is not configured, multicast video streams from all ports are received by default.

- Enable the IVGL mode on the access device.

Figure 9- 3



Remarks	Port Eth0/1 is a multicast router port and Port Eth0/2 is a non-multicast router port. Video servers send same multicast video streams. Hosts A and B can receive multicast streams only from Video Server A.

Deployment

- Enable the source port check function and configure a static multicast router port.
- Enable the IVGL mode on the Layer-2 device.

9.3 Features

Basic Concepts

Multicast Router Port and Member Port

Multicast router ports are classified into dynamic multicast router ports and static multicast router ports. If MLD Snooping is enabled, when the dynamic multicast router port learning function is enabled on a port, after receiving an MLD Query or PIMv6-Hello packet, the port learns the dynamic multicast router port and starts the aging timer of the dynamic multicast router port. A static multicast router port can be added by configuring the **ipv6 mld snooping vlan mrouter** command.

Member ports are classified into dynamic member ports and static member ports. If MLD Snooping is enabled, after receiving an MLD Report packet, a port learns the dynamic member router port and starts the aging timer of the dynamic member port. A static member port can be added by configuring the **ipv6 mld snooping vlan static interface** command.

Fast Leave and Packet Suppression

When the fast leave function is enabled, a port is directly deleted after receiving an MLD Leave packet. The fast leave function is applicable only to scenarios in which only one user is connected to a port, and helps save the bandwidth. When multiple users are connected to a port, if the fast leave function is enabled, other users wanting to receive packets fail to receive any packets.

When the packet suppression function is enabled, only the first MLD Report packet is forwarded within one query period.

Overview

Feature	Description
Globally Enabling MLD Snooping	Globally enables MLD Snooping and configures the work mode.
VLAN-based MLD Snooping	Enables or disables MLD Snooping for a single VLAN when MLD Snooping is globally enabled.
Aging Time of Multicast Router Ports	Adjusts the aging time of dynamic multicast router ports. The default aging time is 300s.
Dynamic Multicast Router Port Learning	After receiving an MLD query packet or a PIMv6 Hello packet, the port is learnt as a dynamic multicast router port.
Fast Leave of Multicast Group Member Ports	A member port can be quickly deleted, instead of being aged and deleted after the query interval of a Group-Specific Query expires.
MLD Report Packet Suppression	Only the first Report packet is processed within one query period, reducing the work load of the module.
Source Port Check	Multicast streams received only from a multicast router port can be forwarded. Packets received from non-multicast router ports cannot be forwarded.
Port-based Specific Multicast Group Filtering	Only multicast group packets that meet the filter conditions can be received.
Maximum Number of Multicast Groups Supported by a Port	Limits the maximum number of multicast groups that a port can join.

9.3.1 Globally Enabling MLD Snooping

Globally enable MLD Snooping and configure the work mode. Multicast forwarding entries can be learnt and multicast streams are forwarded to a specified port.

Working Principle

Enable MLD Snooping. When an MLD Report packet with the time to live (TTL) of 1 is received, a multicast forwarding entry is created and the forwarding egress is this port.

Learning a Dynamic Member Port

After a valid MLD Report packet is received, a dynamic member port is learnt and a forwarding entry is generated. The forwarding egress of this entry is the member port.

Coordinating Parameters

Configure the MLD Report packet suppression function.

Related Configuration

Configure the MLD Report packet suppression function so that only the first Report is processed within one query period, thereby reducing the number of packets in the network.

9.3.1 VLAN-based MLD Snooping

Enable or disable MLD Snooping for a single VLAN. By default, if MLD Snooping is globally enabled, the MLD Snooping function of each VLAN is enabled.

Related Configuration

Globally configure MLD Snooping. Then configure MLD Snooping for a single VLAN.

9.3.2 Aging Time of Multicast Router Ports

Multicast router ports are classified into dynamic multicast router ports and static multicast router ports. By default, the aging time of a dynamic multicast router port is 300s. Static multicast router ports are not aged.

Related Configuration

Ability of learning from dynamic multicast router port learning function

9.3.3 Dynamic Multicast Router Port Learning

By default, all ports support the dynamic multicast router port learning function.

Working Principle

When a port supports the dynamic multicast router port learning function, after receiving an MLD query packet or a PIMv6 Hello packet, the port is learnt as a dynamic multicast router port.

Related Configuration

Configure a port as a static multicast router port.

9.3.4 Aging Time of Dynamic Member Ports

Member ports are classified into dynamic member ports and static member ports. By default, the aging time of a dynamic member port is 260s. Static member ports are not aged.

9.3.5 Fast Leave of Multicast Group Member Ports

By default, the fast leave function of multicast group member ports are disabled. If the fast leave function is enabled, the port is directly deleted after receiving a done packet.

9.3.6 MLD Report Packet Suppression

By default, the MLD report packet suppression function is disabled. If the function is enabled, only the first Report packet is processed within one query interval, thereby reducing the number of packets in the network.

9.3.7 Source Port Check

The source port check function is disabled by default.

Working Principle

When the source port check function is enabled, packets only from multicast router ports are valid and packets from non-multicast router ports are invalid.

Related Configuration

Configure a port as a static multicast router port.

9.3.8 Port-based Specific Multicast Group Filtering

Under certain circumstances, you may use the port filtering function to control a port to forward multicast packets only of a certain range.

9.3.9 Maximum Number of Multicast Groups Supported by a Port

The maximum number of multicast groups that a port is allowed to join can control the maximum number of multicast groups supported by the port.

9.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of MLD Snooping	**ipv6 mld snooping**	Enables MLD Snooping and specifies the work mode.
	ipv6 mld snooping vlan *vlan-id* **mrouter interface** *interface-id*	Configures the static multicast router port.
	ipv6 mld snooping vlan *vlan-id* **static** *ip-addr* **interface** *interface-id*	Configures a static member port.
	ipv6 mld profile *profile-num*	Configures a profile.
	ipv6 mld snooping source-check port	Configures source port check.
	ipv6 mld snooping filter *profile-num*	Configures multicast group filtering for a port.
	ipv6 mld snooping max-groups *num*	Configures the maximum number of multicast groups that a port can join.

9.4.1 Configuring Basic Functions of MLD Snooping

Configuration Effect

- Enable MLD Snooping and configure the work mode.

Notes

- Enable MLD Snooping and set the work node to SVGL. The MLD Snooping SVGL mode cannot coexist with IPv4 or IPv6 Layer-3 multicasting.
- When the work mode is SVGL or IVGL-SVGL, a profile must be associated to specify the multicast group range in which the SVGL mode applies.

Configuration Steps

Enabling IPv6 MLD Snooping

- Mandatory.

Verification

Run the **show ipv6 mld snooping** command to check whether MLD Snooping is enabled.

- Check whether the device can create correct multicast forwarding entries.

Related Commands

Enabling IPv6 MLD Snooping

Command	**ipv6 mld snooping** *mode*
Parameter Description	**mode**:Specifies the work mode.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Profile

Command	**ipv6 mld profile** *profile-num*
Parameter Description	*profile-num*: Indicates the profile number.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure a profile and enter the profile configuration mode.

Configuring a Static Multicast Router Port

Command	**ipv6 mld snooping vlan** *vlan-id* **mrouter interface** *interface-id*
Parameter Description	***vlan-id:*** Indicates the VLAN ID**.** *interface-id:* Indicates interface changes.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Static Member Port

Command	**ipv6 mld snooping vlan** *vlan-id* **static** *ip-addr* **interface** *interface-id*
Parameter Description	***vlan-id:*** Indicates the VLAN ID. *ip-addr:* Indicates the group address. *interface-id:* Indicates interface changes.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring Source Port Check

Command	**ipv6 mld snooping source-check port**
Parameter Description	
Command Mode	Global configuration mode
Usage Guide	-

Configuring Port-based Multicast Group Filtering

Command	**ipv6 mld snooping filter** *profile-num*
Parameter Description	*profile-num:* Indicates the profile number.
Command Mode	Interface configuration port
Usage Guide	N/A

Configuring the Maximum Number of Multicast Groups Supported by a Port

Command	**ipv6 mld snooping max-groups** *num*
Parameter Description	*num:* Indicates the number of groups.
Command Mode	Interface configuration port
Usage Guide	N/A

Configuring Report Packet Suppression

Command	**ipv6 mld snooping suppression enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	When the Report packet suppression function is enabled, only the first Report packet of a specific VLAN and group is forwarded to a multicast router port within one query interval. The subsequent Report packets are forwarded to the multicast router port, so as to reduce the number of packets in the network. This function can only suppress the Report packets of MLDv1. It is invalid on the Report packets of MLDv2.

Configuring Port Fast Leave

Command	**ipv6 mld snooping fast-leave enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	When the port fast leave function is enabled, after receiving a Leave packet, the port is directly deleted from the member

	ports in the corresponding forwarding entries. Later, when receiving a relevant Group-Specific Query packet, the device does not forward the packet to this port. The Leaver packet includes the Leave packet of MLDv1, include type of MLDv2, and Report packet containing no source address. This function is applicable only to scenarios in which only one user is connected to a port, and helps save bandwidth and resources.

Configuring Dynamic Multicast Router Port Learning

Command	**ipv6 mld snooping** [**vlan** *vid*] **mrouter learn**
Parameter Description	vlan-id: Specifies a VLAN ID. This function is applicable to all VLANs by default.
Command Mode	Global configuration mode
Usage Guide	A multicast router port is a port that directly connects an MLD Snooping-enabled multicast device to a neighbor multicast device in which a multicast routing protocol is enabled. By default, when the dynamic multicast router port learning function is enabled, the device automatically listens to the MLD Query/PIM Hello packet and dynamically identifies a multicast router port.

Configuring Aging Time of Dynamic Multicast Router Ports

Command	**ipv6 mld snooping dyn-mr-aging-time** *seconds*
Parameter Description	***seconds***: Indicates the aging time of dynamic multicast router ports. The unit is second and the value ranges from 1 to 3,600.
Command Mode	Global configuration mode
Usage Guide	If a dynamic multicast router port does not receive an MLD General Query packet or a PIM Hello packet before the timeout of its aging time, the device deletes the port from the multicast router port list. When the dynamic multicast router learning function is enabled, you can use this command to adjust the aging time of dynamic multicast router ports. If the aging time is too short, a multicast router port may be added and deleted frequently.

Configuring Aging Time of Dynamic Member Ports

Command	**ipv6 mld snooping host-aging-time** *seconds*
Parameter Description	***seconds***: Indicates the aging time.
Command Mode	Global configuration mode
Usage Guide	The aging time of a dynamic member port refers to the aging time set when a dynamic member port of a device receives from the host an MLD packet of joining a certain IPv6 multicast group. After receiving an MLD Join packet from a dynamic member port, the device resets the aging timer of the dynamic member port and sets the timer time to host-aging-time. If the timer times out, it is deemed that no user host receives multicast packets through this port, and then the multicast device deletes the port from the MLD Snooping member port list. After this command is configured, the aging timer value of dynamic member ports when MLD Join packets are received subsequently is host-aging-time. The aging time takes effect immediately after configuration and the timers of

	started member ports are updated.

Configuring Response Time of Query Packets

Command	**ipv6 mld snooping query-max-response-time** *seconds*
Parameter Description	***seconds***: Indicates the response time.
Command Mode	Global configuration mode
Usage Guide	After receiving an MLD General Query packet from a port, the multicast device resets the aging timers of all dynamic member ports and sets the timer time to query-max-response-time. If the timer times out, it is deemed that no user host receives multicast packets through the port, and then the multicast device deletes the port from the MLD Snooping member port list. After receiving an MLD Group-Specific Query packet from a port, the multicast device resets the aging timers of all dynamic member ports in the specific group and sets the timer time to query-max-response-time. If the timer times out, it is deemed that no user host receives multicast packets through the port, and then the multicast device deletes the port from the MLD Snooping member port list. The configuration takes effect when the a query packet is received next time, and the configuration of currently started timers are not updated. For Group-Specific Query packets of MLDv2, timers are not updated.

Checking Multicast Router Ports

Command	**show ipv6 mld snooping mroute**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, interface configuration mode
Usage Guide	If a multicast router port is successfully configured, the mark "S" is shown in the interface information displayed. For example: FS(config)#show ipv6 mld snooping mrouter Multicast Switching Mroute Port D: DYNAMIC S: STATIC (*, *, 1): VLAN(1) 1 MROUTES: GigabitEthernet 0/1(S)

Checking Dynamic Multicast Router Port Learning

Command	**show ipv6 mld snooping**
Parameter Description	N/A
Command	Privileged EXEC mode, global configuration mode, interface configuration mode

Mode	
Usage Guide	Run the **show ip igmp snooping** command to check the aging time and learning status of dynamic multicast router ports. Dynamic Mroute Aging Time : 300(Seconds) Multicast router learning mode: Enable

Checking Member Ports

Command	**show ipv6 mld snooping gda-table**
Parameter Description	-
Command Mode	Privileged EXEC mode, global configuration mode, interface configuration mode
Usage Guide	If a member port is successfully configured, the mark "S" is shown in the interface information displayed. For example: FS(config)#show ipv6 mld snooping gda-table Multicast Switching Cache Table D: DYNAMIC S: STATIC M: MROUTE (*, FF15::100, 1): VLAN(1) 2 OPORTS: GigabitEthernet 3/7(S)

Checking Other Parameters

Command	**show ipv6 mld snooping**
Parameter Description	-
Command Mode	Privileged EXEC mode, global configuration mode, interface configuration mode
Usage Guide	Run the **show ipv6 mld snooping** command to check the aging time of multicast router ports, aging time of dynamic member ports, response time of query packet, and Report packet suppression, and fast leave parameters. MLD-snooping mode: IVGL Source port check: Disable MLD Fast-Leave: Disable MLD Report suppress: Disable Query Max Response Time: 10 (Seconds) Dynamic Mroute Aging Time: 300(Seconds)

	Dynamic Host Aging Time: 260(Seconds)

9.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears MLD Snooping multicast forwarding entries.	**clear ipv6 mld snooping gda-table**
Clears MLD Snooping statistics.	**clear ipv6 mld snooping statistics**

Displaying

Description	Command
Displays the current MLD Snooping mode.	**show ipv6 mld snooping**
Displays MLD Snooping forwarding entries.	**show ipv6 mld snooping gda-table**
Displays MLD Snooping statistics.	**show ipv6 mld snooping statistics**
Displays MLD Snooping multicast router ports.	**show ipv6 mld snooping mrouter**
Displays MLD Snooping interface information, interface filtering profiles and maximum number of groups that a port can join.	**show ipv6 mld snooping interfaces** *interface-type interface-name*
Displays multicast information about a single VLAN, on which MLD Snooping is configured.	**show ipv6 mld snooping vlan** *vid*
Displays an MLD Profile.	**show ipv6 mld profile** *profile-number*

10 Configuring MSDP

10.1 Overview

Multicast Source Discovery Protocol is used to connect multiple rendezvous points (RPs) on the network and share the multicast source information among these RPs.

- Use MSDP among multiple Protocol Independent Multicast - Sparse-Mode (PIM-SM) domains to share the multicast source information of these PIM-SM domains to implement cross-domain multicast.
- Use MSDP in a PIM-SM domain to share the multicast source information of multiple RPs to implement anycast-RP.

Protocols and Standards

- RFC3618: Multicast Source Discovery Protocol(MSDP)

10.2 Applications

Application	Description
Cross-Domain Multicast	Connect multiple ASs, share the multicast resources among autonomous systems (ASs), and provide the multicast service across ASs.
Anycast-RP	Share the multicast source information among multiple RPs in a single AS.

10.2.1 Cross-Domain Multicast

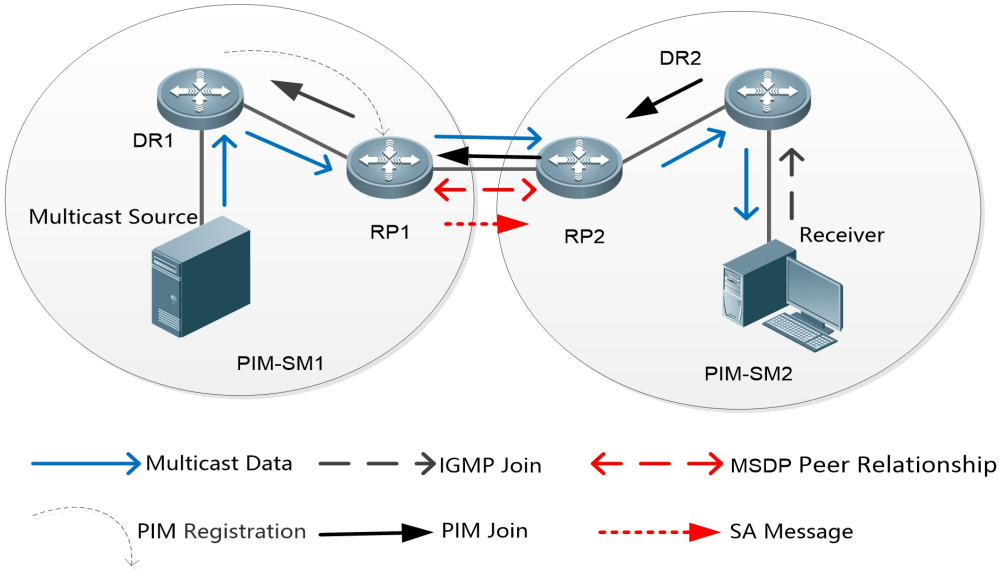
Scenario

Connect multiple ASs, run PPIM-SM within the ASs, and establish an MSDP peer relationship between RPs of different ASs.

As shown in Figure 10- 1, DR 1 connected to the multicast source registers with RP 1 in the local domain. DR 2 connected to the group member host triggers a join towards RP 2 in the local domain. RP 1 uses the SA message to notify RP 2 of the multicast source information. RP 2 continues to trigger a join towards the multicast source to build a multicast distribution tree (MDT).

Cross-domain multicast allows group member hosts to apply for the multicast streams across ASs.

Figure 10- 1



Deployment

- Run Open Shortest Path First (OSPF) within each AS, and run Border Gateway Protocol (BGP) between ASs to implement cross-domain unicast.
- Run PIM-SM within each AS, and run MSDP between ASs to implement cross-domain multicast.

10.2.2Anycast-RP

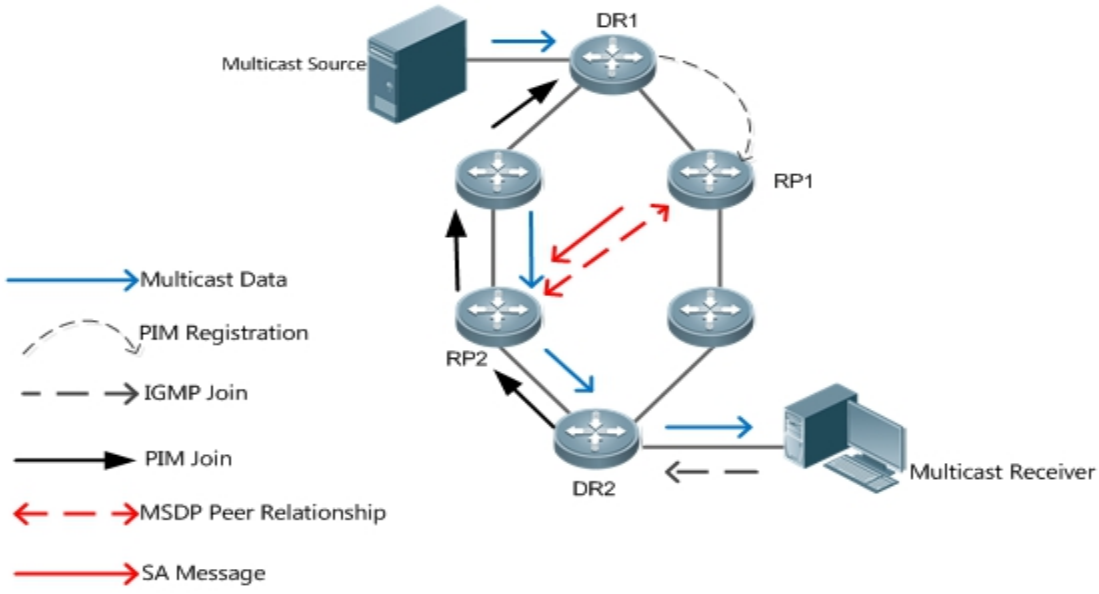
Scenario

PIM-SM runs within each AS. Multiple RPs exist, use the same RP address, and serve the same group. An MSDP peer relationship is established between these RPs.

As shown in Figure 10- 2, DR 1 connected to the multicast source registers with the nearest RP 1 in the local domain. DR 2 connected to the group member host triggers a join towards the nearest RP 2. RP 1 uses the SA message to notify RP 2 of the multicast source information. RP 2 continues to trigger a join towards the multicast source to build an MDT.

Anycast-RP provides redundancy and load balancing for RPs, and helps accelerate convergence of multicast routes.

Figure 10- 2



Deployment

- Run OSPF within each AS to implement intra-domain unicast.
- Run PIM-SM within each AS to implement intra-domain multicast.
- Run MSDP among RPs to share the multicast source information.

10.3 Features

Function	Description
Establishing an MSDP Peer Relationship	Connect multiple RPs to share the multicast source information.
Receiving and Forwarding SA Messages	Prevent SA flooding and suppress SA storms.

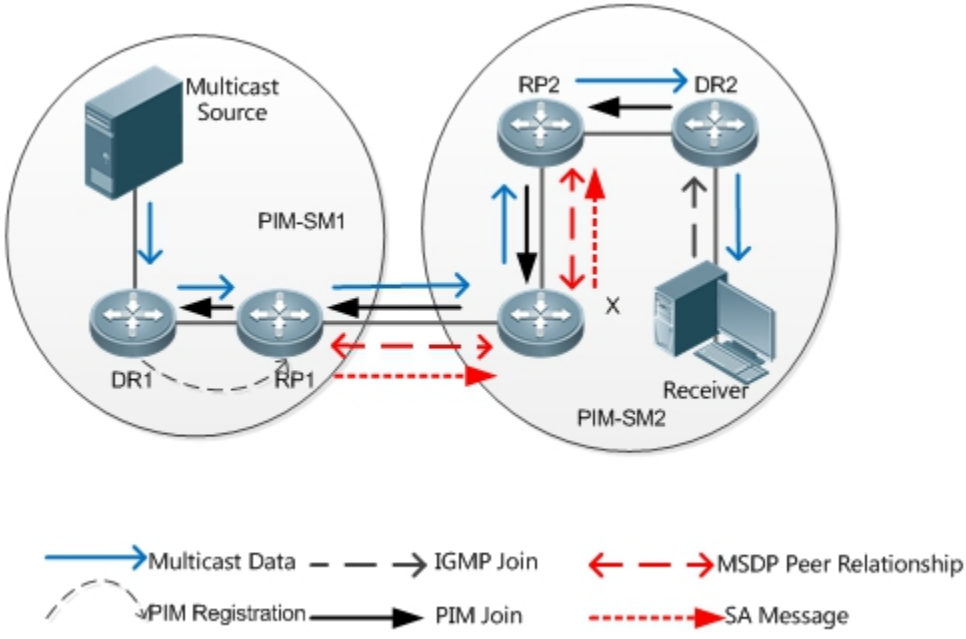
10.3.1 Establishing an MSDP Peer Relationship

Working Principle

Configure one or more pairs of MDSP peers on the network to connect RPs, thereby notifying other RPs of the multicast source information on an RP.

Use the TCP connection between MDSP peers through port 639. So far as the unicast route is reachable, the MSDP peer relationship can be established.

Figure 10- 3



RP Connected to the Multicast Source

Configure the MSDP peer on the RP connected to the multicast source. Then, this RP can use SA messages to send the local multicast source information to other RPs.

As shown in Figure 10- 3, DR 1 registers the multicast source information with RP 1. As a peer relationship is established between RP 1 and RP 2, RP 1 sends the multicast source information to X.

SA Message Forwarder

Non-RPs can also act as MSDP peers, but only forwards SA messages.

As shown in Figure 10- 3, X forwards SA messages sent from RP 1 to RP 2. In this way, the multicast source information is transferred to RP 2.

RP Connected to the Multicast Receiver

Configure the MSDP peer on the RP connected to the multicast receiver. Then, this RP can trigger a join towards the multicast source based on the received SA message.

As shown in Figure 10- 3, DR 2 triggers a join towards RP 2. As RP 2 already obtains the multicast source information, RP 2 continues to trigger a join towards the multicast source, thus establishing an MDT from DR 1 to DR 2.

10.3.2Receiving and Forwarding SA Messages

Working Principle

An SA message contains the multicast source address, multicast group address, and RP address. The RP address is the IP address of the RP with which the multicast source is registered.

- The RP encapsulates the locally registered multicast source information in an SA message, sends the message to all its MSDP peers.
- On receiving the SA message, each MSDP peer performs the Peer-RPF check, compares the SA-Cache, and matches the SA message against the SA incoming and outgoing filtering rules. If the SA message passes the Peer-RPF check, does not exist in the SA SA-Cache, and meets the outgoing filtering rules, this SA message is forwarded to other MSDP peers.

The SA request and SA response messages are also used between MSDP peers to transfer source information of a specific group.

Peer-RPF Check

Any SA message coming from an MSDP peer (address: N) will be checked as follows:

Judge whether the SA message passes the Peer-RPF check in the following sequence. Once the SA message passes the Peer-RPF check, accept the SA message; otherwise, drop the SA message.

4. If N is a member of the mesh group, the SA message passes the Peer-RPF check; otherwise, go to step 2.
5. If N is the only active MSDP peer on the local device, the SA message passes the Peer-RPF check; otherwise, go to step 3.
6. If N is the RP address in the SA message, the SA message passes the Peer-RPF check; otherwise, go to step 4.
7. If an EBGP route to the RP address in the SA message exists on the local device, and the next hop of this route is N, the SA message passes the Peer-RPF check; otherwise, go to step 5.
8. If an optimum route to the RP address in the SA message exists on the local device, check as follows:

If this optimum route is a distance vector route (such as the BGP/RIP route), and this router is advertised by N, the SA message passes the Peer-RPF check.

If this optimum route is a link status route (such as the OSPF/IS-IS route), and the next hop of this router is N, the SA message passes the Peer-RPF check.

Otherwise, go to step 6.

9. If an optimum route to the RP address in the SA message exists on the local device, and this route is a MBGP/BGP route, extract the nearest AS of the AS-Path of this MBGP/BGP route. If the local device has multiple MSDP peers in this AS and N is the MSDP peer with the largest IP address, or N is the only MSDP peer in this AS, the SA message passes the Peer-RPF check; otherwise, go to step 7.
10. If N is the default MSDP peer, the SA message passes the Peer-RPF check; otherwise, go to step 8.

11. The SA message fails in the Peer-RPF check.

The Peer-RPF check helps prevent loops and SA flooding.

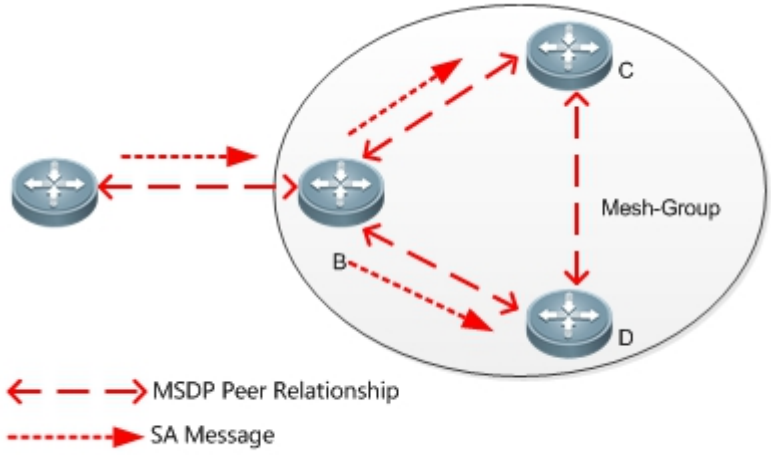
Mesh Group

In a mesh group, an MSDP peer relationship is established on every two members.

- For SA messages coming from entities outside the mesh group, after passing the Peer-RPF check and SA-Cache comparison, these SA messages are forwarded to other members in the group.
- Intra-group SA messages are no longer forwarded to other members in the group.

The mesh group helps reduce the number of SA messages.

Figure 10- 4



SA Cache

The SA cache is used to buffer the SA message status. Expired SA messages will be deleted.

When an MSDP peer receives an SA message, if this message does not exist in the SA cache and passes the Peer-RPF check, the message is stored in the SA cache. If this message already exists in the SA cache, the message is ignored. This helps suppress the SA storms.

When an MSDP peer receives an SA message, if this message already exists in the SA cache, the message is immediately responded. This helps improve the protocol efficiency.

10.4 Configuration

Configuration Item	Description and Command	
Configuring Cross-Domain Multicast	This configuration is mandatory in the cross-domain multicast scenario.	
	ip msdp peer *peer-address* **connect-source** *interface-type interface-number*	Establishes an MSDP peer relationship.
Configuring an Anycast-RP	This configuration is mandatory in the Anycast-RP scenario.	

Configuration Item	Description and Command	
	ip msdp peer *peer-address* **connect-source** *interface-type interface-number*	Establishes an MSDP peer relationship.
	ip msdp originator-id *interface-type interface-number*	Modifies the RP address in the SA message.
Configuring the Peer-RPF Check Green Channel	Optional. It is used to let SA message successfully pass the Peer-RPF check.	
	ip msdp default-peer *peer-address* [**prefix-list** *prefix-list-name*]	Configures the default MSDP peer.
	ip msdp mesh-group *mesh-name peer-address*	Configures an MSDP mesh group.
Enabling Security Measures	Optional. It is used to prevent illegal TCP connections and suppress SA storms.	
	ip msdp password peer *peer-address*[*encryption-type*] *string*	Enables TCP MD5 encryption.
	ip msdp sa-limit *peer-address sa-limit*	Limits the number of SA messages in the SA cache.
Restricting Broadcasting of SA Messages	Optional. It is used to restrict releasing, receiving, and forwarding of SA messages.	
	ip msdp redistribute [**list** *access-list*] [**route-map** *route-map*]	Filters the source information released locally.
	ip msdp filter-sa-request *peer-address*[**list** *access-list*]	Filters received SA requests.
	ip msdp sa-filter in *peer-address* [**list** *access-list*] [**route–map** *route-map*] [**rp-list** *rp-access-list*] [**rp-route-map** *rp-route-map*]	Filters received SA messages.
	ip msdp sa-filter out *peer-address* [**list***access-list*] [**route–map** *route-map*] [**rp-list** *rp-access-list*] [**rp-route-map** *rp-route-map*]	Filters sent SA messages.
Managing MSDP Peers	Optional. It is used to conveniently manage the MSDP peer relationship.	
	ip msdp description *peer-address text*	Adds a description to an MSDP peer.
	ip msdp shutdown *peer-address*	Shuts down an MSDP peer.
Modifying Protocol Parameters	Optional. You are advised not to modify the default values of protocol parameters.	
	ip msdp timer *interval*	Modifies the TCP reconnection interval.
	ip msdp ttl-threshold *peer-address ttl-value*	Modify the TTL value of the multicast data packet carried in the SA message.

10.4.1 Configuring Cross-Domain Multicast

Configuration Effect

Establish the MSDP peer relationship between multiple ASs so that group member hosts can apply for the multicast streams across ASs.

Notes

- The inter-AC unicast route must be reachable.
- Run PIM-SM within each AS, and configure the BSR border.

Configuration Steps

Establishing an MSDP Peer Relationship

- Mandatory.
- Establish a peer relationship between RPs of the corresponding multicast PIM domain.
- Establish an MSDP peer relationship between EBGP devices of different ASs.
- Establish an MSDP peer relationship between the RP and the EBGP device in each AS.

Command	**ip msdp peer** *peer-address* **connect-source** *interface-type interface-number*
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. *interface-type interface-number*: Indicates the local interface, which is used to establish a TCP connection with the remote peer.
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	The peer relationship is a bidirectional relationship. Therefore, this command must be configured on both sides. The IP address and local interface of the MSDP peer must be the same as those of the EBGP peer. To ensure that SA messages can successfully pass the Peer-RPF check, you are advised to: ● Configure a mesh group. ● Configure the default MSDP peer.

Verification

Send a packet from a source (S) close to an RP to the group (G), and enable a host close to another RP to join G.

- Verify that the host can receive the (S, G) packet.
- Run the **show ip msdp summary** command on an RP in another AS to display the status of the MSDP peer.
- Run the **show ip msdp sa-cache** command on an RP in another AS to display the learned MSDP source information.

Displaying the Learned MSDP Source Information

Command	**show ip msdp sa-cache**
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	If no address is specified, all the (S, G) information is displayed by default. If an address is specified, the device checks whether this address is a unicast or multicast address. If the address is a unicast address, this address is treated as the multicast source (S), and all (S, G) information in which the multicast source is S will be displayed. If the address is a multicast address, this address is treated as the multicast group (G), and all (S, G)

	information in which the multicast group is G will be displayed. If this address is neither a unicast or multicast address, no information is displayed. If two addresses are specified, one address is treated as the multicast source (S), and the other as the multicast group (G). If one address is the unicast address, and the other address is the multicast group address, no information is displayed.
	uijie# show ip msdp sa-cache MSDP Source-Active Cache: 2 entries (200.200.200.200, 227.1.2.2), RP: 20.20.20.20, (M)BGP/AS 100, 04:17:09/00:02:05, Peer 200.200.200.2 Learned from peer 200.200.200.2, RPF peer 200.200.200.2, SAs received: 277, Encapsulated data received: 0 (200.200.200.200, 227.1.2.3), RP: 20.20.20.20, (M)BGP/AS 100, 04:17:09/00:02:05, Peer 200.200.200.2 Learned from peer 200.200.200.2, RPF peer 200.200.200.2, SAs received: 277, Encapsulated data received: 0

Displaying the Brief MSDP Peer Information

Command	**show ip msdp summary**
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	N/A

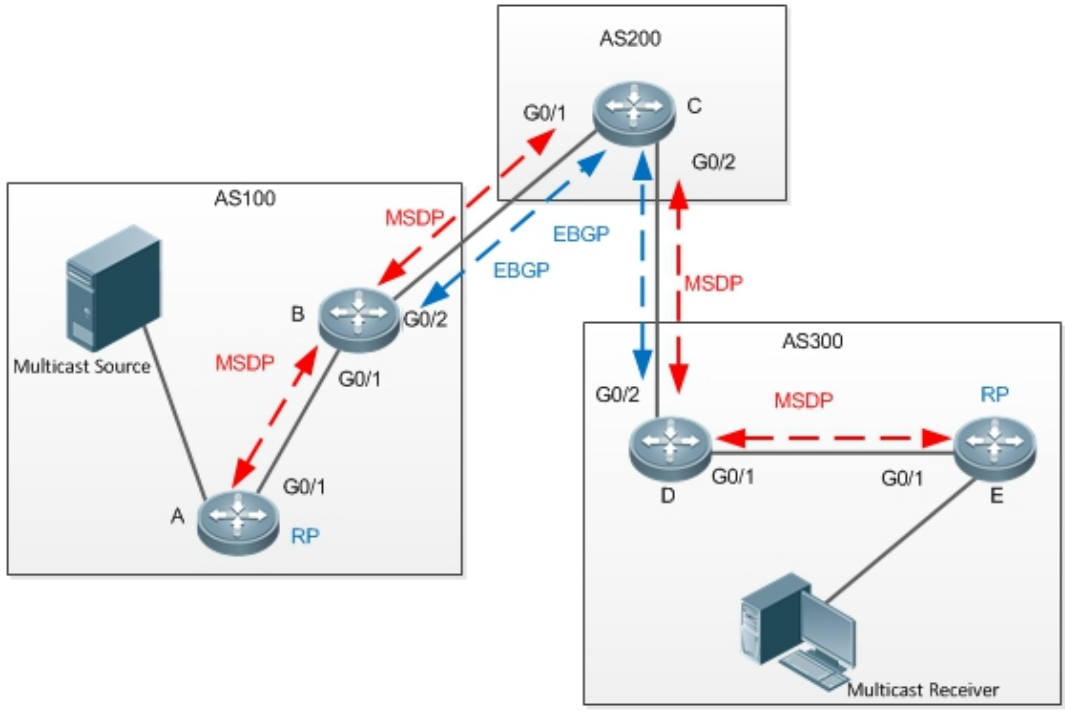
```
FS# show ip msdp summary
Msdp Peer Status Summary
Peer Address     As   State   Uptime/Downtime   Reset-Count   Sa-Count   Peer-description
200.200.200.2    100  Up      04:22:11          10            6616       No description
200.200.200.3    100  Down    19:17:13          4             0          peer-A
```

Configuration Example

Configuring Cross-Domain Multicast

Scenario

Figure 10- 5



The following table lists the interfaces and IP addresses of different devices:

Device	Interface	IP Address	Remark
A	G0/1	100.100.100.1/24	N/A
	Loopback0	10.10.10.10/32	RP address, which is used to establish an MSDP connection.
B	G0/1	100.100.100.2/24	N/A
	G0/2	1.1.1.1/24	BSR border
	Loopback0	20.20.20.20/32	Used to establish the EBGP and MSDP connections.
C	G0/1	1.1.1.2/24	BSR border
	G0/2	2.2.2.1/24	BSR border
	Loopback0	30.30.30.30/32	Used to establish the EBGP and MSDP connections.
D	G0/2	2.2.2.2/24	BSR border
	G0/1	3.3.3.1/24	N/A
	Loopback0	40.40.40.40/32	Used to establish the EBGP and MSDP connections.
E	G0/1	3.3.3.2/24	N/A
	Loopback0	50.50.50.50/32	RP address, which is used to establish an MSDP connection.

Configuration

- Configure IP addresses of interfaces.

Steps	● Enable OSPF in each AS. Set up an EBGP peer relationship between AS 200 and AS 100 and between AS 200 and AS 300. Introduce BGP and OSPF to each other. ● Enable PIM-SM in each AS, configure C-BSR and C-RP, and configure the BSR border. ● Establish the MSDP peer relationship between EBGP peers and between the RP and EBGP peers. The IP address and local interface of the MSDP peer must be the same as those of the EBGP peer.
A	A#configure terminal A(config)#ip multicast-routing A(config)#interface GigabitEthernet 0/1 A(config-if-GigabitEthernet 0/1)#ip pim sparse-mode A(config-if-GigabitEthernet 0/1)# exit A(config)#interface loopback 0 A(config-if-loopback 0)#ip pim sparse-mode A(config-if-loopback 0)# exit A(config)#ip pim rp-candidate loopback 0 A(config)#ip pim bsr-candidate loopback 0 A(config)#ip msdp peer 10.10.10.10 connect-source loopback 0
B	B#configure terminal B(config)#ip multicast-routing B(config)#interface GigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)#ip pim sparse-mode B(config-if-GigabitEthernet 0/1)# exit B(config)#interface GigabitEthernet 0/2 B(config-if-GigabitEthernet 0/2)#ip pim sparse-mode B(config-if-GigabitEthernet 0/2)#ip pim bsr-border B(config-if-GigabitEthernet 0/2)# exit B(config)#interface loopback 0 B(config-if-loopback 0)#ip pim sparse-mode B(config-if-loopback 0)# exit B(config)#ip msdp peer 10.10.10.10 connect-source loopback 0 B(config)#ip msdp peer 30.30.30.30 connect-source loopback 0
C	C#configure terminal C(config)#ip multicast-routing C(config)#interface GigabitEthernet 0/1

	C(config-if-GigabitEthernet 0/1)#ip pim sparse-mode C(config-if-GigabitEthernet 0/1)#ip pim bsr-border C(config-if-GigabitEthernet 0/1)# exit C(config)#interface GigabitEthernet 0/2 C(config-if-GigabitEthernet 0/2)#ip pim sparse-mode C(config-if-GigabitEthernet 0/2)#ip pim bsr-border C(config-if-GigabitEthernet 0/2)# exit C(config)#interface loopback 0 C(config-if-loopback 0)#ip pim sparse-mode C(config-if-loopback 0)# exit C(config)#ip msdp peer 20.20.20.20 connect-source loopback 0 C(config)#ip msdp peer 40.40.40.40 connect-source loopback 0
D	D#configure terminal D(config)#ip multicast-routing D(config)# ip pim ssmdefault D(config)#interface GigabitEthernet 0/1 D(config-if-GigabitEthernet 0/1)#ip pim sparse-mode D(config-if-GigabitEthernet 0/1)# exit D(config)#interface GigabitEthernet 0/2 D(config-if-GigabitEthernet 0/2)#ip pim sparse-mode D(config-if-GigabitEthernet 0/2)#ip pim bsr-border D(config-if-GigabitEthernet 0/2)# exit D(config)#interface loopback 0 D(config-if-loopback 0)#ip pim sparse-mode D(config-if-loopback 0)# exit D(config)#ip msdp peer 30.30.30.30 connect-source loopback 0 D(config)#ip msdp peer 50.50.50.50 connect-source loopback 0
E	E#configure terminal E(config)#ip multicast-routing E(config)#interface GigabitEthernet 0/1 E(config-if-GigabitEthernet 0/1)#ip pim sparse-mode E(config-if-GigabitEthernet 0/1)# exit E(config)#interface loopback 0

	E(config-if-loopback 0)#ip pim sparse-mode E(config-if-loopback 0)# exit E(config)#ip pim rp-candidate loopback 0 E(config)#ip pim bsr-candidate loopback 0 E(config)#ip msdp peer 50.50.50.50 connect-source loopback 0
Verification	Use the multicast source to send the packet (200.200.200.200,225.1.1.1), and enable the host to join the group 225.1.1.1. ● Verify that the host receives this packet. ● On device C, check the status and SA message of the MSDP peer.
D	D# show ip msdp summary Msdp Peer Status Summary Peer Address As State Uptime/Downtime Reset-Count SA-Count Peer-Description 30.30.30.30 200 Up 00:01:420 1 No description D# show ip msdp sa-cache MSDP Source-Active Cache: 1 entries (200.200.200.200,225.1.1.1),RP:10.10.10.10,(M)BGP/AS 100, 00:00:18/00:01:57, Peer 30.30.30.30 Learned from peer 30.30.30.30, RPF peer 30.30.30.30, SAs received: 1, Encapsulated data received: 1

Common Errors

- The BSR border is not configured, or is not configured on a correct interface.
- PIM-SM is not enabled on the local interface used to establish the MSDP peer connection or on the interface of the peer IP address.
- SA messages cannot pass the Peer-RPF check.

10.4.2 Configuring an Anycast-RP

Configuration Effect

Establish the MSDP peer relationship within an AS to provide redundancy and load balancing for RPs.

Notes

- The inter-AC unicast route must be reachable.
- PIM-SM must run within the AS, and multiple RPs using the same IP addresses must be configured.
- The C-RP and C-BSR cannot be configured on the same interface.

Configuration Steps

Establishing an MSDP Peer Relationship

- Mandatory.
- Configure the following command on each RP of the same AS to establish an MSDP peer relationship with each of other RPs:

Command	**ip msdp peer***peer-address***connect-source***interface-typeinterface-number*
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. *interface-type interface-number*: Indicates the local interface, which is used to establish a TCP connection with the remote peer.
Defaults	The MSDP peer relationship is not established.
Command Mode	Global configuration mode
Usage Guide	The peer relationship is a bidirectional relationship. Therefore, this command must be configured on both sides. To ensure that SA messages can successfully pass the Peer-RPF check, you are advised to configure a mesh group.

Modifying the RP Address in the SA Message

- Mandatory.
- Configure the following command on each RP of the same AS:

Command	**ip msdp originator-id** *interface-type interface-number*
Parameter Description	*interface-type interface-number*: Uses the IP address of this interface as the RP address in the SA message.
Defaults	By default, the RP address in the SA message is not modified.
Command Mode	Global configuration mode
Usage Guide	In the anycast-RP application scenario, the RP addresses on all RP devices are the same. If the RP address in an SA message is not modified, the RP device may determine that this SA message is sent by itself and therefore discards this message. Therefore, you need to configure different RP addresses for SA messages sent by different RP devices.

Verification

Send a packet from a source (S) close to an RP to the group (G), and enable a host close to another RP to join G.

- Verify that the host can receive the (S, G) packet.
- Run the **show ip msdp sa-cache** command on an RP in another AS to display the learned MSDP source information.

Displaying the Learned MSDP Source Information

Command	**show ip msdp sa-cache**
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	If no address is specified, all the (S, G) information is displayed by default. If an address is specified, the device checks whether this address is a unicast or multicast address. If the address is a unicast address, this address is treated as the multicast source (S), and all (S, G) information in which the multicast source is S will be displayed. If the address is a multicast address, this address is treated as the multicast group (G), and all (S, G)

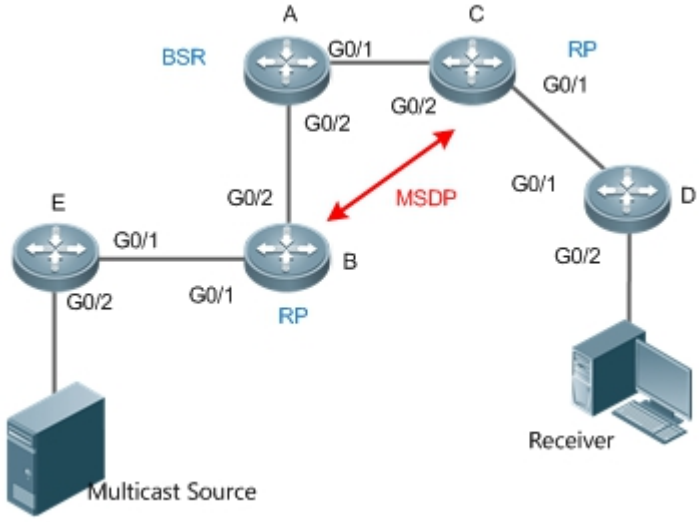
	information in which the multicast group is G will be displayed. If this address is neither a unicast nor multicast address, no information is displayed. If two addresses are specified, one address is treated as the multicast source (S), and the other as the multicast group (G). If one address is the unicast address, and the other address is the multicast group address, no information is displayed.
	FS# show ip msdp sa-cache MSDP Source-Active Cache: 2 entries (200.200.200.200, 227.1.2.2), RP: 20.20.20.20, (M)BGP/AS 100, 04:17:09/00:02:05, Peer 200.200.200.2 Learned from peer 200.200.200.2, RPF peer 200.200.200.2, SAs received: 277, Encapsulated data received: 0 (200.200.200.200, 227.1.2.3), RP: 20.20.20.20, (M)BGP/AS 100, 04:17:09/00:02:05, Peer 200.200.200.2 Learned from peer 200.200.200.2, RPF peer 200.200.200.2, SAs received: 277, Encapsulated data received: 0

Configuration Example

Sharing the Source information Among Anycast-RPs in the Same Multicast Domain

Scenario

Figure 10-6



The following table lists the interfaces and IP addresses of different devices:

Device	Interface	IP Address	Remark
A	G0/2	2.2.2.1/24	
	G0/1	1.1.1.1/24	
	Loopback0	100.100.100.100/32	The C-BSR is configured on this interface.
B	G0/2	2.2.2.2/24	
	G0/1	3.3.3.1/24	
	Loopback1	20.20.20.20/32	Used to establish an MSDP connection and modify the RP address in the SA message.
	Loopback0	10.10.10.10/32	The C-RP is configured on this interface.
C	G0/2	1.1.1.2/24	

	G0/1	4.4.4.1/24	
	Loopback1	30.30.30.30/32	Used to establish an MSDP connection and modify the RP address in the SA message.
	Loopback0	10.10.10.10/32	The C-RP is configured on this interface.
D	G0/1	4.4.4.2/24	
	G0/2	5.5.5.1/24	
E	G0/1	3.3.3.2/24	
	G0/2	6.6.6.1/24	

Configuration Steps

- Configure IP addresses of interfaces.
- Enable OSPF within the AS.
- Enable PIM-SM within the AS, and configure the C-BSR and C-RP.
- Establish the MSDP peer relationship between RPs, and modify the RP address in the SA message.
- Configure a mesh group.

A

```
A#configure terminal
A(config)#ip multicast-routing
A(config)#interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)#interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)#ip pim sparse-mode
A(config-if-GigabitEthernet 0/2)# exit
A(config)#interface loopback 0
A(config-if-loopback 0)#ip pim sparse-mode
A(config-if-loopback 0)# exit
A(config)#ip pim bsr-candidate loopback0
```

B

```
B#configure terminal
B(config)#ip multicast-routing
B(config)#interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)#ip pim sparse-mode
B(config-if-GigabitEthernet 0/1)# exit
B(config)#interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/2)#ip pim sparse-mode
B(config-if-GigabitEthernet 0/2)# exit
B(config)#interface loopback 0
```

	B(config-if-loopback 0)#ip pim sparse-mode B(config-if-loopback 0)# exit B(config)#interface loopback 1 B(config-if-loopback 1)#ip pim sparse-mode B(config-if-loopback 1)# exit B(config)#ip pim rp-candidate loopback 0 B(config)#ip msdp peer 30.30.30.30 connect-source loopback 1 B(config)# ip msdp originator-id loopback 1 B(config)#ip msdp mesh-group mesh-name 30.30.30.30
C	C#configure terminal C(config)#ip multicast-routing C(config)#interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)#ip pim sparse-mode C(config-if-GigabitEthernet 0/1)# exit C(config)#interface GigabitEthernet 0/2 C(config-if-GigabitEthernet 0/2)#ip pim sparse-mode C(config-if-GigabitEthernet 0/2)# exit C(config)#interface loopback 0 C(config-if-loopback 0)#ip pim sparse-mode C(config-if-loopback 0)# exit C(config)#interface loopback 1 C(config-if-loopback 1)#ip pim sparse-mode C(config-if-loopback 1)# exit C(config)#ip pim rp-candidate loopback 0 C(config)#ip msdp peer 20.20.20.20 connect-source loopback 1 C(config)# ip msdp originator-id loopback 1 C(config)#ip msdp mesh-group mesh-name 20.20.20.20
D	D#configure terminal D(config)#ip multicast-routing D(config)#interface GigabitEthernet 0/1 D(config-if-GigabitEthernet 0/1)#ip pim sparse-mode D(config-if-GigabitEthernet 0/1)# exit D(config)#interface GigabitEthernet 0/2

	D(config-if-GigabitEthernet 0/2)#ip pim sparse-mode D(config-if-GigabitEthernet 0/2)# exit
E	E#configure terminal E(config)#ip multicast-routing E(config)#interface GigabitEthernet 0/1 E(config-if-GigabitEthernet 0/1)#ip pim sparse-mode E(config-if-GigabitEthernet 0/1)# exit E(config)#interface GigabitEthernet 0/2 E(config-if-GigabitEthernet 0/2)#ip pim sparse-mode E(config-if-GigabitEthernet 0/2)# exit
Verification	Use the multicast source to send the packet (6.6.6.6,225.1.1.1), and enable the host to join the group 225.1.1.1. ● Verify that the host receives this packet. ● On device C, check the status and SA message of the MSDP peer.
C	C# show ip msdp summary Msdp Peer Status Summary Peer Address As State Uptime/Downtime Reset-Count SA-Count Peer-Description 20.20.20.20 Unknown Up 00:01:420 1 No description C# show ip msdp sa-cache MSDP Source-Active Cache: 1 entries (6.6.6.6,225.1.1.1),RP:10.10.10.10,(M)BGP/AS unknown, 00:00:18/00:01:57, Peer 20.20.20.20 Learned from peer 20.20.20.20, RPF peer 20.20.20.20,

Common Errors

- The C-BSR and C-RP are configured on the same interface.
- The RP address in the SA message is not modified.
- SA messages cannot pass the Peer-RPF check.

10.4.3 Configuring the Peer-RPF Check Green Channel

Configuration Effect

Configure the Peer-RPF check green channel so that all SA messages sent from a specified MSDP peer can pass the Peer-RPF check.

Configure an MSDP mesh group so that all SA messages sent from members of the mesh group can pass the Peer-RPF check.

Notes

- The MSDP peer relationship must be established between devices.

Configuration Steps

Configuring the Default MSDP Peer

- Optional.
- On an MSDP peer, if it is not necessary to perform the Peer-RPF check on SA messages sent from a specified peer, configure this peer as the default peer.

Command	**ip msdp default-peer***peer-address* [**prefix-list** *prefix-list-name*]
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. **prefix-list***prefix-list-name*: Specifies the prefix list, which is used to limit the RPs initiating SA messages.
Defaults	By default, no default peer is configured.
Command Mode	Global configuration mode
Usage Guide	If the command does not contain **prefix-list***prefix-list-name*, all SA messages are accepted. If the command contains **prefix-list***prefix-list-name* but the specified prefix list does not exist, all SA messages are accepted. If the command contains **prefix-list***prefix-list-name*, and the specified prefix list exists, only the SA messages initiated by RPs specified in this prefix list are accepted.

Creating a Mesh Group

- Optional.
- Among multiple MSDP peers, if SA messages coming from any of these peers pass the Peer-RPF check by default, you can add these peers to a mesh group.

Command	**ip msdp mesh-group***mesh-namepeer-address*
Parameter Description	*mesh-name*: Indicates the name of the mesh group. The name is case sensitive. *peer-address*: Indicates the IP address of the MSDP peer to be added to the mesh group.
Defaults	By default, no mesh group is configured.
Command Mode	Global configuration mode
Usage Guide	An MSDP peer relationship must be established between every two MSDP peers added to the same mesh group. All SA messages sent by members of the mesh group can pass the Peer-RPF check.

Verification

- Check whether SA messages sent by the default peer can pass the Peer-RPF check.
- Check the configuration of the mesh group, and check whether all SA messages sent by members of the mesh group can pass the Peer-RPF check.

Displaying Information about the Peer-RPF Check of a Specified MSDP Peer

Command	**show ip msdp rpf-peer***ip-address*

Parameter Description	*peer-address*: Indicates the IP address of the SA message initiator.
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	N/A
	FS# show ip msdp rpf-peer 1.1.1.1 RPF peer information for 1.1.1.1 RPF peer: 200.200.200.2 RPF rule: **Peer is only active peer** RPF route/mask: Not-used RPF type: Not-used

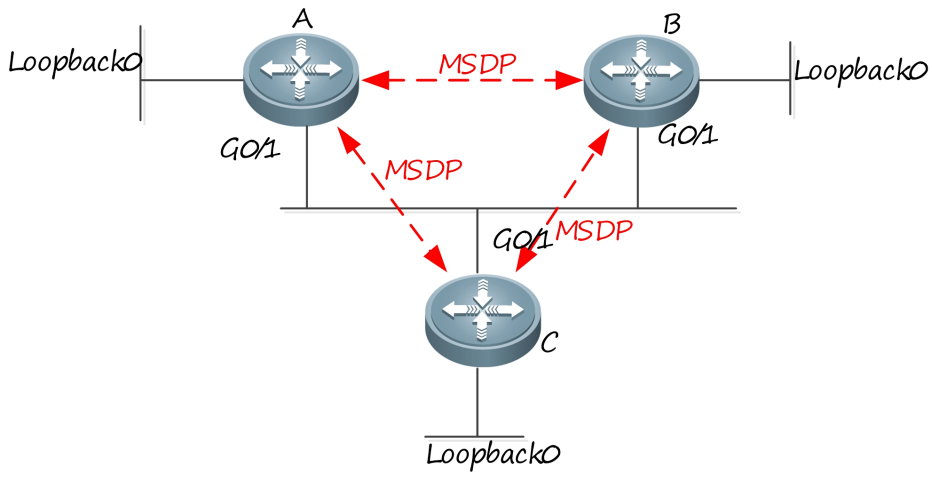
Displaying the Mesh Group Configuration

Command	**show ip msdp mesh-group**
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	N/A
	FS# **show ip msdp mesh-group** MSDP peers in each Mesh-group, <Mesh-group name>:<# peers> msdp-mesh: 1.1.1.2 1.1.1.3

Configuration Example

Configuring the Peer-RPF Check and a Mesh Group

Figure 10-7



The following table lists the interfaces and IP addresses of different devices:

Device	Interface	IP Address	Remark
A	G0/1	20.0.0.3/24	
	Loopback0	10.1.1.1/24	
B	G0/1	20.0.0.4/24	
	Loopback0	40.0.0.1/24	
	G0/1	20.0.0.222/24	
	Loopback0	30.0.0.2/24	

Configuration Steps

- Configure IP addresses of interfaces.
- Enable OSPF within the AS.
- Establish the MSDP peer relationship between A and B and between A and C.
- Enable PIM-SM on the G0/1 interface of device C.
- Before configuration, there are two active MSDP peers on device A, but it is not known which one should be selected as the RPF peer. Therefore, display the RPF peer information. "RPF peer does not exist" is displayed.
- Configure the default MSDP peer, and check whether the configuration is successful.
- Configure a mesh group.

A

```
A#configure terminal
A(config)#ip msdp peer 20.0.0.4 connect-source gi0/1
A(config)#ip msdp peer 30.0.0.2 connect-source loopback 0
```

B

```
B#configure terminal
B(config)#ip msdp peer 20.0.0.3 connect-source gi0/1
```

C

```
C#configure terminal
C(config)#ip msdp peer 10.0.0.1 connect-source loopback 0
C(config)#interface GigabitEthernet 0/1
C(config-if-GigabitEthernet 0/1)#ip pim sparse-mode
```

	C(config-if-GigabitEthernet 0/1)# exit
	● Before configuration, there are two active MSDP peers on device A, but it is not known which one should be selected as the RPF peer. Therefore, display the RPF peer information. "RPF peer does not exist" is displayed. ● Configure the default MSDP peer. Then, display the RPF peer information. " Peer is best default peer" is displayed.
A	A#configure terminal A(config)#ip msdp default-peer 30.0.0.2 ● Cancel the default peer, and send the multicast source information to device C. Information is displayed on device A, indicating that the SA message is received, but does not pass the Peer-RPF check. ● On device A, add 30.0.0.2 to the mesh group. Then, device A can receive the SA message normally.
A	A#configure terminal A(config)#no ip msdp default-peer 30.0.0.2
A	A#configure terminal A(config)#ip msdp mesh-group first 30.0.0.2
Verification	N/A

10.4.4 Enabling Security Measures

Configuration Effect

Enable MD5 encryption on TCP connections between MSDP peers to prevent illegal TCP connections.

Limit the number of SA messages in the SA cache of a specified MSDP peer to suppress SA storms.

Notes

- The MSDP peer relationship must be established between devices.

Configuration Steps

Configuring MD5 Encryption on TCP Connections Between MSDP Peers

- Optional.
- Configure consistent MD5 encryption on MSDP peers that require encryption.

Command	**ip msdp password peer** *peer-address* [*encryption-type*] *string*
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. *encryption-type*: Indicates the encryption level. Currently, only levels 0 to 7 are supported. 0 is the lowest level, and 7 is the highest level. The default value is 0. *string*: Indicates the cipher used for TCP MD5 authentication.
Defaults	By default, MD5 encryption is not configured.
Command Mode	Global configuration mode
Usage Guide	To authenticate the ID of an MSDP peer, enable MD5 encryption on the TCP connection established with this MSDP peer.

	The MSDP peer must have the consistent configuration, and the cipher must be the same; otherwise, the connection fails. If the configuration or cipher changes, the local device does not stop the current session, and will attempt to use a new cipher to retain the current session until timeout. If the encryption level is set to 7, the cipher text length must be an even number equaling to or greater than 4; otherwise, the configuration fails.

Limiting the Number of SA Messages in the SA Cache of a Specified MSDP Peer

- Optional.
- Perform this configuration if you need to limit the number of SA messages in the SA cache of a specified MSDP peer.

Command	**ip msdp sa-limit** *peer-address sa-limit*
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. *sa-limit*: Indicates the maximum number of SA messages in the SA cache.
Defaults	The default value is 1,024.
Command Mode	Global configuration mode
Usage Guide	An MSDP peer relationship must be established between every two MSDP peers added to the same mesh group. Assume that the number of SA messages in the SA cache already exceeds the limit. After the configuration is completed, the number of SA messages in the SA cache does not exceed the limit.

Verification

- Check the connection between peers on which MD5 encryption is configured.
- Send a number of source information packets that exceeds the limit to the peer where the maximum number of SA messages in the SA cache is configured. Check whether all the source information can be learned.

Displaying the Number of SA Messages Learned from a Specified Peer

Command	**show ip msdp count**
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	N/A
	FS# **show ip msdp count** SA State per Peer Counters, <Peer>: <# SA learned> 1.1.1.2 : 0 100.100.100.14 : 0 100.100.100.15 : 0 100.100.100.200: 0

	200.200.200.2 : 2 200.200.200.3 : 0 200.200.200.6 : 0 200.200.200.13 : 0 200.200.200.66 : 0 SA State per ASN Counters, <asn>: <# sources>/<# groups> Total entries: 2 100: 1/2

Configuration Example

Configuring MD5 Encryption on an MSDP Peer and Limiting the Number of SA Messages Sent by This MSDP Peer in the SA Cache

Scenario **Figure 10-8**	
Configuration Steps	● Establish an MSDP peer relationship between A and B. ● Configure MD5 encryption on device A. ● After MSDP timeout, configure the MD5 cipher of the peer on device B, which is the same as the cipher on device A. Then, the session is reconnected. ● On device A, set the maximum number of SA messages sent by the peer 20.0.0.4 in the SA cache to 10.
A	A#configure A(config)# ip msdp password peer 20.0.0.4 0 1234567 A(config)# ip msdp sa-limit 20.0.0.4 10
B	B#configure B(config)# ip msdp password peer 20.0.0.4 0 1234567
Verification	● After MD5 is configured on device A, but is not configured on device B, a message will be displayed, indicating the MD5 encryption failure. At this time, the MSDP peer is in DOWN state. ● A period of time after MD5 is configured on device B, the MSDP peer is in DOWN state. ● Send 20 multicast source packets to device B. A message will be displayed on device A, indicating that the number of SA messages exceeds the limit.
A	A# debug ip msdp sa-cache A# show ip msdp count

10.4.5 Restricting Broadcasting of SA Messages

Configuration Effect

Configure the SA message filtering rules to restricting broadcasting of SA messages.

Notes

- The MSDP peer relationship must be established between devices.

Configuration Steps

Filtering the Source Information Released Locally

- Optional.
- Configure the SA release filtering rule on an MSDP device where releasing of the SA information needs to be limited.

Command	**ip msdp redistribute** [**list** *access-list*] [**route-map** *route-map*]
Parameter Description	**list** *access-list*: Indicates the access control list (ACL) used to control the ranges of S and G. **route-map** *route-map*: Indicates the route map used to control the ranges of S and G.
Defaults	By default, no rule is configured to filter locally released SA information.
Command Mode	Global configuration mode
Usage Guide	After this command is configured, only the accepted (S, G) information (either coming from the local domain or other domains) can be injected to the MSDP. If the command contains **list** *access-list*, only the (S, G) information matching this ACL can be released. If the command contains **route-map** *route-map*, only the (S, G) information matching this route map can be released. If the command contains both parameters, only the (S, G) information matching the ACL and route map can be released. If the command does not contain any parameter, no (S, G) information is released.

Filtering Received SA Requests

- Optional.
- Perform this configuration on the MSDP device where responding to the SA requests needs to be limited.

Command	**ip msdp filter-sa-request** *peer-address*[**list** *access-list*]
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. **list** *access-list*: Indicates the ACL used to control the range of the group address.
Defaults	By default, no rule is configured to filter received SA requests.
Command Mode	Global configuration mode
Usage Guide	Use this command if you need to control the SA requests that can be accepted and responded. If the command does not contain **list** *access-list*, all SA requests will be ignored. If the command contains **list** *access-list*, but this AC does not exist, all SA requests will be ignored. If the command contains **list** *access-list*, and this AC exists, only the SA requests allowed by the ACL will be accepted, and others are ignored.

Filtering Received SA Messages

- Optional.
- Perform this configuration on an MSDP device where the incoming SA information needs to be limited.

Command	**ip msdp sa-filter in** *peer-address* [**list** *access-list*] [**route–map** *route-map*] [**rp-list** *rp-access-list*] [**rp-route-map** *rp-route-map*]
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. **list** *access-list*: Indicates the number or name of the extended IP ACL of a specified (S, G). It is used to control the multicast source information (S, G) that is allowed to pass. **route-map** *route-map*: Indicates the name of the route map of the specified (S, G). The multicast source information (S, G) is allowed to pass only when the AS path of the route on the S matches the AS path in the route map. **rp-list** *rp-access-list*: Indicates the number or name of the standard ACL of a specified RP. It is used to control the RPs, of which the multicast source information (S, G) that is allowed to pass. **rp-route-map** *rp-route-map*: Indicates the name of the route map of a specified RP. The multicast source information (S, G) is allowed to pass only when the AS path of the route on the RP matches the AS path in the route map.
Defaults	By default, no rule is configured to filter incoming SA messages.
Command Mode	Global configuration mode
Usage Guide	If this command is configured, but no ACL or route map is specified, all incoming SA messages will be filtered. If only one keyword (**list** or **route-map**) is specified, and every multicast source record (S, G) in the SA message meets the rule specified by the keyword, the multicast source record (S, G) will be received. If either **rp-list** or **rp-route-map**) is specified, and the RP address contained in the SA message meets the rule specified by this keyword, this SA message will be received. If two or more of the keywords (including **list**, **route-map**, **rp-list**, and **rp-route-map**) are specified, only multicast source record (S, G) in the SA message that meets the rules specified by all the available keywords can be received.

Filtering Sent SA Messages

- Optional.
- Perform this configuration on an MSDP device where the outgoing SA information needs to be limited.

Command	**ip msdp sa-filter out** *peer-address* [**list***access-list*] [**route–map** *route-map*] [**rp-list** *rp-access-list*] [**rp-route-map** *rp-route-map*]
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. **list** *access-list*: Indicates the number or name of the extended IP ACL of the specified (S, G). It is used to control the multicast source information (S, G) that is allowed to pass. **route-map** *route-map*: Indicates the name of the route map of the specified (S, G). The multicast source information (S, G) is allowed to pass only when the AS path of the route on the S matches the AS path in the route map. **rp-list** *rp-access-list*: Indicates the number or name of the standard ACL of a specified RP. It is used to control the RPs, of which the multicast source information (S, G) that is allowed to pass. **rp-route-map** *rp-route-map*: Indicates the name of the route map of a specified RP. The multicast source information (S, G) is allowed to pass only when the AS path of the route on the RP matches the AS path in the route map.
Defaults	By default, no rule is configured to filter outgoing SA messages.
Command Mode	Global configuration mode

Usage Guide	If this command is configured, but no ACL or route map is specified, no SA message will be sent to this MSDP peer. If only one of the keywords (including **list**, **route-map**, **rp-list**, and **rp-route-map**) is specified, any multicast source record (S, G) that meets the rule specified by the keyword will be forwarded to this MSDP peer. If two or more of the keywords (including **list**, **route-map**, **rp-list**, and **rp-route-map**) is specified, any multicast source record (S, G) that meets the rules specified by all the available keywords will be forwarded to this MSDP peer.

Verification

- Check whether SA messages initiated by the local device meet the filtering rules.
- Check whether SA messages learned by the local device meet the filtering rules.

Displaying SA Messages Initiated by the Local Device

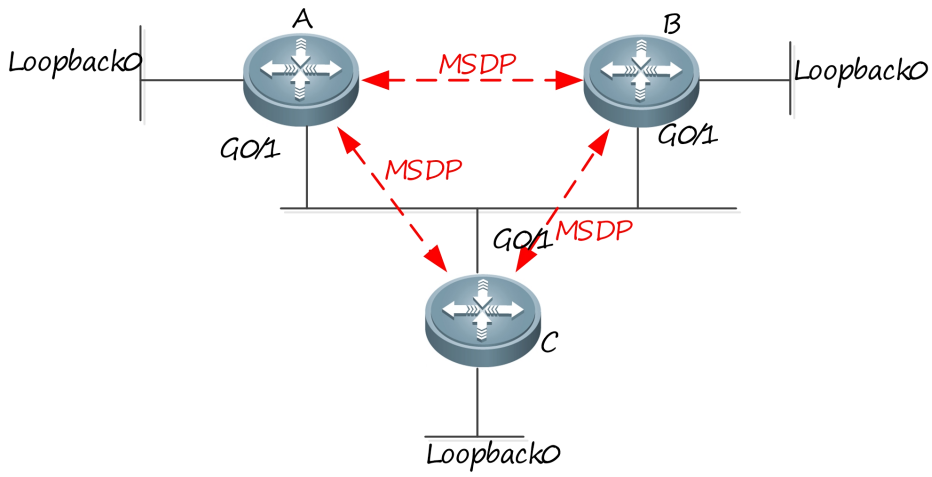
Command	**show ip msdp sa-originated**
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	If the local device is the RP of PIM-SM, multicast source (S, G) information is registered on the RP, and the MSDP peer is configured on the local device, you can run this command to display the (S, G) information initiated by the local device. The (S, G) information displayed by this command has met the criteria specified by the redistribution command **ip msdp redistribute**), but such (S, G) information can be sent to the MSDP peer only when the information meets the outgoing SA information filtering rules specified by the **ip msdp sa-filter out** command.
	FS# show ip msdp sa-originated MSDP Source-Active Originated: 5 entries (192.168.23.78, 225.0.0.1), RP: 192.168.23.249 (192.168.23.79, 225.0.0.2), RP: 192.168.23.249 (192.168.23.80, 225.0.0.3), RP: 192.168.23.249 (192.168.23.81, 225.0.0.4), RP: 192.168.23.249 (192.168.23.82, 225.0.0.5), RP: 192.168.23.249

Configuration Example

Configuring Rules for Filtering Incoming or Outgoing SA Messages

Scenario

Figure 10-9



The following table lists the interfaces and IP addresses of different devices:

Device	Interface	IP Address	Remark
A	G0/1	20.0.0.3/24	
	Loopback0	10.1.1.1/24	
B	G0/1	20.0.0.4/24	
	Loopback0	40.0.0.1/24	
	G0/1	20.0.0.222/24	
	Loopback0	30.0.0.2/24	

Configuration Steps

- Complete the basic configuration, as described in section 10.4.3 "Configuring the Peer-RPF Check Green Channel".
- Configure rules for filtering incoming SA messages on device A.
- Configure rules for filtering outgoing SA messages on device A.
- Send the multicast source information to device C.

A

```
A#configure
A(config)# ip msdp sa-filter in 30.0.0.2
A(config)# ip msdp sa-filter in 30.0.0.2 list 100
A(config)# ip access-list extended 100
A(config-ext-nacl)# permit ip    host 20.0.0.100 host 225.0.0.1
A(config)# ip msdp sa-filter in 30.0.0.2 rp-list rp-acl-1
A(config)# ip access-list standard rp-acl-1
A(config-std-nacl) # permit host 20.0.0.221
A(config)# ip msdp sa-filter in 30.0.0.2 rp-route-map rp-rm-1
A(config)# route-map rp-rm-1
A(config-route-map)#match as-path 1
A(config)# ip as-path access-list 1 permit 2
A#configure
A(config)# ip msdp sa-filter out 30.0.0.2
```

	A(config)# ip msdp sa-filter out 30.0.0.2 list 101 A(config)# ip access-list extended 101 A(config-ext-nacl)# permit ip host 20.0.0.100 host 225.0.0.1 A(config)# ip msdp sa-filter out 30.0.0.2 rp-list rp-acl-2 A(config)# ip access-list standard rp-acl-2 A(config-std-nacl) # permit host 20.0.0.221 A(config)# ip msdp sa-filter out 30.0.0.2 rp-route-map rp-rm-2 A(config)# route-map rp-rm-1 A(config-route-map)#match as-path 1 A(config)# ip as-path access-list 1 permit 2
Verification	● Send the multicast source information to device C in various scenarios. ● On device A, check whether the learned multicast source information meets the incoming requirements. ● On device B, check whether the learned multicast source information meets the outgoing requirements.
A	A#show ip msdp sa-cache
B	B#show ip msdp sa-cache
C	B#show ip msdp sa-originated

10.4.6 Managing MSDP Peers

Configuration Effect

Manage MSDP peers by adding descriptions to a specified MSDP or reset an MSDP peer.

Notes

- MSDP peers must be created in advance.

Configuration Steps

Configuring the Description for an MSDP Peer

- Optional.
- Perform this configuration on an MSDP peer that should be managed.

Command	**ip msdp description** *peer-address text*
Parameter Description	*peer-address*: Indicates the IP address of a remote peer. *text*: Indicates the string that describes the MSDP peer.
Defaults	By default, no description information is configured of an MSDP peer.
Command Mode	Global configuration mode
Usage Guide	N/A

Shutting Down an MSDP Peer

- Optional.
- Perform this configuration when it is required to temporarily shut down the connection with a specified peer.

Command	**ip msdp shutdown** *peer-address*
Parameter Description	*peer-address*: Indicates the IP address of an MSDP peer.
Defaults	By default, an MSDP peer is not shut down.
Command Mode	Global configuration mode
Usage Guide	This command shuts down only the TCP connection with an MSDP peer, but does not delete this MSDP peer or configuration of this MSDP peer.

Verification

- Display information about a specified MSDP peer, and check whether the description and peer status meet the requirements.

Displaying Information about a Specified MSDP Peer

Command	**show ip msdp peer** [*peer-address*]
Parameter Description	N/A
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	N/A

```
FS#show ip msdp peer 20.0.0.1
MSDP PEER 20.0.0.1 (No description), AS unknown
  Connection status:
    State: Listen, Resets: 1, Connection source: GigabitEthernet 0/1 (20.0.0.2)
    Uptime(Downtime): 00:00:25, Message sent/received: 13/19
    Input messages discarded: 0
    Connection and counters cleared 00:13:25 ago
    Local Address of connection: 20.0.0.2
    MD5 signature protection on MSDP TCP connection: enabled
  SA Filtering:
    Input (S,G) Access-list filter: None
    Input (S,G) route-map filter: None
    Input RP Access-list filter: None
    Input RP Route-map filter: None
    Output (S,G) Access-list filter: None
    Output (S,G) Route-map filter: None
    Output RP Access-list filter: None
    Output RP Route-map filter: None
  SA-Requests:
    Input filter: None
  Peer ttl threshold: 0
  SAs learned from this peer: 2, SAs limit: No-limit
  Message counters:
    SA messages discarded: 0
    SA messages in/out: 13/0
    SA Requests discarded/in: 0/0
    SA Responses out: 0
    Data Packets in/out: 6/0
```

Configuration Example

Configuring the Description of an MSDP Peer and Shutting Down the Connection with This Peer

Scenario **Figure 10- 10**	
Configuration Steps	● Establish the MSDP peer relationship between device A and device B. ● Configure the description "peer-router-B" for the peer 20.0.0.4 on device A. ● Wait 60, and shut down the connection with the MSDP peer 20.0.0.4 on device A.
A	A#configure A(config)# ip msdp peer 20.0.0.4 connect-source gi0/1 A(config)# ip msdp description 20.0.0.4 peer-router-B A(config)# end A# show ip msdp peer 20.0.0.4 …… A#configure A(config)# ip msdp shutdown 20.0.0.4 A(config)# show ip msdp peer 20.0.0.4
B	B# configure B(config)# ip msdp peer 20.0.0.3 connect-source gi0/1 B(config)# end
Verification	● Run the **show ip msdp peer** [*peer-address*] command to display the brief information of a specified peer, including the description and connection status of this MSDP peer.
A	A# show ip msdp peer 20.0.0.4

10.4.7 Modifying Protocol Parameters

Configuration Effect

Manage MSDP peers by adding descriptions to a specified MSDP or reset an MSDP peer.

Notes

- MSDP peers must be created in advance.

Configuration Steps

Configuring the TCP Reconnection Interval of an MSDP Peer

- Optional.
- Perform this configuration on the device where the TCP reconnection interval of an MSDP peer needs to be modified.

Command	**ip msdp timer** *interval*
Parameter Description	*interval*: Indicates the TCP reconnection interval. The unit is second. The value ranges from 1 to 60. The default value is 30.
Defaults	By default, the reconnection interval is 30s.
Command Mode	Global configuration mode
Usage Guide	Within the TCP reconnection interval, the MSDP peer on the proactive connection side can initiate at most one TCP connection. In some application scenarios, you can shorten the TCP reconnection interval to accelerate convergence of the MSDP peer relationship.

Configuring the TTL of the Multicast Packet Contained in the SA Message

- Optional.
- Perform this configuration on the MSDP device where inter-RP transfer of multicast packets should be restricted.

Command	**ip msdp ttl-threshold***peer-addressttl-value*
Parameter Description	*peer-address*: Indicates the IP address of an MSDP peer. *peer-addressttl-value*: Indicates the TTL value. The value ranges from 0 to 255. The default value is 0.
Defaults	By default, the TTL value of the multicast packet contained in the SA message is not restricted.
Command Mode	Global configuration mode
Usage Guide	This command restricts the sending of multicast packet encapsulated in the SA message. A multicast packet is sent to the MSDP peer only when the TTL value in the IP header of the multicast packet is equal to or greater than the preset TTL threshold. If the the TTL value in the IP header of the multicast packet is smaller than the preset TTL threshold, the multicast packet will be removed from the SA message and discarded before the SA message is sent to the MSDP peer. This command affects the sending of multicast packet in the SA message, but does not affect the sending of the multicast source information (S, G) in the SA message.

Configuring the MSDP Peer Capacity Supported by a Device

- Optional.
- If the default capacity (64 MSDP peers) is insufficient to support applications, you can modify the capacity on the device.

Command	**ip msdp peer-limit** *peer-limit*
Parameter Description	*peer-limit*: Indicates the maximum number of MSDP peers that can be configured. The value ranges from 1 to 128. The default value is 64.
Defaults	By default, at most 64 peers can be configured.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the maximum number of MSDP peers supported by a device. When configuring this command, if the number of MSDP peers on the device exceeds the value to be configured, a prompt will be displayed, and the configuration fails. The configuration can succeed only after the extra number peers are deleted.

Configuring the SA Cache Capacity Supported by a Device

- Optional.
- Perform this configuration on a device where the SA cache capacity should be adjusted.

Command	**ip msdp global-sa-limit***sa-liit*
Parameter Description	*sa-limit*: Indicates the maximum capacity of the SA cache supported by the device. The value ranges from 1 to 4,096. The default value is 1,024.
Defaults	By default, the SA cache supports 1,024 SA messages.
Command Mode	Global configuration mode
Usage Guide	This command is used to adjust the SA cache capacity of the device. You are advised to configure this command when the device is being started. If the capacity is increased when MSDP is in service, the adjustment does not affect the SA cache that is originally learned. If the capacity is increased when MSDP is in service, all SA caches that are originally learned from other devices or the SA caches initiated by the local devices must be deleted and re-learned.

Verification

- Shut down the connection with an MSDP peer. After the reconnection interval elapses, check whether the MSDP peer is in UP date again.

Configuration Example

Setting the MSDP Peer Reconnection Interval to 20s

Scenario **Figure 10- 11**	
Configuration Steps	- Establish the MSDP peer relationship between device A and device B. - On device A, set the MSDP peer reconnection interval to 20s.
A	A#configure A(config)# ip msdp peer 20.0.0.4 connect-source gi0/1 A(config)# ip msdp description 20.0.0.4 peer-router-B A(config)# end A# show ip msdp peer 20.0.0.4 …… A#configure A(config)# ip msdp timer 20 A(config)# end

B	B# configure B(config)# ip msdp peer 20.0.0.3 connect-source gi0/1 B(config)# end
Verification	● On device B, shut down and then immediately reconnect the connection with the MSDP peer. ● Check whether the MSDP peer is in UP state within 20s.
A	A#debug ip msdp timer
B	B# configure B(config)# show ip msdp peer 20.0.0.3

10.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Resets the TCP connection with a specified MSDP peer.	**clear ip msdp peer***peer-address*
Clears the SA cache.	**clear ip msdp sa-cache**[*group-address*]
Clears the statistics of MSDP peers.	**clear ip msdp statistics** [*peer-address*]

Displaying

Description	Command
Displays the number of sources and number of groups generated by SA messages.	**show ip msdp count**[*as-number*]
Displays information about a mesh group.	**show ip msdp mesh-group**
Displays detailed information about MSDP peers.	**show ip msdp peer** [*peer-address*]
Displays information about the MSDP RPF peer corresponding to the specified initiator address.	**show ip msdp rpf-peer***ip-address*
Displays the learned (S, G) information.	**show ip msdpsa-cache** [*group-address* **|** *source-address*] [*group-address***|** *source-address*] [*as-number*]
Displays the (S, G) information initiated by the local device.	**show ip msdpsa-originated**
Displays brief information about all MSDP peers.	**show ip msdp summary**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs MSDP peers.	**debug ip msdp peer**

Chapter7 Security Configuration

1. Configuring AAA
2. Configuring RADIUS
3. Configuring TACACS+
4. Configuring 802.1X
5. Configuring Web Authentication
6. Configuring SCC
7. Configuring Global IP-MAC Binding
8. Configuring Password Policy
9. Configuring Port Security
10. Configuring Storm Control
11. Configuring SSH
12. Configuration URPF
13. Configuring CPU Protection
14. Configuring DHCP Snooping
15. Configuring DHCPv6 Snooping
16. Configuring ARP Check
17. Configuring Dynamic ARP Inspection
18. Configuring IP Source Guard
19. Configuring IPv6 Source Guard
20. Configuring Gateway-targeted ARP-Spoofing Prevention
21. Configuring NFPP
22. Configuring DoS Protection

1 Configuring AAA

1.1 Overview

Authentication, authorization, and accounting (AAA) provides a unified framework for configuring the authentication, authorization, and accounting services. FS Networks devices support the AAA application.

AAA provides the following services in a modular way:

Authentication: Refers to the verification of user identities for network access and network services. Authentication is classified into local authentication and authentication through Remote Authentication Dial In User Service (RADIUS) and Terminal Access Controller Access Control System+ (TACACS+).

Authorization: Refers to the granting of specific network services to users according to a series of defined attribute-value (AV) pairs. The pairs describe what operations users are authorized to perform. AV pairs are stored on network access servers (NASs) or remote authentication servers.

Accounting: Refers to the tracking of the resource consumption of users. When accounting is enabled, NASs collect statistics on the network resource usage of users and send them in AV pairs to authentication servers. The records will be stored on authentication servers, and can be read and analyzed by dedicated software to realize the accounting, statistics, and tracking of network resource usage.

AAA is the most fundamental method of access control. FS Networks also provides other simple access control functions, such as local username authentication and online password authentication. Compared to them, AAA offers higher level of network security.

AAA has the following advantages:

- Robust flexibility and controllability
- Scalability
- Standards-compliant authentication
- Multiple standby systems

1.2 Applications

Application	Description
Configuring AAA in a Single-Domain Environment	AAA is performed for all the users in one domain.
Configuring AAA in a Multi-Domain Environment	AAA is performed for the users in different domains by using different methods.

1.2.1Configuring AAA in a Single-Domain Environment

Scenario

In the network scenario shown in Figure 1-1, the following application requirements must be satisfied to improve the security management on the NAS:

12. To facilitate account management and avoid information disclosure, each administrator has an individual account with different username and password.

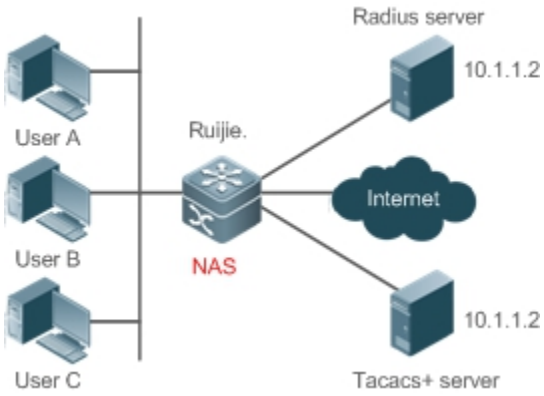
13. Users must pass identity authentication before accessing the NAS. The authentication can be in local or centralized mode. It is recommended to combine the two modes, with centralized mode as active and local mode as standby. As a result, users must undergo authentication by the RADIUS server first. If the RADIUS server does not respond, it turns to local authentication.

14. During the authentication process, users can be classified and limited to access different NASs.

15. Permission management: Users managed are classified into Super User and Common User. Super users have the rights to view and configure the NAS, and common users are only able to view NAS configuration.

16. The AAA records of users are stored on servers and can be viewed and referenced for auditing. (The TACACS+ server in this example performs the accounting.)

Figure 1-1



Remarks	User A, User B, and User C are connected to the NAS in wired or wireless way. The NAS is an access or convergence switch. The RADIUS server can be the Windows 2000/2003 Server (IAS), UNIX system component, and dedicated server software provided by a vendor. The TACACS+ server can be the dedicated server software provided by a vendor.

Deployment

- Enable AAA on the NAS.
- Configure an authentication server on the NAS.
- Configure local users on the NAS.
- Configure the authentication service on the NAS.
- Configure the authorization service on the NAS.
- Configure the accounting service on the NAS.

1.2.2 Configuring AAA in a Multi-Domain Environment

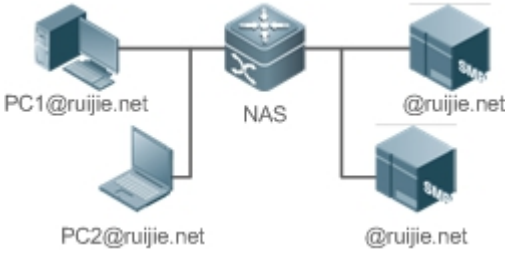
Scenario

Configure the domain-based AAA service on the NAS.

- A user can log in by entering the username PC1@FS.net or PC2@FS.com.cn and correct password on an 802.1X client.
- Permission management: Users managed are classified into Super User and Common User. Super users have the rights to view and configure the NAS, and common users are only able to view NAS configuration.

- The AAA records of users are stored on servers and can be viewed and referenced for auditing.

Figure 1-2



Remarks	The clients with the usernames **PC1@FS.net** and **PC2@FS.com.cn** are connected to the NAS in wired or wireless way. The NAS is an access or convergence switch. The Security Accounts Manager (SAM) server is a universal RADIUS server provided by FS Networks.

Deployment

- Enable AAA on the NAS.
- Configure an authentication server on the NAS.
- Configure local users on the NAS.
- Define an AAA method list on the NAS.
- Enable domain-based AAA on the NAS.
- Create domains and AV sets on the NAS.

1.3 Features

Basic Concepts

Local Authentication and Remote Server Authentication

Local authentication is the process where the entered passwords are verified by the database on the NAS.

Remote server authentication is the process where the entered passwords are checked by the database on a remote server. It is mainly implemented by the RADIUS server and TACACS+ server.

Method List

AAA is implemented using different security methods. A method list defines a method implementation sequence. The method list can contain one or more security protocols so that a standby method can take over the AAA service when the first method fails. On FS devices, the first method in the list is tried in the beginning and then the next is tried one by one if the previous gives no response. This method selection process continues until a security method responds or all the security methods in the list are tried out. Authentication fails if no method in the list responds.

A method list contains a series of security methods that will be queried in sequence to verify user identities. It allows you to define one or more security protocols used for authentication, so that the standby authentication method takes over services when the active security method fails. On FS devices, the first method in the list is tried in the beginning and then the next is tried one by one if the previous gives no response. This method selection process continues until a method responds or all the methods in the method list are tried out. Authentication fails if no method in the list responds.

The next authentication method proceeds on FS devices only when the current method does not respond. When a method denies user access, the authentication process ends without trying other methods.

Figure 1-3

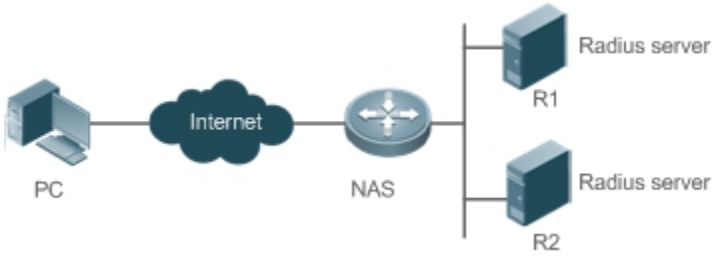


Figure 1-3 shows a typical AAA network topology, where two RADIUS servers (R1 and R2) and one NAS are deployed. The NAS can be the client for the RADIUS servers.

Assume that the system administrator defines a method list, where the NAS selects R1 and R2 in sequence to obtain user identity information and then accesses the local username database on the server. For example, when a remote PC user initiates dial-up access, the NAS first queries the user's identity on R1. When the authentication on R1 is completed, R1 returns an Accept response to the NAS. Then the user is permitted to access the Internet. If R1 returns a Reject response, the user is denied Internet access and the connection is terminated. If R1 does not respond, the NAS considers that the R1 method times out and continues to query the user's identity on R2. This process continues as the NAS keeps trying the remaining authentication methods, until the user request is authenticated, rejected, or terminated. If all the authentication methods are responded with Timeout, authentication fails and the connection will be terminated.

The Reject response is different from the Timeout response. The Reject response indicates that the user does not meet the criteria of the available authentication database and therefore fails in authentication, and the Internet access request is denied. The Timeout response indicates that the authentication server fails to respond to the identity query. When detecting a timeout event, the AAA service proceeds to the next method in the list to continue the authentication process.

This document describes how to configure AAA on the RADIUS server. For details about the configuration on the TACACS+ server, see the *Configuring TACACS+*.

AAA Server Group

You can define an AAA server group to include one or more servers of the same type. If the server group is referenced by a method list, the NAS preferentially sends requests to the servers in the referenced server group when the method list is used to implement AAA.

VRF-enabled AAA Groups

Virtual Private Networks (VPNs) provide a secure way for users to share bandwidth across the ISP backbone. A VPN is a set of sites that share routes. User sites are linked to the service provider network through one or more interfaces, and the VPN routing table, also called the VPN routing/forwarding (VRF) table, allows AAA to assign VRFs to each custom server group.

When AAA operations are performed using a server group with a VRF assigned, request messages are initiated to the remote server. The source address of these request messages is a suitable source IP address based on the server IP address found in the configured VRF.

If you specify the source interface of a request message with the command ip radius/tacacs+ source-interface, the IP address obtained from that source interface takes precedence over the IP address found in the VRF.

Overview

Feature	Description
AAA Authentication	Verifies whether users can access the Internet.
AAA Authorization	Determines what services or permissions users can enjoy.
AAA Accounting	Records the network resource usage of users.
Multi-Domain AAA	Creates domain-specific AAA schemes for 802.1X stations (STAs) in different domains.

1.3.1 AAA Authentication

Authentication, authorization, and accounting are three independent services. The authentication service verifies whether users can access the Internet. During authentication, the username, password, and other user information are exchanged between devices to complete users' access or service requests. You can use only the authentication service of AAA.

To configure AAA authentication, you need to first configure an authentication method list. Applications perform authentication according to the method list. The method list defines the types of authentication and the sequence in which they are performed. Authentication methods are implemented by specified applications. The only exception is the default method list. All applications use the default method list if no method list is configured.

AAA Authentication Scheme

- No authentication (**none**)

The identity of trusted users is not checked. Normally, the no-authentication (None) method is not used.

- Local authentication (**local**)

Authentication is performed on the NAS, which is configured with user information (including usernames, passwords, and AV pairs). Before local authentication is enabled, run the **username password/secret** command to create a local user database.

- Remote server group authentication (**group**)

Authentication is performed jointly by the NAS and a remote server group through RADIUS or TACACS+. A server group consists of one or more servers of the same type. User information is managed centrally on a remote server, thus realizing multi-device centralized and unified authentication with high capacity and reliability. You can configure local authentication as standby to avoid authentication failures when all the servers in the server group fail.

AAA Authentication Types

FS products support the following authentication types:

- Login authentication

Users log in to the command line interface (CLI) of the NAS for authentication through Secure Shell (SSH), Telnet, and File Transfer Protocol (FTP).

- Enable authentication

After users log in to the CLI of the NAS, the users must be authenticated before CLI permission update. This process is called Enable authentication (in Privileged EXEC mode).

- Dot1X (IEEE802.1X) authentication

Dot1X (IEEE802.1X) authentication is performed for users that initiate dial-up access through IEEE802.1X.

- Web (second generation portal) authentication

Web authentication is performed by the second generation portal server.

Related Configuration

Enabling AAA

By default, AAA is disabled.

To enable AAA, run the **aaa new-model** command.

Configuring an AAA Authentication Scheme

By default, no AAA authentication scheme is configured.

Before you configure an AAA authentication scheme, determine whether to use local authentication or remote server authentication. If the latter is to be implemented, configure a RADIUS or TACACS+ server in advance. If local authentication is selected, configure the local user database information on the NAS.

Configuring an AAA Authentication Method List

By default, no AAA authentication method list is configured.

Determine the access mode to be configured in advance. Then configure authentication methods according to the access mode.

1.3.2 AAA Authorization

AAA authorization allows administrators to control the services or permissions of users. After AAA authorization is enabled, the NAS configures the sessions of users according to the user configuration files stored on the NAS or servers. After authorization, users can use only the services or have only the permissions permitted by the configuration files.

AAA Authorization Scheme

- Direct authorization (**none**)

Direct authorization is intended for highly trusted users, who are assigned with the default permissions specified by the NAS.

- Local authorization (**local**)

Local authorization is performed on the NAS, which authorizes users according to the AV pairs configured for local users.

- Remote server-group authorization (**group**)

Authorization is performed jointly by the NAS and a remote server group. You can configure local or direct authorization as standby to avoid authorization failures when all the servers in the server group fail.

AAA Authorization Types

- EXEC authorization

After users log in to the CLI of the NAS, the users are assigned with permission levels (0 to 15).

- Config-commands authorization

Users are assigned with the permissions to run specific commands in configuration modes (including the global configuration mode and sub-modes).

- Console authorization

After users log in through consoles, the users are authorized to run commands.

- Command authorization

Authorize users with commands after login to the CLI of the NAS.

- Network authorization

After users access the Internet, the users are authorized to use the specific session services. For example, after users access the Internet through PPP and Serial Line Internet Protocol (SLIP), the users are authorized to use the data service, bandwidth, and timeout service.

Related Configuration

Enabling AAA

By default, AAA is disabled.

To enable AAA, run the **aaa new-model** command.

Configuring an AAA Authorization Scheme

By default, no AAA authorization scheme is configured.

Before you configure an AAA authorization scheme, determine whether to use local authorization or remote server-group authorization. If remote server-group authorization needs to be implemented, configure a RADIUS or TACACS+ server in advance. If local authorization needs to be implemented, configure the local user database information on the NAS.

Configuring an AAA Authorization Method List

By default, no AAA authorization method list is configured.

Determine the access mode to be configured in advance. Then configure authorization methods according to the access mode.

1.3.3 AAA Accounting

In AAA, accounting is an independent process of the same level as authentication and authorization. During the accounting process, start-accounting, update-accounting, and end-accounting requests are sent to the configured accounting server, which records the network resource usage of users and performs accounting, audit, and tracking of users' activities.

In AAA configuration, accounting scheme configuration is optional.

AAA Accounting Schemes

- No accounting (**none**)

Accounting is not performed on users.

- Local accounting (**local**)

Accounting is completed on the NAS, which collects statistics on and limits the number of local user connections. Billing is not performed.

- Remote server-group accounting (**group**)

Accounting is performed jointly by the NAS and a remote server group. You can configure local accounting as standby to avoid accounting failures when all the servers in the server group fail.

AAA Accounting Types

- EXEC accounting

Accounting is performed when users log in to and out of the CLI of the NAS.

- Command accounting

Records are kept on the commands that users run on the CLI of the NAS.

- Network accounting

Records are kept on the sessions that users set up after completing 802.1X and Web authentication to access the Internet.

Related Configuration

Enabling AAA

By default, AAA is disabled.

To enable AAA, run the **aaa new-model** command.

Configuring an AAA Accounting Scheme

By default, no AAA accounting method is configured.

Before you configure an AAA accounting scheme, determine whether to use local accounting or remote server-group accounting. If remote server-group accounting needs to be implemented, configure a RADIUS or TACACS+ server in advance. If local accounting needs to be implemented, configure the local user database information on the NAS.

Configuring an AAA Accounting Method List

By default, no AAA accounting method list is configured.

Determine the access mode to be configured in advance. Then configure accounting methods according to the access mode.

1.3.4 Multi-Domain AAA

In a multi-domain environment, the NAS can provide the AAA services to users in different domains. The user AVs (such as usernames and passwords, service types, and permissions) may vary with different domains. It is necessary to configure domains to differentiate the user AVs in different domains and configure an AV set (including an AAA service method list, for example, RADIUS) for each domain.

Our products support the following username formats:

1. userid@domain-name
2. domain-name\userid
3. userid.domain-name
4. userid

The fourth format (userid) does not contain a domain name, and it is considered to use the **default** domain name.

The NAS provides the domain-based AAA service based on the following principles:

- Resolves the domain name carried by a user.
- Searches for the user domain according to the domain name.
- Searches for the corresponding AAA method list name according to the domain configuration information on the NAS.
- Searches for the corresponding method list according to the method list name.
- Provides the AAA services based on the method list.

If any of the preceding procedures fails, the AAA services cannot be provided.

Related Configuration

Enabling AAA

By default, AAA is disabled.

To enable AAA, run the **aaa new-model** command.

Configuring an AAA Method List

By default, no AAA method list is configured.

For details, see section 5.2.1, section 5.2.2, and section 5.2.3.

Enabling the Domain-Based AAA Service

By default, the domain-based AAA service is disabled.

To enable the domain-based AAA service, run the **aaa domain enable** command.

Creating a Domain

By default, no domain is configured.

To configure a domain, run the **aaa domain** *domain-name* command.

Configuring an AV Set for a Domain

By default, no domain AV set is configured.

A domain AV set contains the following elements: AAA method lists, the maximum number of online users, whether to remove the domain name from the username, and whether the domain name takes effect.

Displaying Domain Configuration

To display domain configuration, run the **show aaa domain** command.

The system supports a maximum of 32 domains.

1.4 Configuration

Configuration	Description and Command	
Configuring AAA Authentication	Mandatory if user identities need to be verified.	
	aaa new-model	Enables AAA.
	aaa authentication login	Defines a method list of login authentication.
	aaa authentication enable	Defines a method list of Enable authentication.
	aaa authentication dot1x	Defines a method list of 802.1X authentication.
	aaa authentication web-auth	Configures a method list of Web authentication.
	aaa local authentication attempts	Sets the maximum number of login attempts.
	aaa local authentication lockout-time	Sets the maximum lockout time after a login failure.

Configuration	Description and Command	
Configuring AAA Authorization	Mandatory if different permissions and services need to be assigned to users.	
	aaa new-model	Enables AAA.
	aaa authorization exec	Defines a method list of EXEC authorization.
	aaa authorization commands	Defines a method list of command authorization.
	aaa authorization network	Configures a method list of network authorization.
	authorization exec	Applies EXEC authorization methods to a specified VTY line.
	authorization commands	Applies command authorization methods to a specified VTY line.
Configuring AAA Accounting	Mandatory if accounting, statistics, and tracking need to be performed on the network resource usage of users.	
	aaa new-model	Enables AAA.
	aaa accounting exec	Defines a method list of EXEC accounting.
	aaa accounting commands	Defines a method list of command accounting.
	aaa accounting network	Defines a method list of network accounting.
	accounting exec	Applies EXEC accounting methods to a specified VTY line.
	accounting commands	Applies command accounting methods to a specified VTY line.
	aaa accounting update	Enables accounting update.
	aaa accounting update periodic	Configures the accounting update interval.
Configuring an AAA Server Group	Recommended if a server group needs to be configured to handle AAA through different servers in the group.	
	aaa group server	Creates a user-defined AAA server group.
	server	Adds an AAA server group member.
Configuring the Domain-Based AAA Service	Mandatory if AAA management of 802.1X access STAs needs to be performed according to domains.	
	aaa new-model	Enables AAA.
	aaa domain enable	Enables the domain-based AAA service.
	aaa domain	Creates a domain and enters domain configuration mode.
	authentication dot1x	Associates the domain with an 802.1X authentication method list.
	accounting network	Associates the domain with a network accounting method list.
	authorization network	Associates the domain with a network authorization method list.
	state	Configures the domain status.

Configuration	Description and Command	
	username-format	Configures whether to contain the domain name in usernames.
	access-limit	Configures the maximum number of domain users.

1.4.1 Configuring AAA Authentication

Configuration Effect

Verify whether users are able to obtain access permission.

Notes

- If an authentication scheme contains multiple authentication methods, these methods are executed according to the configured sequence.
- When the **none** method is used, users can get access even when no authentication method gets response. Therefore, the **none** method is used only as standby.
- Normally, do not use None authentication. You can use the **none** method as the last optional authentication method in special cases. For example, all the users who may request access are trusted users and the users' work must not be delayed by system faults. Then you can use the **none** method to assign access permissions to these users when the authentication server does not respond. It is recommended that the local authentication method be added before the **none** method.
- If AAA authentication is enabled but no authentication method is configured and the default authentication method does not exist, users can directly log in to the Console without being authenticated. If users log in by other means, the users must pass local authentication.
- When a user enters the CLI after passing login authentication (the **none** method is not used), the username is recorded. When the user performs Enable authentication, the user is not prompted to enter the username again, because the username that the user entered during login authentication is automatically filled in. However, the user must enter the password previously used for login authentication.
- The username is not recorded if the user does not perform login authentication when entering the CLI or the **none** method is used during login authentication. Then, a user is required to enter the username each time when performing Enable authentication.

Configuration Steps

Enabling AAA

- Mandatory.
- Run the **aaa new-model** command to enable AAA.
- By default, AAA is disabled.

Defining a Method List of Login Authentication

- Run the **aaa authentication login** command to configure a method list of login authentication.
- This configuration is mandatory if you need to configure a login authentication method list (including the configuration of the default method list).
- By default, no method list of login authentication is configured.

Defining a Method List of Enable Authentication

- Run the **aaa authentication enable** command to configure a method list of Enable authentication.
- This configuration is mandatory if you need to configure an Enable authentication method list. (You can configure only the default method list.)
- By default, no method list of Enable authentication is configured.

Defining a Method List of 802.1X Authentication

- Run the **aaa authentication dot1x** command to configure a method list of 802.1X authentication.
- This configuration is mandatory if you need to configure an 802.1X authentication method list (including the configuration of the default method list).
- By default, no method list of 802.1X authentication is configured.

Defining a Method List of Web Authentication

- Run the **aaa authentication web-auth** command to configure a method list of Web authentication.
- This configuration is mandatory if you need to configure a Web authentication method list (including the configuration of the default method list).
- By default, no method list of Web authentication is configured.

Setting the Maximum Number of Login Attempts

- Optional.
- By default, a user is allowed to enter passwords up to three times during login.

Setting the Maximum Lockout Time After a Login Failure

- Optional.
- By default, a user is locked for 15 minutes after entering wrong passwords three times.

Verification

- Run the **show aaa method-list** command to display the configured method lists.
- Run the **show aaa lockout** command to display the settings of the maximum number of login attempts and the maximum lockout time after a login failure.
- Run the **show running-config** command to display the authentication method lists associated with login authentication and 802.1X authentication.

Related Commands

Enabling AAA

Command	**aaa new-model**
Parameter Description	N/A
Command Mode	Global configuration mode

Usage Guide	To enable the AAA services, run this command. None of the rest of AAA commands can be effective if AAA is not enabled.

Defining a Method List of Login Authentication

Command	**aaa authentication login** { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	**default:** With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of a login authentication method list in characters. *method*: Indicates authentication methods from **local**, **none**, **group**, and **subs**. A method list contains up to four methods. **local**: Indicates that the local user database is used for authentication. **none**: Indicates that authentication is not performed. **group**: Indicates that a server group is used for authentication. Currently, the RADIUS and TACACS+ server groups are supported. **subs**: Indicates that the subs database is used for authentication.
Command Mode	Global configuration mode
Usage Guide	If the AAA login authentication service is enabled on the NAS, users must perform login authentication negotiation through AAA. Run the **aaa authentication login** command to configure the default or optional method lists for login authentication. In a method list, the next method is executed only when the current method does not receive response. After you configure login authentication methods, apply the methods to the VTY lines that require login authentication; otherwise, the methods will not take effect.

Defining a Method List of Enable Authentication

Command	**aaa authentication enable default** *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of an Enable authentication method list in characters. *method*: Indicates authentication methods from **enable**, **local**, **none**, and **group**. A method list contains up to four methods. **enable**: Indicates that the password that is configured using the **enable** command is used for authentication. **local**: Indicates that the local user database is used for authentication. **none**: Indicates that authentication is not performed. **group**: Indicates that a server group is used for authentication. Currently, the RADIUS and TACACS+ server groups are supported.
Command Mode	Global configuration mode
Usage Guide	If the AAA login authentication service is enabled on the NAS, users must perform Enable authentication negotiation through AAA. Run the **aaa authentication enable** command to configure the default or optional method lists for Enable authentication. In a method list, the next method is executed only when the current method does not receive response.

Defining a Method List of 802.1X Authentication

Command	**aaa authentication dot1x** { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of an 802.1X authentication method list in characters.

	method: Indicates authentication methods from **local**, **none**, and **group**. A method list contains up to four methods. **local**: Indicates that the local user database is used for authentication. **none**: Indicates that authentication is not performed. **group**: Indicates that a server group is used for authentication. Currently, the RADIUS server group is supported.
Command Mode	Global configuration mode
Usage Guide	If the AAA 802.1X authentication service is enabled on the NAS, users must perform 802.1X authentication negotiation through AAA. Run the **aaa authentication dot1x** command to configure the default or optional method lists for 802.1X authentication. In a method list, the next method is executed only when the current method does not receive response.

Defining a Method List of PPP, Web or SSL VPN Authentication

Command	**aaa authentication** { **web-auth** } { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	**web-auth**: Configures a method list of Web authentication. **default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of a PPP authentication method list in characters. *method*: Indicates authentication methods from **local**, **none**, **group**, and **subs**. A method list contains up to four methods. **local**: Indicates that the local user database is used for authentication. **none**: Indicates that authentication is not performed. **group**: Indicates that a server group is used for authentication. Currently, the RADIUS and TACACS+ server groups are supported. **subs**: Specifies the SUBS authentication method using the SUBS database.
Command Mode	Global configuration mode
Usage Guide	If the AAA PPP authentication service is enabled on the NAS, users must perform PPP authentication negotiation through AAA. Run the **aaa authentication ppp** command to configure the default or optional method lists for PPP authentication. In a method list, the next method is executed only when the current method does not receive response.

Setting the Maximum Number of Login Attempts

Command	**aaa local authentication attempts** *max-attempts*
Parameter Description	*max-attempts*: Indicates the maximum number of login attempts. The value ranges from 1 to 2,147,483,647.
Command Mode	Global configuration mode
Usage Guide	Use this command to set the maximum number of times a user can attempt to login.

Setting the Maximum Lockout Time After a Login Failure

Command	**aaa local authentication lockout-time** *lockout-time*
Parameter Description	*lockout-time*: Indicates the time during which a user is locked after entering wrong passwords up to the specified times. The value ranges from 1 to 2,147,483,647, in the unit of minutes.
Command	Global configuration mode

Mode	
Usage Guide	Use this command to set the maximum time during which a user is locked after entering wrong passwords up to the specified times.

Configuration Example

Configuring AAA Login Authentication

Configure a login authentication method list on the NAS containing **group** *radius* and **local** methods in order.

Scenario **Figure 1-5**	
Configuration Steps	Step 1: Enable AAA. Step 2: Configure a RADIUS or TACACS+ server in advance if group-server authentication needs to be implemented. Configure the local user database information on the NAS if local authentication needs to be implemented. (This example requires the configuration of a RADIUS server and local database information.) Step 3: Configure an AAA authentication method list for login authentication users. (This example uses **group** *radius* and **local** in order.) Step 4: Apply the configured method list to an interface or line. Skip this step if the default authentication method is used.
NAS	FS#configure terminal FS(config)#username user password pass FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key FS FS(config)#aaa authentication login list1 group radius local FS(config)#line vty 0 20 FS(config-line)#login authentication list1 FS(config-line)#exit
Verification	Run the **show aaa method-list** command on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication login list1 group radius local Accounting method-list:

	Authorization method-list:
	Assume that a user remotely logs in to the NAS through Telnet. The user is prompted to enter the username and password on the CLI. The user must enter the correct username and password to access the NAS.
User	User Access Verification Username:user Password:pass

Configuring AAA Enable Authentication

Configure an Enable authentication method list on the NAS containing **group** *radius*, **local**, and then **enable** methods in order.

Scenario **Figure 1-6**	
Configuration Steps	Step 1: Enable AAA. Step 2: Configure a RADIUS or TACACS+ server in advance if group-server authentication needs to be implemented. Configure the local user database information on the NAS if local authentication needs to be implemented. Configure Enable authentication passwords on the NAS if you use Enable password authentication. Step 3: Configure an AAA authentication method list for Enable authentication users. You can define only one Enable authentication method list globally. You do not need to define the list name but just default it. After that, it will be applied automatically.
NAS	FS#configure terminal FS(config)#username user privilege 15 password pass FS(config)#enable secret w FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key FS FS(config)#aaa authentication enable default group radius local enable
Verification	Run the **show aaa method-list** command on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication enable default group radius local enable

	Accounting method-list: Authorization method-list:
	The CLI displays an authentication prompt when the user level is updated to level 15. The user must enter the correct username and password to access the NAS.
NAS	FS>enable Username:user Password:pass FS#

Configuring AAA 802.1X Authentication

Configure an 802.1X authentication method list on the NAS containing **group** *radius*, and then **local** methods in order.

Scenario **Figure 1-7**	
Configuration Steps	Step 1: Enable AAA. Step 2: Configure a RADIUS server in advance if group-server authentication needs to be implemented. Configure the local user database information on the NAS if local authentication needs to be implemented. (This example requires the configuration of a RADIUS server and local database information.) Currently, 802.1X authentication does not support TACACS+. Step 3: Configure an AAA authentication method list for 802.1X authentication users. (This example uses **group** *radius* and **local** in order.) Step 4: Apply the AAA authentication method list. Skip this step if the default authentication method is used. Step 5: Enable 802.1X authentication on an interface.
NAS	FS#configure terminal FS(config)#username user1 password pass1 FS(config)#username user2 password pass2 FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key FS FS(config)#aaa authentication dot1x default group radius local FS(config)#interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)#dot1 port-control auto FS(config-if-gigabitEthernet 0/1)#exit

Verification	Run the **show aaa method-list** command on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication dot1x default group radius local Accounting method-list: Authorization method-list:

Common Errors

- No RADIUS server or TACACS+ server is configured.
- Usernames and passwords are not configured in the local database.

1.4.2 Configuring AAA Authorization

Configuration Effect

- Determine what services or permissions authenticated users can enjoy.

Notes

- EXEC authorization is often used with login authentication, which can be implemented on the same line. Authorization and authentication can be performed using different methods and servers. Therefore, the results of the same user may be different. If a user passes login authentication but fails in EXEC authorization, the user cannot enter the CLI.
- The authorization methods in an authorization scheme are executed in accordance with the method configuration sequence. The next authorization method is executed only when the current method does not receive response. If authorization fails using a method, the next method will be not tried.
- Command authorization is supported only by TACACS+.
- Console authorization: The FSOS can differentiate between the users who log in through the Console and the users who log in through other types of clients. You can enable or disable command authorization for the users who log in through the Console. If command authorization is disabled for these users, the command authorization method list applied to the Console line no longer takes effect.

Configuration Steps

Enabling AAA

- Mandatory.
- Run the **aaa new-model** command to enable AAA.
- By default, AAA is disabled.

Defining a Method List of EXEC Authorization

- Run the **aaa authorization exec** command to configure a method list of EXEC authorization.
- This configuration is mandatory if you need to configure an EXEC authorization method list (including the configuration of the default method list).
- By default, no EXEC authorization method list is configured.

The default access permission level of EXEC users is the lowest. (Console users can connect to the NAS through the Console port or Telnet. Each connection is counted as an EXEC user, for example, a Telnet user and SSH user.)

Defining a Method List of Command Authorization

- Run the **aaa authorization commands** command to configure a method list of command authorization.
- This configuration is mandatory if you need to configure a command authorization method list (including the configuration of the default method list).
- By default, no command authorization method list is configured.

Configuring a Method List of Network Authorization

- Run the **aaa authorization network** command to configure a method list of network authorization.
- This configuration is mandatory if you need to configure a network authorization method list (including the configuration of the default method list).
- By default, no authorization method is configured.

Applying EXEC Authorization Methods to a Specified VTY Line

- Run the **authorization exec** command in line configuration mode to apply EXEC authorization methods to a specified VTY line.
- This configuration is mandatory if you need to apply an EXEC authorization method list to a specified VTY line.
- By default, all VTY lines are associated with the default authorization method list.

Applying Command Authorization Methods to a Specified VTY Line

- Run the **authorization commands** command in line configuration mode to apply command authorization methods to a specified VTY line.
- This configuration is mandatory if you need to apply a command authorization method list to a specified VTY line.
- By default, all VTY lines are associated with the default authorization method list.

Enabling Authorization for Commands in Configuration Modes

- Run the **aaa authorization config-commands** command to enable authorization for commands in configuration modes.
- By default, authorization is disabled for commands in configuration modes.

Enabling Authorization for the Console to Run Commands

- Run the **aaa authorization console** command to enable authorization for console users to run commands.
- By default, authorization is disabled for the Console to run commands.

Verification

Run the **show running-config** command to verify the configuration.

Related Commands

Enabling AAA

Command	**aaa new-model**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	To enable the AAA services, run this command. None of the rest of AAA commands can be effective if AAA is not enabled.

Defining a Method List of EXEC Authorization

Command	**aaa authorization exec** { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of an EXEC authorization method list in characters. *method*: Specifies authentication methods from **local**, **none**, and **group**. A method list contains up to four methods. **local**: Indicates that the local user database is used for EXEC authorization. **none**: Indicates that EXEC authorization is not performed. **group**: Indicates that a server group is used for EXEC authorization. Currently, the RADIUS and TACACS+ server groups are supported.
Command Mode	Global configuration mode
Usage Guide	The FSOS supports authorization of the users who log in to the CLI of the NAS to assign the users CLI operation permission levels (0 to 15). Currently, EXEC authorization is performed only on the users who have passed login authentication. If a user fails in EXEC authorization, the user cannot enter the CLI. After you configure EXEC authorization methods, apply the methods to the VTY lines that require EXEC authorization; otherwise, the methods will not take effect.

Defining a Method List of Command Authorization

Command	**aaa authorization commands** *level* { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of a command authorization method list in characters. *method*: Indicates authentication methods from **none** and **group**. A method list contains up to four methods. **none**: Indicates that command authorization is not performed. **group**: Indicates that a server group is used for command authorization. Currently, the TACACS+ server group is supported.
Command Mode	Global configuration mode
Usage Guide	The FSOS supports authorization of the commands executable by users. When a user enters a command, AAA sends the command to the authentication server. If the authentication server permits the execution, the command is executed. If the authentication server forbids the execution, the command is not executed and a message is displayed showing that the execution is rejected. When you configure command authorization, specify the command level, which is used as the default level. (For example, if a command above Level 14 is visible to users, the default level of the command is 14.)

	After you configure command authorization methods, apply the methods to the VTY lines that require command authorization; otherwise, the methods will not take effect.

Configuring a Method List of Network Authorization

Command	**aaa authorization network** { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of a network authorization method list in characters. *method*: Indicates authentication methods from **none** and **group**. A method list contains up to four methods. **none**: Indicates that authentication is not performed. **group**: Indicates that a server group is used for network authorization. Currently, the RADIUS and TACACS+ server groups are supported.
Command Mode	Global configuration mode
Usage Guide	The FSOS supports authorization of network-related service requests such as PPP and SLIP requests. After authorization is configured, all authenticated users or interfaces are authorized automatically. You can configure three different authorization methods. The next authorization method is executed only when the current method does not receive response. If authorization fails using a method, the next method will be not tried. RADIUS or TACACS+ servers return a series of AV pairs to authorize authenticated users. Network authorization is based on authentication. Only authenticated users can perform network authorization.

Enabling Authorization for Commands in Configuration Modes (Including the Global Configuration Mode and Sub-Modes)

Command	**aaa authorization config-commands**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	If you need to enable authorization for commands only in non-configuration modes (for example, privileged EXEC mode), disable authorization in configuration modes by using the **no** form of this command. Then users can run commands in configuration mode and sub-modes without authorization.

Enabling Authorization for the Console to Run Commands

Command	**aaa authorization console**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	The FSOS can differentiate between the users who log in through the Console and the users who log in through other types of clients. You can enable or disable command authorization for the users who log in through the Console. If command authorization is disabled for these users, the command authorization method list applied to the Console line no longer takes effect.

Configuration Example

Configuring AAA EXEC Authorization

Configure login authentication and EXEC authorization for users on VTY lines 0 to 4. Login authentication is performed in local mode, and EXEC authorization is performed on a RADIUS server. If the RADIUS server does not respond, users are redirected to the local authorization.

Scenario **Figure 1-8**	
Configuration Steps	Step 1: Enable AAA. Step 2: Configure a RADIUS or TACACS+ server in advance if remote server-group authorization needs to be implemented. If local authorization needs to be implemented, configure the local user database information on the NAS. Step 3: Configure an AAA authorization method list according to different access modes and service types. Step 4: Apply the configured method list to an interface or line. Skip this step if the default authorization method is used. EXEC authorization is often used with login authentication, which can be implemented on the same line.
NAS	FS#configure terminal FS(config)#username user password pass FS(config)#username user privilege 6 FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key test FS(config)#aaa authentication login list1 group local FS(config)#aaa authorization exec list2 group radius local FS(config)#line vty 0 4 FS(config-line)#login authentication list1 FS(config-line)# authorization exec list2 FS(config-line)#exit
Verification	Run the **show run** and **show aaa method-list** commands on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication login list1 group local Accounting method-list: Authorization method-list: aaa authorization exec list2 group radius local

```
FS# show running-config
aaa new-model
!
aaa authorization exec list2 group local
aaa authentication login list1 group radius local
!
username user password    pass
username user privilege 6
!
radius-server host 10.1.1.1
radius-server key 7 093b100133
!
line con 0
line vty 0 4
  authorization exec list2
  login authentication list1
!
End
```

Configuring AAA Command Authorization

Provide command authorization for login users according to the following default authorization method: Authorize level-15 commands first by using a TACACS+ server. If the TACACS+ server does not respond, local authorization is performed. Authorization is applied to the users who log in through the Console and the users who log in through other types of clients.

Scenario **Figure 1-9**	
Configuration Steps	Step 1: Enable AAA. Step 2: Configure a RADIUS or TACACS+ server in advance if remote server-group authorization needs to be implemented. If local authorization needs to be implemented, configure the local user database information on the NAS. Step 3: Configure an AAA authorization method list according to different access modes and service types. Step 4: Apply the configured method list to an interface or line. Skip this step if the default authorization method is used.
NAS	FS#configure terminal FS(config)#username user1 password pass1 FS(config)#username user1 privilege 15

	FS(config)#aaa new-model FS(config)#tacacs-server host 192.168.217.10 FS(config)#tacacs-server key aaa FS(config)#aaa authentication login default local FS(config)#aaa authorization commands 15 default group tacacs+ local FS(config)#aaa authorization console
Verification	Run the **show run** and **show aaa method-list** commands on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication login default local Accounting method-list: Authorization method-list: aaa authorization commands 15 default group tacacs+ local
	FS#show run ! aaa new-model ! aaa authorization console aaa authorization commands 15 default group tacacs+ local aaa authentication login default local ! ! nfpp ! vlan 1 ! username user1 password 0 pass1 username user1 privilege 15 no service password-encryption !

	tacacs-server host 192.168.217.10 tacacs-server key aaa ! line con 0 line vty 0 4 ! ! end

Configuring AAA Network Authorization

Scenario **Figure 1-10**	
Configuration Steps	Step 1: Enable AAA. Step 2: Configure a RADIUS or TACACS+ server in advance if remote server-group authorization needs to be implemented. If local authorization needs to be implemented, configure the local user database information on the NAS. Step 3: Configure an AAA authorization method list according to different access modes and service types. Step 4: Apply the configured method list to an interface or line. Skip this step if the default authorization method is used.
NAS	FS#configure terminal FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key test FS(config)#aaa authorization network default group radius none FS(config)# end
Verification	Run the **show aaa method-list** command on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: Accounting method-list: Authorization method-list: aaa authorization network default group radius none

Common Errors

N/A

1.4.3 Configuring AAA Accounting

Configuration Effect

- Record the network resource usage of users.
- Record the user login and logout processes and the commands executed by users during device management.

Notes

About accounting methods:

- If an accounting scheme contains multiple accounting methods, these methods are executed according to the method configuration sequence. The next accounting method is executed only when the current method does not receive response. If accounting fails using a method, the next method will be not tried.
- After the default accounting method list is configured, it is applied to all VTY lines automatically. If a non-default accounting method list is applied to a line, it will replace the default one. If you apply an undefined method list to a line, the system will display a message indicating that accounting on this line is ineffective. Accounting will take effect only when a defined method list is applied.

EXEC accounting:

- EXEC accounting is performed only when login authentication on the NAS is completed. EXEC accounting is not performed if login authentication is not configured or the **none** method is used for authentication. If Start accounting is not performed for a user upon login, Stop accounting will not be performed when the user logs out.

Command accounting

- Only the TACACS+ protocol supports command accounting.

Configuration Steps

Enabling AAA

- Mandatory.
- Run the **aaa new-model** command to enable AAA.
- By default, AAA is disabled.

Defining a Method List of EXEC Accounting

- Run the **aaa accounting exec** command to configure a method list of EXEC accounting.
- This configuration is mandatory if you need to configure an EXEC accounting method list (including the configuration of the default method list).
- The default access permission level of EXEC users is the lowest. (Console users can connect to the NAS through the Console port or Telnet. Each connection is counted as an EXEC user, for example, a Telnet user and SSH user.)
- By default, no EXEC accounting method list is configured.

Defining a Method List of Command Accounting

- Run the **aaa accounting commands** command to configure a method list of command accounting.
- This configuration is mandatory if you need to configure a command accounting method list (including the configuration of the default method list).
- By default, no command accounting method list is configured. Only the TACACS+ protocol supports command accounting.

Defining a Method List of Network Accounting

- Run the **aaa accounting network** command to configure a method list of network accounting.
- This configuration is mandatory if you need to configure a network accounting method list (including the configuration of the default method list).
- By default, no network accounting method list is configured.

Applying EXEC Accounting Methods to a Specified VTY Line

- Run the **accounting exec** command in line configuration mode to apply EXEC accounting methods to a specified VTY line.
- This configuration is mandatory if you need to apply an EXEC accounting method list to a specified VTY line.
- You do not need to run this command if you apply the default method list.
- By default, all VTY lines are associated with the default accounting method list.

Applying Command Accounting Methods to a Specified VTY Line

- Run the **accounting commands** command in line configuration mode to apply command accounting methods to a specified VTY line.
- This configuration is mandatory if you need to apply a command accounting method list to a specified VTY line.
- You do not need to run this command if you apply the default method list.
- By default, all VTY lines are associated with the default accounting method list.

Applying 802.1X Network Accounting Methods

- Run the **dot1x accounting network** command to configure 802.1X network accounting methods.
- This configuration is mandatory if you need to specify 802.1X network accounting methods.
- You do not need to run this command if you apply the default method list.
- By default, all VTY lines are associated with the default accounting method list.

Enabling Accounting Update

- Optional.
- It is recommended that accounting update be configured for improved accounting accuracy.
- By default, accounting update is disabled.

Configuring the Accounting Update Interval

- Optional.
- It is recommended that the accounting update interval not be configured unless otherwise specified.

Verification

Run the **show running-config** command to verify the configuration.

Related Commands

Enabling AAA

Command	**aaa new-model**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	To enable the AAA services, run this command. None of the rest of AAA commands can be effective if AAA is not enabled.

Defining a Method List of EXEC Accounting

Command	**aaa accounting exec** { **default** | *list-name* } **start-stop** *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of an EXEC accounting method list in characters. *method*: Indicates authentication methods from **none** and **group**. A method list contains up to four methods. **none**: Indicates that EXEC accounting is not performed. **group**: Indicates that a server group is used for EXEC accounting. Currently, the RADIUS and TACACS+ server groups are supported.
Command Mode	Global configuration mode
Usage Guide	The FSOS enables EXEC accounting only when login authentication is completed. EXEC accounting is not performed if login authentication is not performed or the **none** authentication method is used. After accounting is enabled, when a user logs in to the CLI of the NAS, the NAS sends a start-accounting message to the authentication server. When the user logs out, the NAS sends a stop-accounting message to the authentication server. If the NAS does not send a start-accounting message when the user logs in, the NAS will not send a stop-accounting message when the user logs out. After you configure EXEC accounting methods, apply the methods to the VTY lines that require EXEC accounting; otherwise, the methods will not take effect.

Defining a Method List of Command Accounting

Command	**aaa accounting commands** *level* { **default** | *list-name* } **start-stop** *method1* [*method2*...]
Parameter Description	*level*: Indicates the command level for which accounting will be performed. The value ranges from 0 to 15. After a command of the configured level is executed, the accounting server records related information based on the received accounting packet. **default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of a command accounting method list in characters. *method*: Indicates authentication methods from **none** and **group**. A method list contains up to four methods. **none**: Indicates that command accounting is not performed. **group**: Indicates that a server group is used for command accounting. Currently, the TACACS+ server group is supported.
Command Mode	Global configuration mode
Usage Guide	The FSOS enables command accounting only when login authentication is completed. Command accounting is not

	performed if login authentication is not performed or the **none** authentication method is used. After accounting is enabled, the NAS records information about the commands of the configured level that users run and sends the information to the authentication server. After you configure command accounting methods, apply the methods to the VTY lines that require command accounting; otherwise, the methods will not take effect.

Defining a Method List of Network Accounting

Command	**aaa accounting network** { **default** | *list-name* } **start-stop** *method1* [*method2*...]
Parameter Description	**default**: With this parameter used, the configured method list will be defaulted. *list-name*: Indicates the name of a network accounting method list in characters. **start-stop**: Indicates that a start-accounting message and a stop-accounting message are sent when a user accesses a network and when the user disconnects from the network respectively. The start-accounting message indicates that the user is allowed to access the network, regardless of whether accounting is successfully enabled. *method*: Indicates authentication methods from **none** and **group**. A method list contains up to four methods. **none**: Indicates that network accounting is not performed. **group**: Indicates that a server group is used for network accounting. Currently, the RADIUS and TACACS+ server groups are supported.
Command Mode	Global configuration mode
Usage Guide	The FSOS sends record attributes to the authentication server to perform accounting of user activities. The **start-stop** keyword is used to configure user accounting options.

Enabling Accounting Update

Command	**aaa accounting update**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Accounting update cannot be used if the AAA services are not enabled. After the AAA services are enabled, run this command to enable accounting update.

Configuring the Accounting Update Interval

Command	**aaa accounting update periodic** *interval*
Parameter Description	*Interval*: Indicates the accounting update interval, in the unit of minutes. The shortest is 1 minute.
Command Mode	Global configuration mode
Usage Guide	Accounting update cannot be used if the AAA services are not enabled. After the AAA services are enabled, run this command to configure the accounting update interval.

Configuration Example

Configuring AAA EXEC Accounting

Configure login authentication and EXEC accounting for users on VTY lines 0 to 4. Login authentication is performed in local mode, and

EXEC accounting is performed on a RADIUS server.

Scenario **Figure 1-11**	
Configuration Steps	Step 1: Enable AAA. If remote server-group accounting needs to be implemented, configure a RADIUS or TACACS+ server in advance. Step 2: Configure an AAA accounting method list according to different access modes and service types. Step 3: Apply the configured method list to an interface or line. Skip this step if the default accounting method is used.
NAS	FS#configure terminal FS(config)#username user password pass FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key test FS(config)#aaa authentication login list1 group local FS(config)#aaa accounting exec list3 start-stop group radius FS(config)#line vty 0 4 FS(config-line)#login authentication list1 FS(config-line)# accounting exec list3 FS(config-line)#exit
Verification	Run the **show run** and **show aaa method-list** commands on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication login list1 group local Accounting method-list: aaa accounting exec list3 start-stop group radius Authorization method-list:
	FS# show running-config aaa new-model ! aaa accounting exec list3 start-stop group radius

```
aaa authentication login list1 group local
!
username user password    pass
!
radius-server host 10.1.1.1
radius-server key 7 093b100133
!
line con 0
line vty 0 4
  accounting    exec list3
  login authentication list1
!
End
```

Configuring AAA Command Accounting

Configure command accounting for login users according to the default accounting method. Login authentication is performed in local mode, and command accounting is performed on a TACACS+ server.

Scenario **Figure 1-12**	
Configuration Steps	Step 1: Enable AAA. If remote server-group accounting needs to be implemented, configure a RADIUS or TACACS+ server in advance. Step 2: Configure an AAA accounting method list according to different access modes and service types. Step 3: Apply the configured method list to an interface or line. Skip this step if the default accounting method is used.
NAS	FS#configure terminal FS(config)#username user1 password pass1 FS(config)#username user1 privilege 15 FS(config)#aaa new-model FS(config)#tacacs-server host 192.168.217.10 FS(config)#tacacs-server key aaa FS(config)#aaa authentication login default local FS(config)#aaa accounting commands 15 default start-stop group tacacs+
Verification	Run the **show aaa method-list** command on the NAS to display the configuration.

NAS	FS#show aaa method-list Authentication method-list: aaa authentication login default local Accounting method-list: aaa accounting commands 15 default start-stop group tacacs+ Authorization method-list:
	FS#show run ! aaa new-model ! aaa authorization config-commands aaa accounting commands 15 default start-stop group tacacs+ aaa authentication login default local ! ! nfpp ! vlan 1 ! username user1 password 0 pass1 username user1 privilege 15 no service password-encryption ! tacacs-server host 192.168.217.10 tacacs-server key aaa ! line con 0 line vty 0 4 ! !

	end

Configuring AAA Network Accounting

Configure a network accounting method list for 802.1X STAs, and configure a RADIUS remote server for authentication and accounting.

Scenario **Figure 1-13**	
Configuration Steps	Step 1: Enable AAA. Step 2: If remote server-group accounting needs to be implemented, configure a RADIUS server in advance. Step 3: Configure an AAA accounting method list according to different access modes and service types. Step 4: Apply the configured AAA accounting method list. Skip this step if the default accounting method is used. Accounting is performed only when 802.1X authentication is completed.
NAS	FS#configure terminal FS(config)#username user password pass FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key test FS(config)#aaa authentication dot1x aut1x group radius local FS(config)#aaa accounting network acc1x start-stop group radius FS(config)#dot1x authentication aut1x FS(config)#dot1x accounting acc1x FS(config)#interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#dot1 port-control auto FS(config-if-GigabitEthernet 0/1)#exit
Verification	Run the **show aaa method-list** command on the NAS to display the configuration.
NAS	FS#show aaa method-list Authentication method-list: aaa authentication dot1x aut1x group radius local Accounting method-list: aaa accounting network acc1x start-stop group radius Authorization method-list:

Common Errors

- For radius servers that use non-default authentication and accounting ports, specify both the authentication port or the accounting port when adding a server using the command server.
- Currently, only server groups of the RADIUS type can be configured with the VRF attribute.

1.4.4 Configuring an AAA Server Group

Configuration Effect

- Create a user-defined server group and add one or more servers to the group.
- When you configure authentication, authorization, and accounting method lists, name the methods after the server group name so that the servers in the group are used to handle authentication, authorization, and accounting requests.
- Use self-defined server groups to separate authentication, authorization, and accounting.

Notes

In a user-defined server group, you can specify and apply only the servers in the default server group.

Configuration Steps

Creating a User-Defined AAA Server Group

- Mandatory.
- Assign a meaningful name to the user-defined server group. Do not use the predefined **radius** and **tacacs+** keywords in naming.

Adding an AAA Server Group Member

- Mandatory.
- Run the **server** command to add AAA server group members.
- By default, a user-defined server group does not have servers.

Configuring the VRF Attribute of an AAA Server Group

- Optional.
- Run the **ip vrf forwarding** command to configure the VRF attribute of an AAA server group.
- By default, the AAA server group belongs to the global VRF table.

Verification

Run the **show aaa group** command to verify the configuration.

Related Commands

Creating a User-Defined AAA Server Group

Command	**aaa group server** {**radius** | **tacacs+**} *name*
Parameter Description	*name*: Indicates the name of the server group to be created. The name must not contain the **radius** and **tacacs+** keywords because they are the names of the default RADIUS and TACACS+ server groups.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure an AAA server group. Currently, the RADIUS and TACACS+ server groups are supported.

Adding an AAA Server Group Member

Command	**server** *ip-addr* [**auth-port** *port1*] [**acct-port** *port2*]
Parameter Description	*ip-addr*: Indicates the IP address of a server. *port1*: Indicates the authentication port of a server. (This parameter is supported only by the RADIUS server group.) *port2*: Indicates the accounting port of a server. (This parameter is supported only by the RADIUS server group.)
Command Mode	Server group configuration mode
Usage Guide	When you add servers to a server group, the default ports are used if you do not specify ports.

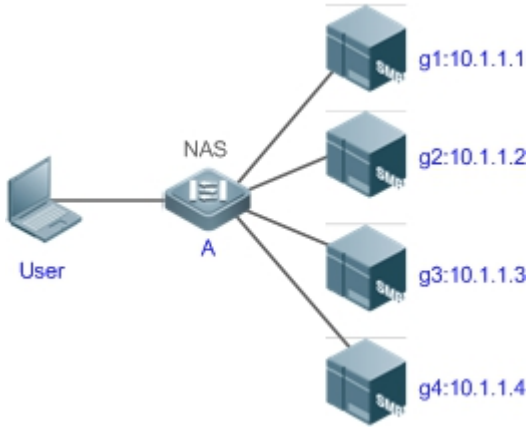
Configuring the VRF Attribute of an AAA Server Group

Command	**ip vrf forwarding** *vrf_name*
Parameter Description	*vrf_name*: Indicates the name of a VRF table.
Command Mode	Server group configuration mode
Usage Guide	Use this command to assign a VRF table to the specified server group.

Configuration Example

Creating an AAA Server Group

Create RADIUS server groups named g1 and g2. The IP addresses of the servers in g1 are 10.1.1.1 and 10.1.1.2, and the IP addresses of the servers in g2 are 10.1.1.3 and 10.1.1.4.

Scenario **Figure 1-14**	
Prerequisites	1. The required interfaces, IP addresses, and VLANs have been configured on the network, network connections have been set up, and the routes from the NAS to servers are reachable. 2. Enable AAA.
Configuration Steps	Step 1: Configure a server (which belongs to the default server group). Step 2: Create user-defined AAA server groups. Step 3: Add servers to the AAA server groups.

NAS	FS#configure terminal FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server host 10.1.1.2 FS(config)#radius-server host 10.1.1.3 FS(config)#radius-server host 10.1.1.4 FS(config)#radius-server key secret FS(config)#aaa group server radius g1 FS(config-gs-radius)#server 10.1.1.1 FS(config-gs-radius)#server 10.1.1.2 FS(config-gs-radius)#exit FS(config)#aaa group server radius g2 FS(config-gs-radius)#server 10.1.1.3 FS(config-gs-radius)#server 10.1.1.4 FS(config-gs-radius)#exit
Verification	Run the **show aaa group** and **show run** commands on the NAS to display the configuration.
NAS	FS#show aaa group Type Reference Name ---------- ---------- ---------- radius 1 radius tacacs+ 1 tacacs+ radius 1 g1 radius 1 g2
	FS#show run ! radius-server host 10.1.1.1 radius-server host 10.1.1.2 radius-server host 10.1.1.3 radius-server host 10.1.1.4 radius-server key secret ! aaa group server radius g1 server 10.1.1.1

```
 server 10.1.1.2
!
aaa group server radius g2
 server 10.1.1.3
 server 10.1.1.4
!
!
```

Common Errors

- For RADIUS servers that use non-default authentication and accounting ports, when you run the **server** command to add servers, specify the authentication or accounting port.

1.4.5 Configuring the Domain-Based AAA Service

Configuration Effect

Create AAA schemes for 802.1X users in different domains.

Notes

About referencing method lists in domains:

- The AAA method lists that you select in domain configuration mode should be defined in advance. If the method lists are not defined in advance, when you select them in domain configuration mode, the system prompts that the configurations do not exist.
- The names of the AAA method lists selected in domain configuration mode must be consistent with those of the method lists defined for the AAA service. If they are inconsistent, the AAA service cannot be properly provided to the users in the domain.

About the default domain:

- Default domain: After the domain-based AAA service is enabled, if a username does not carry domain information, the AAA service is provided to the user based on the default domain. If the domain information carried by the username is not configured in the system, the system determines that the user is unauthorized and will not provide the AAA service to the user. If the default domain is not configured initially, it must be created manually.
- When the domain-based AAA service is enabled, the default domain is not configured by default and needs to be created manually. The default domain name is **default**. It is used to provide the AAA service to the users whose usernames do not carry domain information. If the default domain is not configured, the AAA service is not available for the users whose usernames do not carry domain information.

About domain names:

- The domain names carried by usernames and those configured on the NAS are matched in the longest matching principle. For example, if two domains, **domain.com** and **domain.com.cn** are configured on a NAS and a user sends a request carrying **aaa@domain.com**, the NAS determines that the user belongs to **domain.com**, instead of **domain.com.cn**.
- If the username of an authenticated user carries domain information but the domain is not configured on the NAS, the AAA service is not provided to the user.

Configuration Steps

Enabling AAA

- Mandatory.
- Run the **aaa new-model** command to enable AAA.
- By default, AAA is disabled.

Enabling the Domain-Based AAA Service

- Mandatory.
- Run the **aaa domain enable** command to enable the domain-based AAA service.
- By default, the domain-based AAA service is disabled.

Creating a Domain and Entering Domain Configuration Mode

- Mandatory.
- Run the **aaa domain** command to create a domain or enter the configured domain.
- By default, no domain is configured.

Associating the Domain with an 802.1X Authentication Method List

- Run the **authentication dot1x** command to associate the domain with an 802.1X authentication method list.
- This configuration is mandatory if you need to apply a specified 802.1X authentication method list to the domain.
- Currently, the domain-based AAA service is applicable only to 802.1X access.

Associating the Domain with a Network Accounting Method List

- Run the **accounting network** command to associate the domain with a network accounting method.
- This configuration is mandatory if you need to apply a specified network accounting method list to the domain.
- If a domain is not associated with a network accounting method list, by default, the global default method list is used for accounting.

Associating the Domain with a Network Authorization Method List

- Run the **authorization network** command to associate the domain with a network authorization method list.
- This configuration is mandatory if you need to apply a specified network authorization method list to the domain.
- If a domain is not associated with a network authorization method list, by default, the global default method list is used for authorization.

Configuring the Domain Status

- Optional.
- When a domain is in Block state, the users in the domain cannot log in.
- By default, after a domain is created, its state is Active, indicating that all the users in the domain are allowed to request network services.

Configuring Whether to Contain the Domain Name in Usernames

- Optional.
- By default, the usernames exchanged between the NAS and an authentication server carry domain information.

Configuring the Maximum Number of Domain Users

- Optional.
- By default, the maximum number of access users allowed in a domain is not limited.

Verification

Run the **show aaa domain** command to verify the configuration.

Related Commands

Enabling AAA

Command	**aaa new-model**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	To enable the AAA services, run this command. None of the rest of AAA commands can be effective if AAA is not enabled.

Enabling the Domain-Based AAA Service

Command	**aaa domain enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Use this command to enable the domain-based AAA service.

Creating a Domain and Entering Domain Configuration Mode

Command	**aaa domain** { **default** | *domain-name* }
Parameter Description	**default**: Uses this parameter to configure the default domain. *domain-name*: Indicates the name of the domain to be created.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure a domain to provide the domain-based AAA service. The **default** parameter specifies the default domain. If a username does not carry domain information, the NAS uses the method list associated with the default domain to provide the AAA service to the user. The *domain-name* parameter specifies the name of the domain to be created. If the domain name carried by a username matches the configured domain name, the NAS uses the method list associated with this domain to provide the AAA service to the user. The system supports a maximum of 32 domains.

Associating the Domain with an 802.1X Authentication Method List

Command	**authentication dot1x** { **default** | *list-name* }
Parameter	**default**: Indicates that the default method list is used.

Description	*list-name*: Indicates the name of the method list to be associated.
Command Mode	Domain configuration mode
Usage Guide	Use this command to associate the domain with a 802.1X authentication method list.

Associating the Domain with a Web Authentication Method List

Command	**authentication web-auth** { **default** | *list-name* }
Parameter Description	**default**: Indicates that the default method list is used. *list-name*: Indicates the name of the method list to be associated.
Command Mode	Domain configuration mode
Usage Guide	Use this command to associate the domain with a Web authentication method list.

Associating the Domain with a Network Accounting Method List

Command	**accounting network** { **default |** *list-name* }
Parameter Description	**default**: Indicates that the default method list is used. *list-name*: Indicates the name of the method list to be associated.
Command Mode	Domain configuration mode
Usage Guide	Use this command to associate the domain with a network accounting method list.

Associating the Domain with a Network Authorization Method List

Command	**authorization network** { **default** | *list-name* }
Parameter Description	**default**: Indicates that the default method list is used. *list-name*: Indicates the name of the method list to be associated.
Command Mode	Domain configuration mode
Usage Guide	

Configuring the Domain Status

Command	**state** { **block** | **active** }
Parameter Description	**block**: Indicates that the configured domain is invalid. **active**: Indicates that the configured domain is valid.
Command Mode	Domain configuration mode
Usage Guide	Use this command to make the configured domain valid or invalid.

Configuring Whether to Contain the Domain Name in Usernames

Command	**username-format** { **without-domain** | **with-domain** }
Parameter Description	**without-domain**: Indicates to remove domain information from usernames. **with-domain**: Indicates to keep domain information in usernames.
Command Mode	Domain configuration mode

Usage Guide	Use this command in domain configuration mode to determine whether to include domain information in usernames when the NAS interacts with authentication servers in a specified domain.

Configuring the Maximum Number of Domain Users

Command	**access-limit** *num*
Parameter Description	*num*: Indicates the maximum number of access users allowed in a domain. This limit is applicable only to 802.1X STAs.
Command Mode	Domain configuration mode
Usage Guide	Use this command to limit the number of access users in a domain.

Configuration Example

Configuring the Domain-Based AAA Services

Configure authentication and accounting through a RADIUS server to 802.1X users (username: *user*@domain.com) that access the NAS. The usernames that the NAS sends to the RADIUS server do not carry domain information, and the number of access users is not limited.

Scenario Figure 1-15	
Configuration Steps	The following example shows how to configure RADIUS authentication and accounting, which requires the configuration of a RADIUS server in advance. Step 1: Enable AAA. Step 2: Define an AAA method list. Step 3: Enable the domain-based AAA service. Step 4: Create a domain. Step 5: Associate the domain with the AAA method list. Step 6: Configure the domain attribute.
NAS	FS#configure terminal FS(config)#aaa new-model FS(config)#radius-server host 10.1.1.1 FS(config)#radius-server key test FS(config)#aaa authentication dot1x default group radius FS(config)#aaa accounting network list3 start-stop group radius FS(config)# aaa domain enable FS(config)# aaa domain domain.com FS(config-aaa-domain)# authentication dot1x default FS(config-aaa-domain)# accounting network list3 FS(config-aaa-domain)# username-format without-domain

Verification	Run the **show run** and **show aaa domain** command on the NAS to display the configuration.

NAS	FS#show aaa domain domain.com
	
	==========Domain domain.com==========
	State: Active
	Username format: With-domain
	Access limit: No limit
	802.1X Access statistic: 0
	
	Selected method list:
	  authentication dot1x default
	  accounting network list3

	FS#show run
	
	Building configuration...
	Current configuration : 1449 bytes
	version FSOS 10.4(3) Release(101069)(Wed Oct 20 09:12:40 CST 2010 -ngcf67)
	co-operate enable
	!
	aaa new-model
	aaa domain enable
	!
	aaa domain domain.com
	  authentication dot1x default
	  accounting network list3
	!
	aaa accounting network list3 start-stop group radius
	aaa authentication dot1x default group radius
	!
	nfpp
	!
	no service password-encryption
	!

	radius-server host 10.1.1.1 radius-server key test ! line con 0 line vty 0 4 ! end

Common Errors

N/A

1.5 Monitoring

Clearing

Description	Command
Clears the locked users.	**clear aaa local user lockout** {**all** | **user-name** *username* }

Displaying

Description	Command
Displays the accounting update information.	**show aaa accounting update**
Displays the current domain configuration.	**show aaa domain**
Displays the current lockout configuration.	**show aaa lockout**
Displays the AAA server groups.	**show aaa group**
Displays the AAA method lists.	**show aaa method-list**
Displays the AAA users.	**show aaa user**

2 Configuring RADIUS

2.1 Overview

The Remote Authentication Dial-In User Service (RADIUS) is a distributed client/server system.

RADIUS works with the Authentication, Authorization, and Accounting (AAA) to conduct identity authentication on users who attempt to access a network, to prevent unauthorized access. In FSOS implementation, a RADIUS client runs on a device or Network Access Server (NAS) and transmits identity authentication requests to the central RADIOUS server, where all user identity authentication information and network service information are stored. In addition to the authentication service, the RADIUS server provides authorization and accounting services for access users.

RADIUS is often applied in network environments that have high security requirements and allow the access of remote users. RADIUS is a completely open protocol and the RADIUS server is installed on many operating systems as a component, for example, on UNIX, Windows 2000, and Windows 2008. Therefore, RADIUS is the most widely applied security server currently.

The Dynamic Authorization Extensions to Remote Authentication Dial In User Service is defined in the IETF RFC3576. This protocol defines a user offline management method. Devices communicate with the RADIUS server through the Disconnect-Messages (DMs) to bring authenticated users offline. This protocol implements compatibility between devices of different vendors and the RADIUS server in terms of user offline processing.

In the DM mechanism, the RADIUS server actively initiates a user offline request to a device, the device locates a user according to the user session information, user name, and other information carried in the request and brings the user offline. Then, the device returns a response packet that carries the processing result to the RADIUS server, thereby implementing user offline management of the RADIUS server.

Protocols and Standards

- RFC2865: Remote Authentication Dial In User Service (RADIUS)
- RFC2866: RADIUS Accounting
- RFC2867: RADIUS Accounting Modifications for Tunnel Protocol Support
- RFC2868: RADIUS Attributes for Tunnel Protocol Support
- RFC2869: RADIUS Extensions
- RFC3576: Dynamic Authorization Extensions to Remote Authentication Dial In User Service (RADIUS)

2.2 Applications

Application	Description
Providing Authentication, Authorization, and Accounting Services for Access Users	Authentication, authorization, and accounting are conducted on access users on a network, to prevent unauthorized access or operations.
Forcing Users to Go Offline	The server forces an authenticated user to go offline.

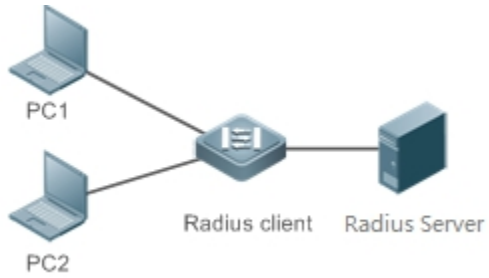
2.2.1 Providing Authentication, Authorization, and Accounting Services for Access Users

Scenario

RADIUS is typically applied in the authentication, authorization, and accounting of access users. A network device serves as a RADIUS client and transmits user information to a RADIUS server. After completing processing, the RADIUS server returns the authentication

acceptance/authentication rejection/accounting response information to the RADIUS client. The RADIUS client performs processing on the access user according to the response from the RADIUS server.

Figure 2- 1 Typical RADIUS Networking Topology



Remarks	PC 1 and PC 2 are connected to the RADIUS client as access users in wired mode, and initiate authentication and accounting requests. The RADIUS client is usually an access switch or aggregate switch. The RADIUS server can be a component built in the Windows 2000/2003, Server (IAS), or UNIX operating system or dedicated server software provided by vendors.

Deployment

- Configure access device information on the RADIUS server, including the IP address and shared key of the access devices.
- Configure the AAA method list on the RADIUS client.
- Configure the RADIUS server information on the RADIUS client, including the IP address and shared key.
- Enable access control on the access port of the RADIUS client.
- Configure the network so that the RADIUS client communicates with the RADIUS server successfully.

2.2.2 Forcing Users to Go Offline

Scenario

The RADIUS server forces authenticated online users to go offline for the sake of management.

See Figure 2- 1 for the networking topology.

Deployment

- Add the following deployment on the basis of 1.2.1 "Deployment".
- Enable the RADIUS dynamic authorization extension function on the RADIUS client.

2.3 Features

Basic Concepts

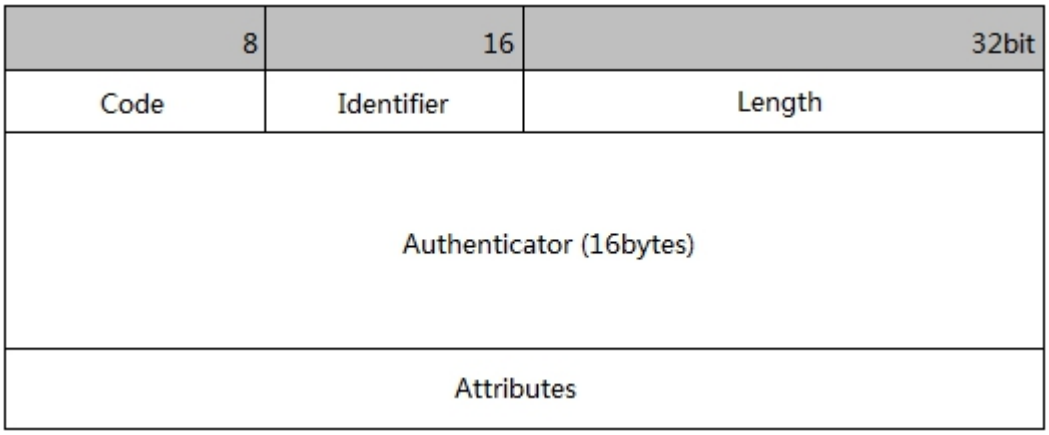
Client/Server Mode

- Client: A RADIUS client initiates RADIUS requests and usually runs on a device or NAS. It transmits user information to the RADIUS server, receives responses from the RADIUS server, and performs processing accordingly. The processing includes accepting user access, rejecting user access, or collecting more user information for the RADIUS server.
- Server: Multiple RADIUS clients map to one RADIUS server. The RADIUS server maintains the IP addresses and shared keys of all

RADIUS clients as well as information on all authenticated users. It receives requests from a RADIUS client, conducts authentication, authorization, and accounting, and returns processing information to the RADIUS client.

Structure of RADIUS Packets

The following figure shows the structure of RADIUS packets.



- Code: Identifies the type of RADIUS packets, which occupies one byte. The following table lists the values and meanings.

Code	Packet Type	Code	Packet Type
1	Access-Request	4	Accounting-Request
2	Access-Accept	5	Accounting-Response
3	Access-Reject	11	Access-Challenge

- Identifier: Indicates the identifier for matching request packets and response packets, which occupies one byte. The identifier values of request packets and response packets of the same type are the same.
- Length: Identifies the length of a whole RADIUS packet, which includes **Code**, **Identifier**, **Length**, **Authenticator**, and **Attributes**. It occupies two bytes. Bytes that are beyond the **Length** field will be truncated. If the length of a received packet is smaller than the value of **Length**, the packet is discarded.
- Authenticator: Verifies response packets of the RADIUS server by a RADIUS client, which occupies 16 bytes. This field is also used for encryption/decryption of user passwords.
- Attributes: Carries authentication, authorization, and accounting information, with the length unfixed. The **Attributes** field usually contains multiple attributes. Each attribute is represented in the Type, Length, Value (TLV) format. Type occupies one byte and indicates the attribute type. The following table lists common attributes of RADIUS authentication, authorization, and accounting. Length occupies one byte and indicates the attribute length, with the unit of bytes. Value indicates the attribute information.

Attribute No.	Attribute Name	Attribute No.	Attribute Name
1	User-Name	43	Acct-Output-Octets
2	User-Password	44	Acct-Session-Id
3	CHAP-Password	45	Acct-Authentic
4	NAS-IP-Address	46	Acct-Session-Time
5	NAS-Port	47	Acct-Input-Packets
6	Service-Type	48	Acct-Output-Packets
7	Framed-Protocol	49	Acct-Terminate-Cause
8	Framed-IP-Address	50	Acct-Multi-Session-Id

Attribute No.	Attribute Name	Attribute No.	Attribute Name
9	Framed-IP-Netmask	51	Acct-Link-Count
10	Framed-Routing	52	Acct-Input-Gigawords
11	Filter-ID	53	Acct-Output-Gigawords
12	Framed-MTU	55	Event-Timestamp
13	Framed-Compression	60	CHAP-Challenge
14	Login-IP-Host	61	NAS-Port-Type
15	Login-Service	62	Port-Limit
16	Login-TCP-Port	63	Login-LAT-Port
18	Reply-Message	64	Tunnel-Type
19	Callback-Number	65	Tunnel-Medium-Type
20	Callback-ID	66	Tunnel-Client-Endpoint
22	Framed-Route	67	Tunnel-Server-Endpoint
23	Framed-IPX-Network	68	Acct-Tunnel-Connection
24	State	69	Tunnel-Password
25	Class	70	ARAP-Password
26	Vendor-Specific	71	ARAP-Features
27	Session-Timeout	72	ARAP-Zone-Access
28	Idle-Timeout	73	ARAP-Security
29	Termination-Action	74	ARAP-Security-Data
30	Called-Station-Id	75	Password-Retry
31	Calling-Station-Id	76	Prompt
32	NAS-Identifier	77	Connect-Info
33	Proxy-State	78	Configuration-Token
34	Login-LAT-Service	79	EAP-Message
35	Login-LAT-Node	80	Message-Authenticator
36	Login-LAT-Group	81	Tunnel-Private-Group-id
37	Framed-AppleTalk-Link	82	Tunnel-Assignment-id
38	Framed-AppleTalk-Network	83	Tunnel-Preference
39	Framed-AppleTalk-Zone	84	ARAP-Challenge-Response
40	Acct-Status-Type	85	Acct-Interim-Interval
41	Acct-Delay-Time	86	Acct-Tunnel-Packets-Lost
42	Acct-Input-Octets	87	NAS-Port-Id

Shared Key

A RADIUS client and a RADIUS server mutually confirm their identities by using a shared key during communication. The shared key cannot be transmitted over a network. In addition, user passwords are encrypted for transmission for the sake of security.

RADIUS Server Group

The RADIUS security protocol, also called RADIUS method, is configured in the form of a RADIUS server group. Each RADIUS method corresponds to one RADIUS server group and one or more RADIUS severs can be added to one RADIUS server group. For details about the RADIUS method, see the *Configuring AAA*. If you add multiple RADIUS servers to one RADIUS server group, when the communication between a device and the first RADIUS server in this group fails or the first RADIUS server becomes unreachable, the device automatically

attempts to communicate with the next RADIUS server till the communication is successful or the communication with all the RADIUS servers fails.

RADIUS Attribute Type

- Standard attributes

The RFC standards specify the RADIUS attribute numbers and attribute content but do not specify the format of some attribute types. Therefore, the format of attribute contents needs to be configured to adapt to different RADIUS server requirements. Currently, the format of the RADIUS Calling-Station-ID attribute (attribute No.: 31) can be configured.

The RADIUS Calling-Station-ID attribute is used to identify user identities when a network device transmits request packets to the RADIUS server. The RADIUS Calling-Station-ID attribute is a string, which can adopt multiple formats. It needs to uniquely identify a user. Therefore, it is often set to the MAC address of a user. For example, when IEEE 802.1X authentication is used, the Calling-Station-ID attribute is set to the MAC address of the device where the IEEE 802.1X client is installed. The following table describes the format of MAC addresses.

Format	Description
Ietf	Indicates the standard format specified in the IETF standard (RFC3580), which is separated by the separator (-). Example: 00-D0-F8-33-22-AC
Normal	Indicates the common format that represents a MAC address (dotted hexadecimal format), which is separated by the separator (.). Example: 00d0.f833.22ac
Unformatted	Indicates the format without separators. This format is used by default. Example: 00d0f83322ac

- Private attributes

RADIUS is an extensible protocol. According to RFC2865, the Vendor-Specific attribute (attribute No.: 26) is used by device vendors to extend the RADIUS protocol to implement private functions or functions that are not defined in the standard RADIUS protocol. Table 1-3 lists private attributes supported by FS products. The **TYPE** column indicates the default configuration of private attributes of FS products and the **Extended TYPE** column indicates the default configuration of private attributes of other non-FS products.

ID	Function	TYPE	Extended TYPE
1	max-down-rate	1	76
2	port-priority	2	77
3	user-ip	3	3
4	vlan-id	4	4
5	last-supplicant-version	5	5
6	net-ip	6	6
7	user-name	7	7
8	password	8	8
9	file-directory	9	9
10	file-count	10	10
11	file-name-0	11	11
12	file-name-1	12	12
13	file-name-2	13	13
14	file-name-3	14	14

ID	Function	TYPE	Extended TYPE
15	file-name-4	15	15
16	max-up-rate	16	16
17	current-supplicant-version	17	17
18	flux-max-high32	18	18
19	flux-max-low32	19	19
20	proxy-avoid	20	20
21	dailup-avoid	21	21
22	ip-privilege	22	22
23	login-privilege	42	42
26	ipv6-multicast-address	79	79
27	ipv4-multicast-address	87	87
62	sdg-type	62	62
85	sdg-zone-name	85	85
103	sdg-group-name	103	103

Overview

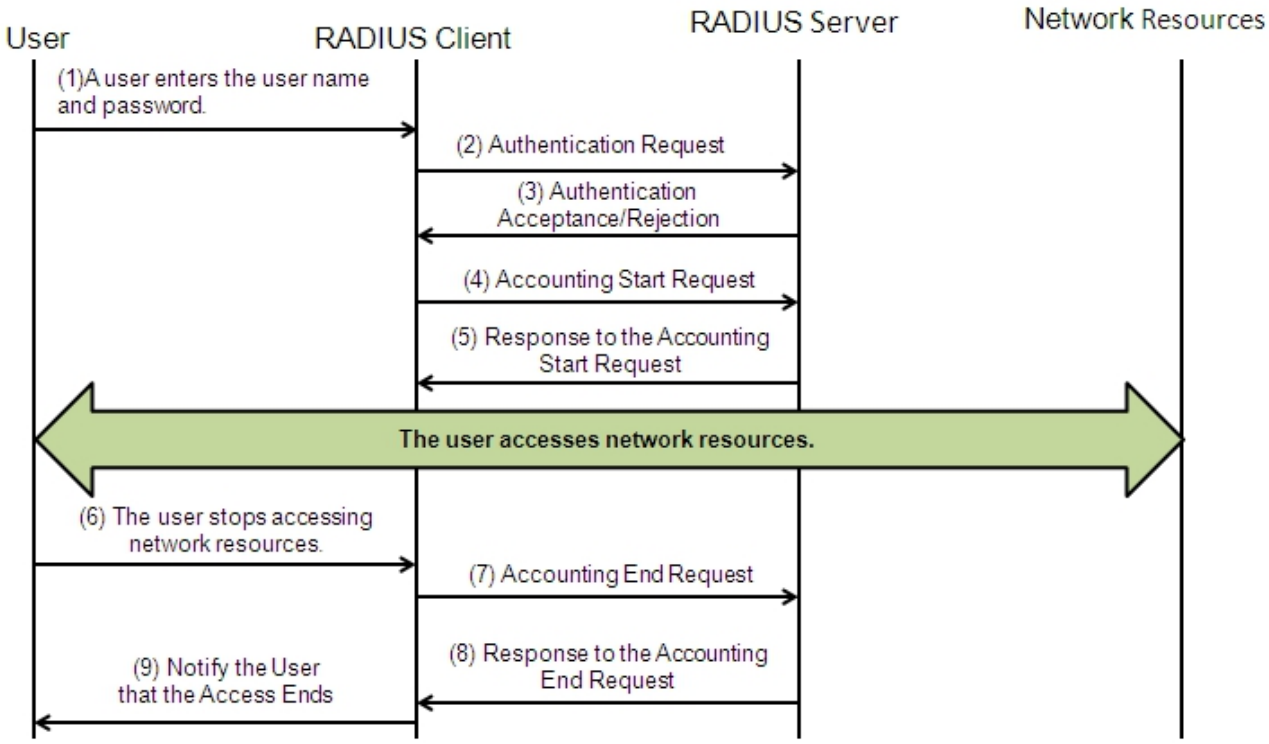
Feature	Description
RADIUS Authentication, Authorization, and Accounting	Conducts identity authentication and accounting on access users, safeguards network security, and facilitates management for network administrators.
Source Address of RADIUS Packets	Specifies the source IP address used by a RADIUS client to transmit packets to a RADIUS server.
RADIUS Timeout Retransmission	Specifies the packet retransmission parameter for a RADIUS client when a RADIUS server does not respond to packets transmitted from the RADIUS client within a period of time.
RADIUS Server Accessibility Detection	Enables a RADIUS client to actively detect whether a RADIUS server is reachable and maintain the accessibility of each RADIUS server. A reachable RADIUS server is selected preferentially to improve the handling performance of RADIUS services.
RADIUS Forced Offline	Enables a RADIUS server to actively force authenticated users to go offline.

2.3.1 RADIUS Authentication, Authorization, and Accounting

Conduct identity authentication and accounting on access users, safeguard network security, and facilitate management for network administrators.

Working Principle

Figure 2- 2



The RADIUS authentication and authorization process is described as follows:

1. A user enters the user name and password and transmits them to the RADIUS client.
2. After receiving the user name and password, the RADIUS client transmits an authentication request packet to the RADIUS server. The password is encrypted for transmission. For the encryption method, see RFC2865.
3. The RADIUS server accepts or rejects the authentication request according to the user name and password. When accepting the authentication request, the RADIUS server also issues authorization information apart from the authentication acceptance information. The authorization information varies with the type of access users.

The RADIUS accounting process is described as follows:

1. If the RADIUS server returns authentication acceptance information in Step (3), the RADIUS client sends an accounting start request packet to the RADIUS server immediately.
2. The RADIUS server returns the accounting start response packet, indicating accounting start.
3. The user stops accessing network resources and requests the RADIUS client to disconnect the network connection.
4. The RADIUS client transmits the accounting end request packet to the RADIUS server.
5. The RADIUS server returns the accounting end response packet, indicating accounting end.
6. The user is disconnected and cannot access network resources.

Related Configuration

Configuring RADIUS Server Parameters

No RADIUS server is configured by default.

You can run the **radius-server host** command to configure a RADIUS server.

At least one RADIUS server must be configured so that RADIUS services run normally.

Configuring the AAA Authentication Method List

No AAA authentication method list is configured by default.

You can run the **aaa authentication** command to configure a method list for different user types and select **group radius** when setting the authentication method.

The RADIUS authentication can be conducted only after the AAA authentication method list of relevant user types is configured.

Configuring the AAA Authorization Method List

No AAA authorization method list is configured by default.

You can run the **aaa authorization** command to configure an authorization method list for different user types and select **group radius** when setting the authorization method.

The RADIUS authorization can be conducted only after the AAA authorization method list of relevant user types is configured.

Configuring the AAA Accounting Method List

No AAA accounting method list is configured by default.

You can run the **aaa accounting** command to configure an accounting method list for different user types and select **group radius** when setting the accounting method.

The RADIUS accounting can be conducted only after the AAA accounting method list of relevant user types is configured.

2.3.2 Source Address of RADIUS Packets

Specify the source IP address used by a RADIUS client to transmit packets to a RADIUS server.

Working Principle

When configuring RADIUS, specify the source IP address to be used by a RADIUS client to transmit RADIUS packets to a RADIUS server, in an effort to reduce the workload of maintaining a large amount of NAS information on the RADIUS server.

Related Configuration

The global routing is used to determine the source address for transmitting RADIUS packets by default.

Run the **ip radius source-interface** command to specify the source interface for transmitting RADIUS packets. The device uses the first IP address of the specified interface as the source address of RADIUS packets.

2.3.3 RADIUS Timeout Retransmission

Working Principle

After a RADIUS client transmits a packet to a RADIUS server, a timer is started to detect the response of the RADIUS server. If the RADIUS server does not respond within a certain period of time, the RADIUS client retransmits the packet.

Related Configuration

Configuring the RADIUS Server Timeout Time

The default timeout time is 5 seconds.

You can run the **radius-server timeout** command to configure the timeout time. The value ranges from 1 second to 1,000 seconds.

The response time of a RADIUS server is relevant to its performance and the network environment. Set an appropriate timeout time according to actual conditions.

Configuring the Retransmission Count

The default retransmission count is 3.

You can run the **radius-server retransmit** command to configure the retransmission count. The value ranges from 1 to 100.

Configuring Whether to Retransmit Accounting Update Packets

Accounting update packets are not retransmitted by default.

You can run the **radius-server account update retransmit** command to configure retransmission of accounting update packets for authenticated users.

2.3.4 RADIUS Server Accessibility Detection

Working Principle

A RADIUS client actively detects whether a RADIUS server is reachable and maintains the accessibility of each RADIUS server. A reachable RADIUS server is selected preferentially to improve the handling performance of RADIUS services.

Related Configuration

Configuring the Criteria for the Device to Judge That a RADIUS Server Is Unreachable

The default criteria configured for judging that a RADIUS server is unreachable meet the two conditions simultaneously: 1. The device does not receive a correct response packet from the RADIUS security server within 60 seconds. 2. The device transmits the request packet to the same RADIUS security server for consecutive 10 times.

You can run the **radius-server dead-criteria** command to configure the criteria for the device to judge that the RADIUS security server is unreachable.

Configuring the Test User Name for Actively Detecting the RADIUS Security Server

No test user name is specified for actively detecting the RADIUS security server by default.

You can run the **radius-server host x.x.x.xtestusername xxx** command to configure the test user name.

2.3.5 RADIUS Forced Offline

Working Principle

Figure 2- 3 DM Message Exchange of the RADIUS Dynamic Authorization Extension Protocol

```
+----------+   Disconnect-Request     +----------+
|          |   <--------------------  |          |
|    NAS   |                          |  RADIUS  |
|          |   Disconnect-Response    |  Server  |
|          |   --------------------->  |          |
+----------+                          +----------+
```

The preceding figure shows the exchange of DM messages between the RADIUS server and the device. The RADIUS server transmits the Disconnect-Request message to UDP Port 3799 of the device. After processing, the device returns the Disconnect-Response message that carries the processing result to the RADIUS server.

Related Configuration

N/A

2.4 Configuration

Configuration	Description and Command	
RADIUS Basic Configuration	(Mandatory) It is used to configure RADIUS authentication, authorization, and accounting.	
	radius-server host	Configures the IP address of the remote RADIUS security server.
	radius-server key	Configures the shared key for communication between the device and the RADIUS server.
	radius-server retransmit	Configures the request transmission count, after which the device confirms that a RADIUS server is unreachable.
	radius-server timeout	Configures the waiting time, after which the device retransmits a request.
	radius-server account update retransmit	Configures retransmission of accounting update packets for authenticated users.
	ip radius source-interface	Configures the source address of RADIUS packets.
Configuring the RADIUS Attribute Type	(Optional) It is used to define attribute processing adopted when the device encapsulates and parses RADIUS packets.	
	radius-serverattribute31	Configures the MAC address format of RADIUS attribute No. 31 (Calling-Station-ID).
	radius-server attribute class	Configures the parsing mode of the RADIUS Class attribute.
	radius attribute	Configures the RADIUS private attribute type.
	radius set qos cos	Sets the private attribute port-priority issued by the server to the COS value of an interface. For COS-relevant concepts, see the *Configuring QoS*.
	radius support cui	Configures the device to support the CUI attribute.
	radius-server authentication attribute	Configures whether RADIUS authentication request packets carry a specified attribute.
	radius-server account attribute	Configures whether RADIUS accounting request packets carry a specified attribute.

Configuration	Description and Command	
	radius-server authentication vendor	Configures whether RADIUS authentication request packets carry the private attributes of other vendors.
	radius-server account vendor	Configures whether RADIUS accounting request packets carry the private attributes of other vendors.
Configuring RADIUS Accessibility Detection	(Optional) It is used to detect whether a RADIUS server is reachable and maintain the accessibility of the RADIUS server.	
	radius-server dead-criteria	Configures the global criteria for judging that a RADIUS security server is unreachable.
	radius-server deadtime	Configures the duration for the device to stop transmitting request packets to an unreachable RADIUS server.
	radius-server host	Configures the IP address of the remote RADIUS security server, authentication port, accounting port, and active detection parameters.

2.4.1 RADIUS Basic Configuration

Configuration Effect

- RADIUS authentication, authorization, and accounting can be conducted after RADIUS basic configuration is complete.

Notes

- Before configuring RADIUS on the device, ensure that the network communication of the RADIUS server is in good condition.
- When running the **ip radius source-interface** command to configure the source address of RADIUS packets, ensure that the device of the source IP address communicates with the RADIUS server successfully.
- When conducting RADIUS IPv6 authentication, ensure that the RADIUS server supports RADIUS IPv6 authentication.

Configuration Steps

Configuring the Remote RADIUS Security Server

- Mandatory.
- Configure the IP address, authentication port, accounting port, and shard key of the RADIUS security server.

Configuring the Shared Key for Communication Between the Device and the RADIUS Server

- Optional.
- Configure a shared key in global configuration mode for servers without a shared key.

The shared key on the device must be consistent with that on the RADIUS server.

Configuring the Request Transmission Count, After Which the Device Confirms That a RADIUS Server Is Unreachable

- Optional.
- Configure the request transmission count, after which the device confirms that a RADIUS server is unreachable, according to the actual network environment.

Configuring the Waiting Time, After which the Device Retransmits a Request

- Optional.
- Configure the waiting time, after which the device retransmits a request, according to the actual network environment.

In an 802.1X authentication environment that uses the RADIUS security protocol, if a network device serves as the 802.1X authenticator and FS SU is used as the 802.1X client software, it is recommended that **radius-server timeout** be set to 3 seconds (the default value is 5 seconds) and **radius-server retransmit** be set to 2 (the default value is 3) on the network device.

Configuring Retransmission of Accounting Update Packets for Authenticated Users

- Optional.
- Determine whether to enable the function of retransmitting accounting update packets of authenticated users according to actual requirements.

Configuring the Source Address of RADIUS Packets

- Optional.
- Configure the source address of RADIUS packets according to the actual network environment.

Verification

- Configure the AAA method list that specifies to conduct authentication, authorization, and accounting on users by using RADIUS.
- Enable the device to interact with the RADIUS server. Conduct packet capture to confirm that the device communicates with the RADIUS server over the RADIUS protocol.

Related Commands

Configuring the Remote RADIUS Security Server

Command	**radius-server host** 【**oob**】 { *ipv4-address* | *ipv6-address* } [**auth-port** *port-number*] [**acct-port** *port-number*] [**test username** *name* [**idle-time** *time*] [**ignore-auth-port**] [**ignore-acct-port**]] [**key** [**0** | **7**] *text-string*]
Parameter Description	*oob : oob authentication, i.e. the source interface is the mgmt port when sending messages to this server.* *ipv4-address*: Indicates the IPv4 address of the RADIUS security server. *ipv6-address*: Indicates the IPv6 address of the RADIUS security server. **auth-port** *port-number*: Indicates the UDP port for RADIUS identity authentication. The value ranges from 0 to 65,535. If it is set to **0**, the host does not conduct identity authentication. **acct-port** *port-number*: Indicates the UDP port for RADIUS accounting. The value ranges from 0 to 65,535. If it is set to **0**, the host does not conduct accounting. **test username** *name*: Enables the function of actively detecting the RADIUS security server and specifies the user name used for active detection. **idle-time** *time*: Indicates the interval for the device to transmit test packets to a reachable RADIUS security server. The default value is 60 minutes. The value ranges from 1 minute to 1,440 minutes (24 hours). **ignore-auth-port**: Disables the function of detecting the authentication port of the RADIUS security server. It is enabled by default. **ignore-acct-port**: Disables the function of detecting the accounting port of the RADIUS security server. It is enabled by default. **key**[**0** | **7**] *text-string* : Configures the shared key of the server. The global shared key is used if it is not configured.
Command Mode	Global configuration mode

Usage Guide	A RADIUS security server must be defined to implement the AAA security service by using RADIUS. You can run the **radius-server host** command to define one or more RADIUS security servers.

Configuring the Shared Key for Communication Between the Device and the RADIUS Server

Command	**radius-server key** [**0** | **7**]*text-string*
Parameter Description	*text-string*: Indicates the text of the shared key. **0 | 7**: Indicates the encryption type of the key. The value **0** indicates no encryption and **7**indicates simple encryption. The default value is **0**.
Command Mode	Global configuration mode
Usage Guide	A shared key is the basis for correct communication between the device and the RADIUS security server. The same shared key must be configured on the device and RADIUS security server so that they can communicate with each other successfully.

Configuring the Request Transmission Count, After Which the Device Confirms That a RADIUS Server Is Unreachable

Command	**radius-server retransmit** *retries*
Parameter Description	*retries*: Indicates the RADIUS retransmission count. The value ranges from 1 to 100.
Command Mode	Global configuration mode
Usage Guide	The prerequisite for AAA to use the next user authentication method is that the current security server used for authentication does not respond. The criteria for the device to judge that a security server does not respond are that the security server does not respond within the RADIUS packet retransmission duration of the specified retransmission count. There is an interval between consecutive two retransmissions.

Configuring the Waiting Time, After which the Device Retransmits a Request

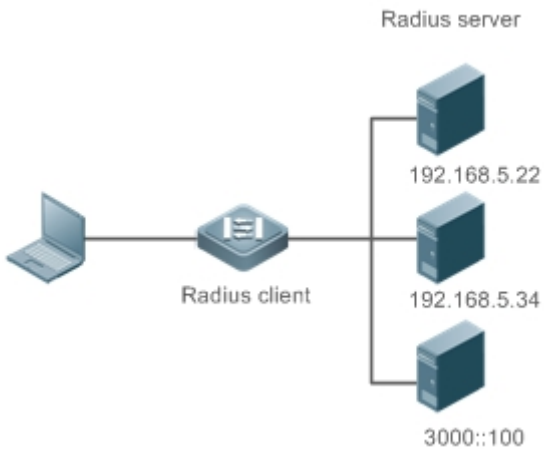
Command	**radius-server timeout** *seconds*
Parameter Description	*seconds*: Indicates the timeout time, with the unit of seconds. The value ranges from 1 second to 1,000 seconds.
Command Mode	Global configuration mode
Usage Guide	Use this command to adjust the packet retransmission timeout time.

Configuring Retransmission of Accounting Update Packets for Authenticated Users

Command	**radius-server account update retransmit**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure retransmission of accounting update packets for authenticated users. Accounting update packets are not retransmitted by default. The configuration does not affect users of other types.

Configuration Example

Using RADIUS Authentication, Authorization, and Accounting for Login Users

Scenario **Figure 2- 4**	
Configuration Steps	● Enable AAA. ● Configure the RADIUS server information. ● Configure to use the RADIUS authentication, authorization, and accounting methods. ● Apply the configured authentication method on the interface.
RADIUS Client	FS#configure terminal FS (config)#aaa new-model FS (config)# radius-server host 192.168.5.22 FS (config)#radius-server host 3000::100 FS (config)# radius-server key aaa FS (config)#aaa authentication login test group radius FS (config)#aaa authorizationexectest group radius FS (config)#aaa accountingexectest start-stop group radius FS (config)# line vty 0 4 FS (config-line)#login authentication test FS (config-line)# authorization exec test FS (config-line)# accounting exec test
Verification	Telnet to a device from a PC. The screen requesting the user name and password is displayed. Enter the correct user name and password to log in to the device. After obtaining a certain access level granted by the server, only run commands under this access level. Display the authentication log of the user on the RADIUS server. Perform management operations on the device as the user and then log out. Display the accounting information on the user on the RADIUS server.
	FS#show running-config ! radius-server host 192.168.5.22

```
radius-server host 3000::100
radius-server key aaa
aaa new-model
aaa accounting exec test start-stop group radius
aaa authorization exec test group radius
aaa authentication login test group radius
no service password-encryption
iptcp not-send-rst
!
vlan 1
!
line con 0
line vty 0 4
accounting exec test
authorization exec test
login authentication test
!
```

Common Errors

- The key configured on the device is inconsistent with that configured on the server.
- No method list is configured.

2.4.2 Configuring the RADIUS Attribute Type

Configuration Effect

- Define the attribute processing adopted when the device encapsulates and parses RADIUS packets.

Notes

- Private attributes involved in "Configuring the RADIUS Attribute Type" refer to FS private attributes.

Configuration Steps

Configuring the MAC Address Format of RADIUS Attribute No. 31 (Calling-Station-ID)

- Optional.
- Set the MAC address format of **Calling-Station-Id** to a type supported by the server.

Configuring the Parsing Mode of the RADIUS Class Attribute

- Optional.

- Configure the parsing mode of the Class attribute according to the server type.

Configuring the RADIUS Private Attribute Type

- Optional.
- If the server is a FS application server, the RADIUS private attribute type needs to be configured.

Setting the Private Attribute port-priority Issued by the Server to the COS Value of an Interface

- Optional.
- Set the private attribute **port-priority** issued by the server to the COS value of an interface as required.

Configures the Device to Support the CUI Attribute

- Optional.
- Configure whether the device supports the RADIUS CUI attribute as required.

Configuring the Mode of Parsing Private Attributes by the Device

- Optional.
- Configure the index of a FS private attribute parsed by the device as required.

Configuring Whether RADIUS Authentication Request Packets Carry a Specified Attribute

- Optional.
- Configure whether to specify the attribute type for RADIUS authentication request packets as required.

Configuring Whether RADIUS Accounting Request Packets Carry a Specified Attribute

- Optional.
- Configure whether to specify the attribute type for RADIUS accounting request packets as required.

Configuring Whether RADIUS Authentication Request Packets Carry the Private Attribute of a Specified Vendor

- Optional.
- Configure whether RADIUS authentication request packets carry the private attribute of a specified vendor as required.

Configuring Whether RADIUS Accounting Request Packets Carry the Private Attribute of a Specified Vendor

- Optional.
- Configure whether RADIUS accounting request packets carry the private attribute of a specified vendor as required.

Configuring Whether RADIUS Server Parses the Private Attribute of Cisco, Huawei or Microsoft

- Optional.
- Configure whether RADIUS server parses the private attribute of Cisco, Huawei or Microsoft.

Configuring the Nas-Port-Id Encapsulation Format for RADIUS Packets

- Optional.
- In either QINQ or non-QINQ scenarios, configure the nas-nort-id encapsulation format for RADIUS packets. By default, the packets

are encapsulated in the normal format.

Verification

- Configure the AAA method list that specifies to conduct authentication, authorization, and accounting on users by using RADIUS.
- Enable the device to interact with the RADIUS server. Conduct packet capture to display the MAC address format of Calling-Station-Id.
- Enable the device to interact with the RADIUS server. Display the debug information of the device to check that FS private attributes are correctly parsed by the device.
- Enable the device to interact with the RADIUS server. Display the debug information of the device to check that the CUI attribute is correctly parsed by the device.

Related Commands

Configuring the MAC Address Format of RADIUS Attribute No. 31 (Calling-Station-ID)

Command	**radius-server attribute 31 mac format** {**ietf** | **normal** | **unformatted** }
Parameter Description	**ietf**: Indicates the standard format specified in the IETF standard (RFC3580), which is separated by the separator (-). Example: 00-D0-F8-33-22-AC. normal: Indicates the common format that represents a MAC address (dotted hexadecimal format), which is separated by the separator (.). Example: 00d0.f833.22ac. **unformatted**: Indicates the format without separators. This format is used by default. Example: 00d0f83322ac.
Command Mode	Global configuration mode
Usage Guide	Some RADIUS security servers (mainly used for 802.1X authentication) can identify only MAC addresses in the IETF format. In this case, set the MAC address format of Calling-Station-ID to IETF.

Configuring the Parsing Mode of the RADIUS Class Attribute

Command	**radius-server attribute class**
Parameter Description	user-flow-control: Configure the speed limit configuration to be parsed from the class attribute. format-16bytes: Configure the format of the speed limit value in the class attribute to be 16 bytes. format-32bytes: Configure the format of the speed limit value in the class attribute to be 32 bytes.
Command Mode	Global configuration mode
Usage Guide	Configure this command if the server needs to issue the rate limit value by using the Class attribute.

Setting the Private Attribute port-priority Issued by the Server to the COS Value of an Interface

Command	**radius set qos cos**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure this command to use the issued QoS value as the CoS value. The QoS value is used as the DSCP value by default.

Configures the Device to Support the CUI Attribute

Command	**radius support cui**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure this command to enable the RADIUS-compliant device to support the CUI attribute.

Configuring Whether RADIUS Authentication Request Packets Carry a Specified Attribute

Command	**radius-server authentication attribute** *type* **package** **radius-server authentication attribute** *type* **unpackage**
Parameter Description	***type***: Indicates the RADIUS attribute type. The value ranges from 1 to 255.
Command Mode	Global configuration mode
Usage Guide	Use this command to specify the attribute to be carried in authentication request packets.

Configuring Whether RADIUS Accounting Request Packets Carry a Specified Attribute

Command	**radius-server account attribute** *type* **package** **radius-server account attribute** *type* **unpackage**
Parameter Description	***type***: Indicates the RADIUS attribute type. The value ranges from 1 to 255.
Command Mode	Global configuration mode
Usage Guide	Use this command to specify the attribute to be carried in accounting request packets.

Configuring Whether RADIUS Authentication Request Packets Carry the Private Attribute of a Specified Vendor

Command	**radius-server authentication vendor** *vendor_name* **package**
Parameter Description	*vendor_name*: Indicates the vendor name. It can be set to **cmcc**, **microsoft**, or **cisco**.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure whether authentication request packets carry the private attribute of a specified vendor.

Configuring Whether RADIUS Accounting Request Packets Carry the Private Attribute of a Specified Vendor

Command	**radius-server account vendor** *vendor_name* **package**
Parameter Description	*vendor_name*: Indicates the vendor name. It can be set to **cmcc**, **Microsoft**, or **cisco**.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure whether accounting request packets carry the private attribute of a specified vendor.

Configuration Example

Configuring the RADIUS Attribute Type

Scenario	One authentication device
Configuration Steps	● Configure the MAC address format of RADIUS Calling-Station-Id. ● Configure the RADIUS private attribute type. ● Set the QoS value issued by the RADIUS server as the COS value of the interface. ● Configure the RADIUS function to support the CUI attribute. ● Configure the device to support private attributes of other vendors. ● Configure authentication requests not to carry the NAS-PORT-ID attribute. ● Configure accounting requests to carry the CMCC private attribute. ● Configure the RAIUDS server not to parse Cisoc's private attributes contained in packets. ● Configure application of the nas-port-id encapsulation format in a QINQ scenario.
	FS(config)#radius-server attribute 31 mac format ietf FS(config)#radius attribute 16 vendor-type 211 FS(config)#radius set qos cos FS(config)#radiussupport cui FS(config)#radius-server authentication attribute 87 unpackage FS(config)#radius-server account vendor cmcc package
Verification	Conduct packet capture or display debug information of the device to check whether the RADIUS standard attributes and private attributes are encapsulated/parsed correctly.

2.4.3 Configuring RADIUS Accessibility Detection

Configuration Effect

The device maintains the accessibility status of each configured RADIUS server: reachable or unreachable. The device will not transmit authentication, authorization, and accounting requests of access users to an unreachable RADIUS server unless all the other servers in the same RADIUS server group as the unreachable server are all unreachable.

The device actively detects a specified RADIUS server. The active detection function is disabled by default. If the active detection function is enabled for a specified RADIUS server, the device will, according to the configuration, periodically transmits detection requests (authentication requests or accounting requests) to the RADIUS server. The transmission interval is as follows:

- For a reachable RADIUS server, the interval is the active detection interval of the reachable RADIUS server (the default value is 60 minutes).
- For an unreachable RADIUS server, the interval is always 1 minute.

Notes

All the following conditions need to be met before the active detection function is enabled for a specified RADIUS server:

- The test user name of the RADIUS server is configured on the device.
- At least one tested port (authentication port or accounting port) of the RADIUS server is configured on the device.

If the following two conditions are all met, it is deemed that a reachable RADIUS server becomes unreachable:

- After the previous correct response is received from the RADIUS server, the time set in **radius-server dead-criteria time** *seconds* has elapsed.
- After the previous correct response is received from the RADIUS server, the count that the device transmits requests to the RADIUS server but fails to receive correct responses (including retransmission) reaches the value set in **radius-server dead-criteria tries** *number*.

If any of the following conditions is met, it is deemed that an unreachable RADIUS server becomes reachable:

- The device receives correct responses from the RADIUS server.
- The duration that the RADIUS server is in the unreachable state exceeds the time set in **radius-server deadtime** and the active detection function is disabled for the RADIUS server.
- The authentication port or accounting port of the RADIUS server is updated on the device.

Configuration Steps

Configuring the Global Criteria for Judging That a RADIUS Security Server Is Unreachable

- Mandatory.
- Configuring the global criteria for judging that a RADIUS security server is unreachable is a prerequisite for enabling the active detection function.

Configuring the IP Address of the Remote RADIUS Security Server, Authentication Port, Accounting Port, and Active Detection Parameters

- Mandatory.
- Configuring active detection parameters of the RADIUS server is a prerequisite for enabling the active detection function.

Configuring the Duration for the Device to Stop Transmitting Request Packets to an Unreachable RADIUS Server

- Optional.
- The configured duration for the device to stop transmitting request packets to an unreachable RADIUS server takes effect only when the active detection function is disabled for the RADIUS server.

Verification

- Run the **show radius server** command to display the accessibility information of each RADIUS server.

Related Commands

Configuring the Global Criteria for Judging That a RADIUS Security Server Is Unreachable

Command	**radius-server dead-criteria** { **time** *seconds* [**tries** *number*] | **tries** *number* }
Parameter Description	**time** *seconds*: Indicates the time condition parameter. If the device fails to receive a correct response packet from a RADIUS security server within the specified time, it is deemed that the RADIUS security server meets the inaccessibility duration condition. The value ranges from 1 second to 120 seconds. **tries** *number*: Indicates the consecutive request timeout count. If the timeout count of request packets transmitted by the device to the same RADIUS security server reaches the preset count, it is deemed that the RADIUS security server meets the consecutive timeout count condition of inaccessibility. The value ranges from 1 to 100.
Command Mode	Global configuration mode

Usage Guide	If a RADIUS security server meets both the duration condition and the consecutive request timeout count condition, it is deemed that the RADIUS security server is unreachable. Users can use this command to adjust parameter values in the duration condition and consecutive request timeout count condition.

Configuring the Duration for the Device to Stop Transmitting Request Packets to an Unreachable RADIUS Server

Command	**Radius-server deadtime** *minutes*
Parameter Description	*minutes*: Indicates the duration for the device to stop transmitting requests to an unreachable RADIUS security server, with the unit of minutes. The value ranges from 1 minute to 1,440 minutes (24 hours).
Command Mode	Global configuration mode
Usage Guide	If the active detection function is enabled for a RADIUS security server on the device, the time parameter in **radius-server deadtime** does not take effect on the RADIUS server. If the active detection function is disabled for a RADIUS security server, the device automatically restores the RADIUS security server to the reachable state when the duration that the RADIUS security server is in the unreachable state exceeds the time specified in **radius-server deadtime**.

Configuration Example

Configuring Accessibility Detection on the RADIUS Server

Scenario **Figure 2- 5**	
Configuration Steps	● Configure the global criteria for judging that a RADIUS security server is unreachable. ● Configure the IP address of the remote RADIUS security server, authentication port, accounting port, and active detection parameters.
RADIUS Client	FS(config)#radius-server dead-criteria time120 tries 5 FS(config)# radius-server host 192.168.5.22 test username test ignore-acct-port idle-time 90
Verification	Disconnect the network communication between the device and the server with the IP address of 192.168.5.22.Conduct RADIUS authentication through the device. After 120 seconds, run the **show radius server** command to check that the server state is **dead**.
	FS#show running-config … radius-server host 192.168.5.22 test username test ignore-acct-port idle-time 90 radius-server dead-criteria time 120 tries 5 …

2.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears statistics of the RADIUS dynamic authorization extension function and restarts statistics.	**clear radius dynamic-authorization-extension statistics**

Displaying

Description	Command
Displays global parameters of the RADIUS server.	**show radius parameter**
Displays the configuration of the RADIUS server.	**show radius server**
Displays statistics relevant to the RADIUS dynamic authorization extension function.	**show radius dynamic-authorization-extension statistics**
Displays statistics relevant to RADIUS authentication.	**show radius auth statistics**
Displays statistics relevant to RADIUS accounting.	**show radius acct statistics**
Displays configuration of RADIUS server groups.	**show radius group**
Displays RADIUS standard attributes.	**show radius attribute**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the RADIUS event.	**debugradiusevent**
Debugs RADIUS packet printing.	**debugradiusdetail**
Debugs the RADIUS dynamic authorization extension function.	**debug radiusextension event**
Debugs the RADIUS dynamic authorization extension packet printing.	**debug radius extension detail**

3 Configuring TACACS+

3.1 Overview

TACACS+ is a security protocol enhanced in functions based on the Terminal Access Controller Access Control System (TACACS) protocol. It is used to implement the authentication, authorization, and accounting (AAA) of multiple users.

Protocols and Standards

- RFC 1492 Terminal Access Controller Access Control System

3.2 Applications

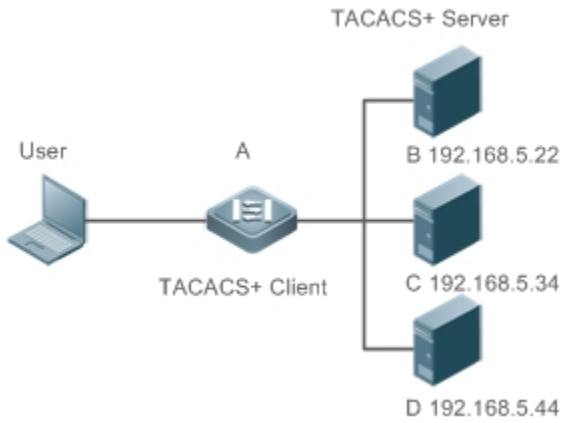
Application	Description
Managing and Controlling Login of End Users	Password verification and authorization need to be conducted on end users.

3.2.1 Managing and Controlling Login of End Users

Scenario

TACACS+ is typically applied in the login management and control of end users. A network device serves as the TACACS+ client and sends a user name and password to the TACACS+ server for verification. The user is allowed to log in to the network device and perform operations after passing the verification and obtaining authorization. See the following figure.

Figure 3- 1



Remarks	● A is a client that initiates TACACS+ requests. ● B, C, and D are servers that process TACACS+ requests.

Deployment

- Start the TACACS+ server on Server B, Server C, and Server D, and configure information on the access device (Device A) so that the servers provide TACACS+-based AAA function for the access device. Enable the AAA function on Device A to start authentication for the user login.
- Enable the TACACS+ client function on Device A, add the IP addresses of the TACACS+ servers (Server B, Server C, and Server D) and the shared key so that Device A communicates with the TACACS+ servers over TACACS+ to implement the AAA function.

3.3 Features

Basic Concepts

Format of TACACS+ Packets

Figure 3- 2

4	8	16	24	32 bit
Major	Minor	Packet type	Sequence no.	Flags
Session ID				
Length				

- Major Version: Indicates the major TACACS+ version number.
- Minor Version: Indicates the minor TACACS+ version number.
- Packet Type: Indicates the type of packets, with the options including:
  TAC_PLUS_AUTHEN: = 0x01 (authentication);
  TAC_PLUS_AUTHOR: = 0x02 (authorization);
  TAC_PLUS_ACCT: = 0x03 (accounting)
- Sequence Number: Indicates the sequence number of a data packet in the current session. The sequence number of the first TACACS+ data packet in a session must be 1 and the sequence number of subsequent each data packet increases by one. Therefore, the client sends data packets only with an odd sequence number and TACACS+ Daemon sends packets only with an even sequence number.
- Flags: Contains various bitmap format flags. One of the bits in the value specifies whether data packets need to be encrypted.
- Session ID: Indicates the ID of a TACACS+ session.
- Length: Indicates the body length of a TACACS+ data packet (excluding the header). Packets are encrypted for transmission on a network.

Overview

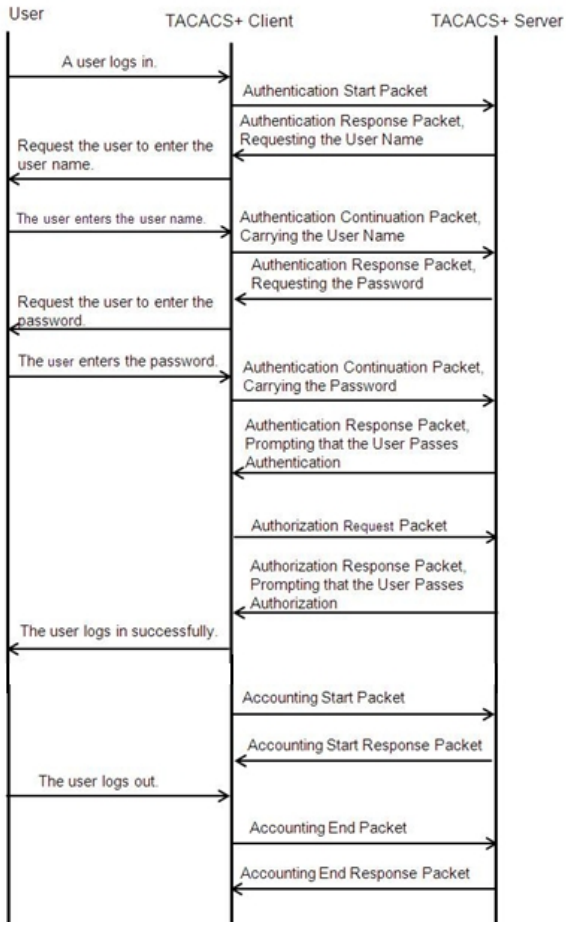
Feature	Description
TACACS+ Authentication, Authorization, and Accounting	Conducts authentication, authorization, and accounting on end users.

3.3.1 TACACS+ Authentication, Authorization, and Accounting

Working Principle

The following figure uses basic authentication, authorization, and accounting of user login to describe interaction of TACACS+ data packets.

Figure 3- 3



The entire basic message interaction process includes three sections:

1. The authentication process is described as follows:

    1) A user requests to log in to a network device.

    2) After receiving the request, the TACACS+ client sends an authentication start packet to the TACACS+ server.

    3) The TACACS+ server returns an authentication response packet, requesting the user name.

    4) The TACACS+ client requests the user to enter the user name.

    5) The user enters the login user name.

    6) After receiving the user name, the TACACS+ client sends an authentication continuation packet that carries the user name to the TACACS+ server.

    7) The TACACS+ server returns an authentication response packet, requesting the login password.

    8) The TACACS+ client requests the user to enter the login password.

    9) The user enters the login password.

    10) After receiving the login password, the TACACS+ client sends an authentication continuation packet that carries the login password to the TACACS+ server.

    11) The TACACS+ server returns an authentication response packet, prompting that the user passes authentication.

2. The user authorization starts after successful authentication:

    1) The TACACS+ client sends an authorization request packet to the TACACS+ server.

2) The TACACS+ server returns an authorization response packet, prompting that the user passes authorization.

3) After receiving the authorization success packet, the TACACS+ client outputs the network device configuration screen for the user.

3. Accounting and audit need to be conducted on the login user after successful authorization:

1) The TACACS+ client sends an accounting start packet to the TACACS+ server.

2) The TACACS+ server returns an accounting response packet, prompting that the accounting start packet has been received.

3) The user logs out.

4) The TACACS+ client sends an accounting end packet to the TACACS+ server.

5) The TACACS+ server returns an accounting response packet, prompting that the accounting end packet has been received.

3.4 Configuration

Configuration	Description and Command	
Configuring TACACS+ Basic Functions	(Mandatory) It is used to enable the TACACS+ security service.	
	tacacs-server host	Configures the TACACS+ server.
	tacacs-server key	Specifies the key shared by the server and network device.
	tacacs-server timeout	Configures the global waiting timeout time of the TACACS+ server for communication between a network device and the TACACS+ server.
Configuring Separate Processing of Authentication, Authorization, and Accounting of TACACS+	(Optional) It is used to separately process authentication, authorization, and accounting requests.	
	aaa group server tacacs+	Configures TACACS+ server groups and divides TACACS+ servers into different groups.
	server	Adds servers to TACACS+ server groups.

3.4.1 Configuring TACACS+ Basic Functions

Configuration Effect

- The TACACS+ basic functions are available after the configuration is complete. When configuring the AAA method list, specify the method of using TACACS+ to implement TACACS+ authentication, authorization, and accounting.
- When authentication, authorization, and accounting operations are performed, TACACS+ initiates the authentication, authorization, and accounting requests to configured TACACS+ servers according to the configured sequence. If response timeout occurs on a TACACS+ server, TACACS+ traverses the TACACS+ server list in sequence.

Notes

- The TACACS+ security service is a type of AAA service. You need to run the **aaa new-model** command to enable the security service.
- Only one security service is provided after TACACS+ basic functions are configured. To make the TACACS+ functions take effect, specify the TACACS+ service when configuring the AAA method list.

Configuration Steps

Enabling AAA

- Mandatory. The AAA method list can be configured only after AAA is enabled. TACACS+ provides services according to the AAA method list.

Command	**aaa new-model**
Parameter Description	N/A
Defaults	The AAA function is disabled.
Command Mode	Global configuration mode
Usage Guide	The AAA method list can be configured only after AAA is enabled. TACACS+ provides services according to the AAA method list.

Configuring the IP Address of the TACACS+ Server

- Mandatory. Otherwise, a device cannot communicate with the TACACS+ server to implement the AAA function.

Command	**tacacs-server host [oob]** {*ipv4-address* | *ipv6-address*} [**port** *integer*] [**timeout** *integer*] [**key** [**0** | **7**] *text-string*]
Parameter Description	*oob: TACACS+ communication uses mgmt port as source interface, default is to use non-mgmt port for communication.* *ipv4-address*: Indicates the IPv4 address of the TACACS+ server. *ipv6-address*: Indicates the IPv6 address of the TACACS+ server. **port** *integer*: Indicates the TCP port used for TACACS+ communication. The default TCP port is 49. **timeout** *integer*: Indicates the timeout time of the communication with the TACACS+ server. The global timeout time is used by default. **key** [**0** | **7**] *text-string*: Indicates the shared key of the server. The global key is used if it is not configured. An encryption type can be specified for the configured key. The value **0** indicates no encryption and **7** indicates simple encryption. The default value is **0**.
Defaults	No TACACS+ server is configured.
Command Mode	Global configuration mode
Usage Guide	1. You can specify the shared key of the server when configuring the IP address of the server. If no shared key is specified, the global key configured using the **tacacs-server key** command is used as the shared key of the server. The shared key must be completely the same as that configured on the server. 2. You can specify the communication port of the server when configuring the IP address. 3. You can specify the communication timeout time of the server when configuring the IP address.

Configuring the Shared Key of the TACACS+ Server

- Optional.
- If no global communication protocol is configured using this command, set **key** to specify the shared key of the server when running the **tacacs-server host** command to add server information. Otherwise, a device cannot communicate with the TACACS+ server.
- If no shared key is specified by using **key** when you run the **tacacs-server host** command to add server information, the global key is used.

Command	**tacacs-server** [**key** [**0** | **7**] *text-string*]

Parameter Description	*text-string*: Indicates the text of the shared key. **0 | 7**: Indicates the encryption type of the key. The value **0** indicates no encryption and **7** indicates simple encryption.
Defaults	No shared key is configured for any TACACS+ server.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure a global shared key for servers. To specify a different key for each server, set **key** when running the **tacacs-server host** command.

Configuring the Timeout Time of the TACACS+ Server

- Optional.
- You can set the timeout time to a large value when the link between the device and the server is unstable.

Command	**tacacs-server timeout** *seconds*
Parameter Description	*seconds*: Indicates the timeout time, with the unit of seconds. The value ranges from 1 second to 1,000 seconds.
Defaults	The default value is 5 seconds.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the global server response timeout time. To set different timeout time for each server, set **timeout** when running the **tacacs-server host** command.

Verification

Configure the AAA method list that specifies to conduct authentication, authorization, and accounting on users by using TACACS+.

- Enable the device to interact with the TACACS+ server and conduct packet capture to check the TACACS+ interaction process between the device and the TACACS+ server.
- View server logs to check whether the authentication, authorization, and accounting are normal.

Configuration Example

Using TACACS+ for Login Authentication

Scenario **Figure 3- 4**	
Remarks	- A is a client that initiates TACACS+ requests. - B is a server that processes TACACS+ requests.
Configuration Steps	- Enable AAA. - Configure the TACACS+ server information.

	● Configure the method of using TACACS+ for authentication. ● Apply the configured authentication method on an interface.
A	FS# configure terminal FS(config)# aaa new-model FS(config)# tacacs-server host 192.168.5.22 FS(config)# tacacs-server key aaa FS(config)# aaa authentication login test group tacacs+ FS(config)# line vty 0 4 FS(config-line)# login authentication test
Verification	Telnet to a device from a PC. The screen requesting the user name and password is displayed. Enter the correct user name and password to log in to the device. View the authentication log of the user on the TACACS+ server.

Common Errors

- The AAA security service is disabled.
- The key configured on the device is inconsistent with the key configured on the server.
- No method list is configured.

3.4.2 Configuring Separate Processing of Authentication, Authorization, and Accounting of TACACS+

Configuration Effect

- The authentication, authorization, and accounting in the security service are processed by different TACACS+ servers, which improves security and achieves load balancing to a certain extent.

Notes

- The TACACS+ security service is a type of AAA service. You need to run the **aaa new-model** command to enable the security service.
- Only one security service is provided after TACACS+ basic functions are configured. To make the TACACS+ functions take effect, specify the TACACS+ service when configuring the AAA method list.

Configuration Steps

Configuring TACACS+ Server Groups

- Mandatory. There is only one TACACS+ server group by default, which cannot implement separate processing of authentication, authorization, and accounting.
- Three TACACS+ server groups need to be configured for separately processing authentication, authorization, and accounting.

Command	**aaa group server tacacs+***group-name*
Parameter Description	*group-name*: Indicates the name of a group. A group name cannot be radius or tacacs+, which are the names of embedded groups.
Defaults	No TACACS+ server group is configured.

Command Mode	Global configuration mode
Usage Guide	Group TACACS+ servers so that authentication, authorization, and accounting are completed by different server groups.

Adding Servers to TACACS+ Server Groups

- Mandatory. If no server is added to a server group, a device cannot communicate with TACACS+ servers.
- In server group configuration mode, add the servers that are configured using the **tacacs-server host** command.

Command	**server** {*ipv4-address* | *ipv6-address*}
Parameter Description	*ipv4-address*: Indicates the IPv4 address of the TACACS+ server. *ipv6-address*: Indicates the IPv6 address of the TACACS+ server.
Defaults	No server is configured.
Command Mode	TACACS+ server group configuration mode
Usage Guide	Before configuring this command, you must run the **aaa group server tacacs+** command to enter the TACACS+ server group configuration mode. For the address of a server configured in a TACACS+ server group, the server must be configured using the **tacacs-server host** command in global configuration mode. If multiple servers are added to one server group, when one server does not respond, the device continues to send a TACACS+ request to another server in the server group.

Configuring VRF of a TACACS+ Server Group

- Optional. Configure Virtual Routing and Forwarding (VRF) if a device needs to send TACACS+ packets through a specified address.
- In server group configuration mode, use a configured VRF name to specify the routing for the communication of servers in this group.

Command	**ip vrf forwarding** *vrf-name*
Parameter Description	*vrf-name*: Indicates the VRF name.
Defaults	No VRF is specified by default.
Command Mode	TACACS+ server group configuration mode
Usage Guide	Before configuring this command, you must run the **aaa group server tacacs+** command to enter the TACACS+ server group configuration mode. For VRF configured in a TACACS+ server group, a valid name must be configured for VRF by using the **vrf definition** command in global configuration mode.

Configuring oob of a TACACS+ Server Group

- Optional. Configure oob if a device needs to send TACACS+ packets through a specified MGMT port.
- In server group configuration mode, specify routing for the communication of servers in the group.

Command	**ip oob** **ip oob via** *mgmt.-name* **ip vrf forwarding** *vrf-name*

Parameter Description	**ip oob**: Indicates the MGMT0 port. *mgmt.-name*: Name of management port. vrf-name: Indicates the VRF name.
Defaults	No oob is specified by default.
Command Mode	TACACS+ server group configuration mode
Usage Guide	Before configuring this command, you must run the **aaa group server tacacs+** command to enter the TACACS+ server group configuration mode. If no MGMT port is specified, the MGMT0 port is used by default.

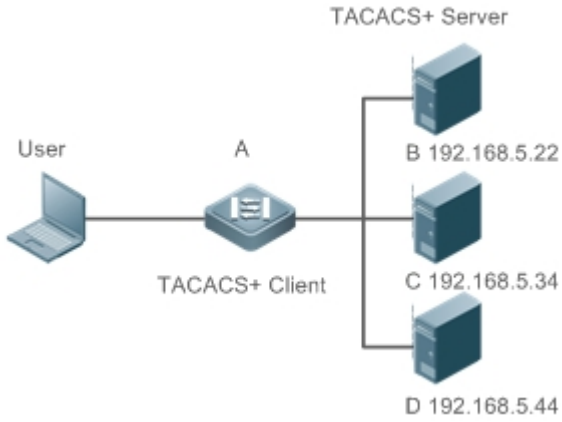
Verification

Configure the AAA method list that specifies to conduct authentication, authorization, and accounting on users by using TACACS+.

- Enable a device to interact with TACACS+ servers. Conduct packet capture, check that the authentication, authorization, and accounting packets are interacted with different servers, and check the source addresses in packets.

Configuration Example

Configuring Different TACACS+ Server Groups for Separately Processing Authentication, Authorization, and Accounting

Scenario **Figure 3- 5**	
Remarks	- A is a client that initiates TACACS+ requests. - B is a server that processes TACACS+ authentication requests. - C is a server that processes TACACS+ authorization requests. - D is a server that processes TACACS+ accounting requests.
Configuration Steps	- Enable AAA. - Configure the TACACS+ server information. - Configure TACACS+ server groups. - Add servers to TACACS+ server groups. - Configure the method of using TACACS+ for authentication. - Configure the method of using TACACS+ for authorization. - Configure the method of using TACACS+ for accounting.

	● Apply the configured authentication method on an interface. ● Apply the configured authorization method on an interface. ● Apply the configured accounting method on an interface.
	FS# configure terminal FS(FS(config)# aaa new-model FS(config)# tacacs-server host 192.168.5.22 FS(config)# tacacs-server host 192.168.5.34 FS(config)# tacacs-server host 192.168.5.44 FS(config)# tacacs-server key aaa FS(config)# aaa group server tacacs+ tacgrp1 FS(config-gs-tacacs)# server 192.168.5.22 FS(config-gs-tacacs)# exit FS(config)# aaa group server tacacs+ tacgrp2 FS(config-gs-tacacs)# server 192.168.5.34 FS(config-gs-tacacs)# exit FS(config)# aaa group server tacacs+ tacgrp3 FS(config-gs-tacacs)# server 192.168.5.44 FS(config-gs-tacacs)# exit FS(config)# aaa authentication login test1 group tacacs+ FS(config)# aaa authentication enable default group tacgrp1 FS(config)# aaa authorization exec test2 group tacgrp2 FS(config)# aaa accounting commands 15 test3 start-stop group tacgrp3 FS(config)# line vty 0 4 FS(config-line)# login authentication test1 FS(config-line)#authorization exec test2 FS(config-line)# accounting commands 15 test3
Verification	Telnet to a device from a PC. The screen requesting the user name and password is displayed. Enter the correct user name and password to log in to the device. Enter the **enable** command and enter the correct **enable** password to initiate **enable** authentication. Enter the privilege EXEC mode after passing the authentication. Perform operations on the device and then exit the device. View the authentication log of the user on the server with the IP address of 192.168.5.22. View the **enable** authentication log of the user on the server with the IP address of 192.168.5.22. View the **exec** authorization log of the user on the server with the IP address of 192.168.5.34. View the command accounting log of the user on the server with the IP address of 192.168.5.44.

Common Errors

- The AAA security service is disabled.
- The key configured on the device is inconsistent with the key configured on the server.
- Undefined servers are added to a server group.
- No method list is configured.

3.5 Monitoring

Displaying

Description	Command
Displays interaction with each TACACS+ server.	**show tacacs**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs TACACS+.	**debug tacacs+**

4 Configuring 802.1X

4.1 Overview

IEEE 802.1X is a standard for port-based network access control that provides secure access service for local area networks (LANs).

In IEEE 802-compliant LANs, users connecting to the network access devices (NASs) can access network resources without authentication and authorization, bringing security risks to the network. IEEE 802.1X was proposed to resolve security problems of such LANs.

802.1X supports three security applications: authentication, authorization, and accounting, which are called AAA.

- Authentication: Checks whether to allow user access and restricts unauthorized users.
- Authorization: Grants specified services to users and controls permissions of authorized users.
- Accounting: Records network resource status of users to provide statistics for charges.

802.1X can be deployed in a network to realize user authentication, authorization and other functions.

Protocols and Standards

- IEEE 802.1X: Port-Based Network Access Control

4.2 Applications

Application	Description
Wired 802.1X Authentication	To ensure secure admission on the campus network, 802.1X authentication is deployed on access switches.

4.2.1 Wired 802.1X Authentication

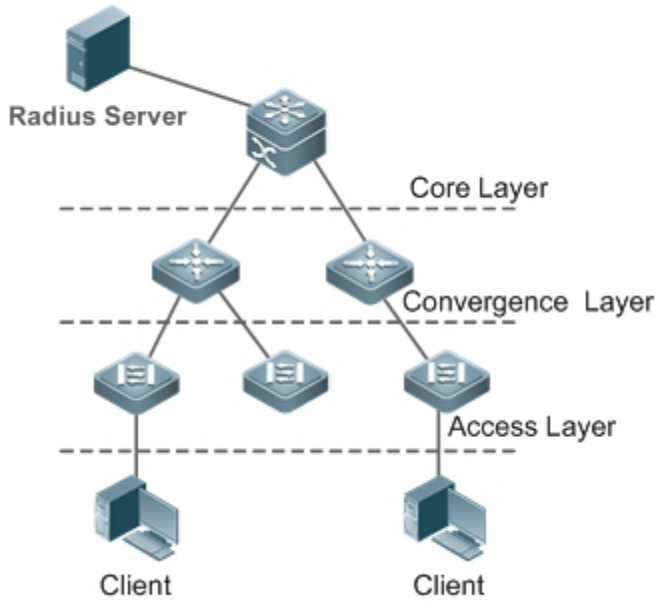
Scenario

The campus network is deployed at the access, convergence, and core layers. 802.1X is deployed on access switches connected to dormitories to perform secure admission. Dormitory users must pass 802.1X authentication before accessing the campus network.

As shown in Figure 4-1:

- User ends must be installed with 802.1X clients (which can come with the operating system, or others like FS Supplicant).
- Access switches support 802.1X.
- One or multiple Remote Authentication Dial-In User Service (RADIUS) servers perform authentication.

Figure 4-1



Remarks	The supplicant software installed on the user ends (or software coming with the operating system) performs 802.1X authentication. 802.1X authentication is deployed on access switches, convergence switches, or core switches. The RADIUS server runs the RADIUS server software to perform identity verification.

Deployment

- Enable 802.1X authentication on ports between access switches and users to make ports controllable. Only authenticated users on one port can access the network.
- Configure an AAA authentication method list so that 802.1X can adopt the appropriate method and authentication server.
- Configure RADIUS parameters to ensure proper communication between a switch and the RADIUS server. For details, see the *Configuring RDS*.
- If a FS RADIUS server is used, configure SNMP parameters to allow the RADIUS server to manage devices, such as querying and setting.
- Configure the port between the access switch and the RADIUS server as an uncontrolled port to ensure proper communication between them.
- Create an account on the RADIUS server, register the IP address of an access switch, and configure RADIUS-related parameters. Only in this case, can the RADIUS server respond to the requests of the switch.

4.2.2 MAB Auto Authentication

Scenario

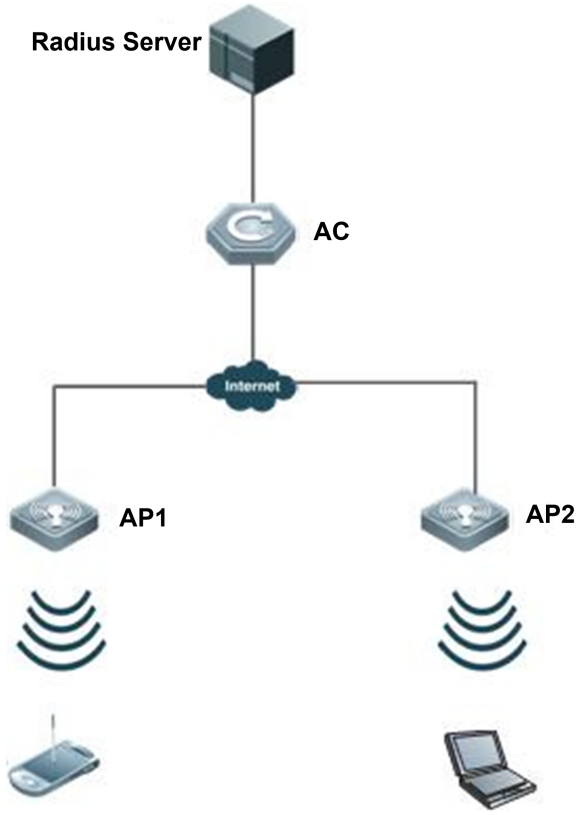
MAC address bypass (MAB) auto authentication indicates that MAB authentication is performed together with Web authentication. In the original wireless Web authentication scenario, it is complained that the ease-to-use performance of Web authentication is poor. During each Web authentication, a user needs to associate the STA with an SSID, open the browser, and enter the user name and password. In addition, if the STA drops out of the network, the STA cannot automatically access the network again. To ensure that all Web

authenticated STAs are always online and access the network imperceptibly, MAB auto authentication is proposed. After a STA passes Web authentication, the STA can access the network again imperceptibly without Web authentication.

As shown in Figure 4- 1:

- Only the browser is mandatory on the client.
- The AC supports Web authentication and MAB authentication.
- One or multiple RADIUS servers provide authentication. In addition, the authentication server supports the authentication mode of using the MAC address as the user name and password.

Figure 4-2



Remarks	Wireless MAB authentication is triggered by a STA advertisement. When a STA is already online, MAB authentication will not be triggered again. If MAB authentication fails, it can be triggered again only after the STA goes offline and reconnects to the network.

Deployment

- Enable Web authentication, DOT1X authentication, and MAB authentication on the interface of the AC. MAB authentication can be performed only after DOT1X authentication is enabled. (For details about MAB authentication, see section 4.4.5 "Configuring MAB Auto Authentication". For details about Web authentication, see the WEB-AUTH-SCG document.)
- Configure an AAA authentication method list, so that a correct method and authentication server can be used for MAB/Web authentication. (For details about the AAA authentication method list configuration, see the AAA-SCG document.)
- Configure RADIUS parameters to ensure proper communication between the AC and the RADIUS server. In addition, configure the RADIUS server to support the authentication mode of using the MAC address as the user name and password. For details about the

RADIUS configuration, see the corresponding configuration guide.

- If a FS RADIUS server is used, configure SNMP parameters to allow the RADIUS server to perform operations such as querying and setting on the AP.
- Create an account on the RADIUS server, register the IP address of the AC, and configure RADIUS-related parameters. The RADIUS server can respond to the requests of the AP and AC only after the foregoing settings are completed.

4.3 Features

Basic Concepts

User

In wired environment, 802.1X is a LAN-based protocol. It identifies users based on physical information but not accounts. In a LAN, a user is identified by the MAC address and VLAN ID (VID). Except them, all other information such as the account ID and IP address can be changed.

RADIUS

RADIUS is a remote authentication protocol defined in RFC2865, which get wide practice. Using this protocol, the authentication server can remotely deploy and perform authentication. During 802.1X deployment, the authentication server is remotely deployed, and 802.1X authentication information between the NAS and the authentication server is transmitted through RADIUS.

Timeout

During authentication, an NAS needs to communicate with the authentication client and server. If the authentication client or server times out, not responding within the time specified by 802.1X, authentication will fail. During deployment, ensure that the timeout specified by 802.1X is longer than that specified by RADIUS.

MAB

MAC address bypass (MAB) authentication means that the MAC address is used as the user name and password for authentication. Since FS Supplicant cannot be installed on some dumb ends such as network printers, use MAB to perform security control.

EAP

802.1X uses Extensible Authentication Protocol (EAP) to carry authentication information. Defined in RFC3748, EAP provides a universal authentication framework, in which multiple authentication modes are embedded, including Message Digest Algorithm 5 (MD5), Challenge Handshake Authentication Protocol (CHAP), Password Authentication Protocol (PAP), and Transport Layer Security (TLS). FS 802.1X authentication supports various modes including MD5, CHAP, PAP, PEAP-MSCHAP, and TLS.

Authorization

Authorization means to bind specified services to authenticated users, such as IP address, VLAN, Access Control List (ACL), and Quality of Service (QoS).

Accounting

Accounting performs network audit on network usage duration and traffic for users, which facilitates network operation, maintenance, and management.

Overview

Feature	Description
Authentication	Provides secure admission for users. Only authenticated users can access the network.
Authorization	Grants network access rights to authenticated users, such as IP address binding and ACL binding
Accounting	Provides online record audit, such as online duration and traffic.

4.3.1 Authentication

Authentication aims to check whether users are authorized and prevent unauthorized users from accessing the network. Users must pass authentication to obtain the network access permission. They can access the network only after the authentication server verifies the account. Before user authentication succeeds, only EAPOL packets (Extensible Authentication Protocol over LAN, 802.1X packets) can be transmitted over the network for authentication.

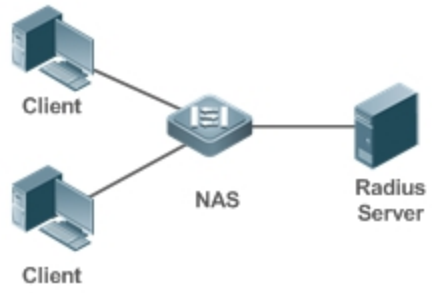
Working Principle

802.1X authentication is very simple. After a user submits its account information, the NAS sends the account information to the remote RADIUS server for identity authentication. If the authentication succeeds, the user can access the network.

Roles in Authentication

802.1X authentication involves three roles: supplicant, authenticator, and server. In real applications, their respective roles are client, network access server (NAS), and authentication server (mostly RADIUS server).

Figure 4-3



- Supplicant

The supplicant is the role of end users, usually a PC. It requests to access network services and replies to the request packets of the authenticator. The supplicant must run software compliant with the 802.1X standard. Except the typical 802.1X client support embedded in the operating system, FS has launched a FS Supplicant compliant with the 802.1X standard.

- Authenticator

The authenticator is usually an NAS such as a switch access hotspot. It controls the network connection of a client based on the client's authentication status. As a proxy between the client and the authentication server, the authenticator requests the user name from the client, verifies the authentication information from the authentication server, and forwards it to the client. Except as the 802.1X authenticator, the so-called NAS also acts as a RADIUS Client. It encapsulates the replies of the client into the RADIUS-format packets and forwards the packets to the RADIUS server. After receiving the information from the RADIUS server, it interprets the information and forwards it to the client.

The authenticator has two types of ports: controlled port and uncontrolled port. Users connected to controlled ports can access network

resources only when authenticated. Users connected to uncontrolled ports can directly access network resources without authentication. We can connect users to controlled ports to control users. Uncontrolled ports are mainly used to connect the authentication server to ensure proper communication between the authentication server and the NAS.

- Authentication server

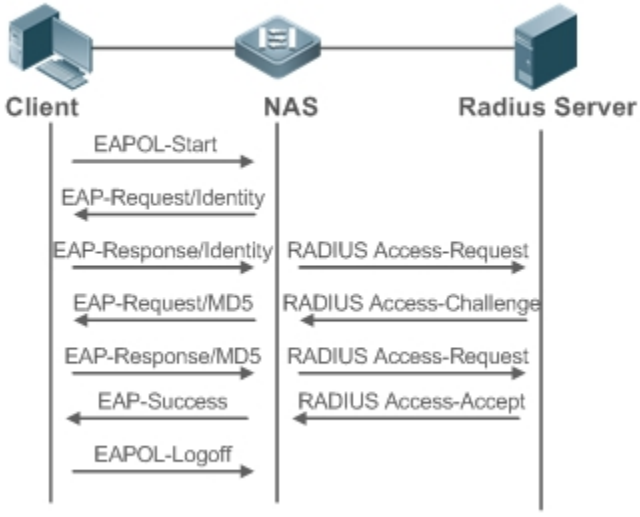
The authenticator server is usually an RADIUS server. It cooperates with the authenticator to provide authentication service for users. The authentication server saves the user names, passwords, and related authorization information. One server can provides authentication service for multiple authenticators to achieve centralized user management. The authentication server also manages accounting data received from authenticators.

Authentication Process and Packet Exchange

The supplicant exchanges information with the authenticator through EAPOL while exchanges information with the authentication server through RADIUS. EAPOL is encapsulated on the MAC layer, with the type number of 0x888E. IEEE assigned a multicast MAC address 01-80-C2-00-00-03 for EAPOL to exchange packets during initial authentication. FS Supplicant may also use 01-D0-F8-00-00-03 to for initial authentication packets.

Figure 4-4 shows the typical authentication process of a wired user.

Figure 4-4



This is a typical authentication process initiated by a user. In special cases, the NAS, may take place of the user to initiate an authentication request.

Authenticating User Status

802.1X determines whether a user on a port can access the network based on the authentication status of the port. FS products extend the 802.1X and realizes access control based on users (identify a wired user by the MAC address) by default. FS 802.1X can also be enabled in interface configuration mode. For details, see the chapter "Configuration."

All users on an uncontrolled port can access network resources, while users on a controlled port can access network resources only after authorized. When a user initiates authentication, its status remains Unauthorized and cannot access the network yet. After it passes authentication, its status changes to Authorized and can access network resources.

If the user connected to a controlled port does not support 802.1X, it will not respond to the NAS requesting the user name of the user.

That means, the user remains Unauthorized and cannot access network resources.

In the case of 802.1X-enabled user and 802.1X-disabled NAS, if the user does not receive any responses after sending a specified number of EAPOL-Start packets, it regards the connected port uncontrolled and directly accesses network resources.

On 802.1X-enabled devices, all ports are uncontrolled by default. We can configure a port as controlled so that all users on this port have to be authorized.

If a user passes authentication (that is, the NAS receives a success packet from the RADIUS server), the user becomes Authorized and can freely access network resources. If the user fails in authentication, it remains Unauthorized and re-initiates authentication. If the communication between the NAS and the RADIUS server fails, the user remains Unauthorized and cannot access network resources.

When a user sends an EAPOL-LOGOFF packet, the user's status changes from Authorized to Unauthorized.

When a port of the NAS goes down, all users on this port will become Unauthorized.

When the NAS restarts, all users on it become Unauthorized.

Deploying the Authentication Server

802.1X authentication uses the RADIUS server as the authentication server. Therefore, when 802.1X secure admission is deployed, the RADIUS server also needs to be deployed. Common RADIUS servers include Microsoft IAS/NPS, Cisco ACS, and FS-SAM/SMP. For details about the deployment procedure, see related software description.

Configuring Authentication Parameters

To use 802.1X authentication, enable 802.1X authentication on the access port and configure AAA authentication method list and RADIUS server parameters. To ensure the accessibility between the NAS and RADIUS server, the 802.1X server timeout should be longer than the RADIUS server timeout.

Supplicant

A user should start FS Supplicant to enter the user name and initiate authentication. If the operating system brings an own authentication client and the network is available, a dialog box will be displayed, asking the user to enter the user name. Different clients may have different implementation processes and Graphical User Interfaces (GUIs). It is recommended to use FS Supplicant as the authentication client. If other software is used, see related software description.

Offline

If a user does not want to access the network, it can choose to go offline by multiple approaches, such as powering off the device, connecting the port to the network, and offline function provided by some supplicants.

4.3.2 Authorization

After a user passes authentication, the NAS restricts the accessible network resources of the user in multiple approaches, such as binding the IP address and the MAC address, and specifying the maximum online time or period, accessible VLANs, and bandwidth limit.

Working Principle

Authorization means to bind the permissions with the users. A user is identified based on the MAC address and VLAN ID, as mentioned before. Besides MAC-VID binding, some other information such as the IP address and VLAN ID are bound with a user to implement authorization.

IP Authorization

802.1X does not support IP address identification. FS 802.1X authentication extends 802.1X to support IP-MAC binding, which is called IP authorization. IP authorization supports four modes:

Supplicant authorization: The IP address is provided by FS Supplicant.

RADIUS authorization: After successful authentication, the RADIUS server delivers the IP address to the NAS.

DHCP authorization: In such case, an authenticated user will initiate a DHCP request to obtain an IP address, and then bind the IP address with the MAC address of the client.

Mixed authorization: IP-MAC binding is configured for users in the following sequence: Supplicant authorization -> RADIUS authorization -> DHCP authorization. That is, the IP address provided by FS Supplicant preferred, then the IP address provided by the RADIUS server, and finally the IP address provided by DHCP.

ACL Authorization

After user authentication is complete, the authentication server delivers the ACL or ACE to users. The ACL must be configured on the authentication server before delivery while no extra configuration is required for ACE delivery. ACL authorization delivers the ACL based on RADIUS attributes such as standard attributes, FS-proprietary attributes, and Cisco-proprietary attributes. For details, see the software description related to the RADIUS server.

4.3.3 Accounting

Accounting allows the network operators to audit the network access or fees of accessed users, including the online time and traffic.

Working Principle

Accounting is enabled on the NAS. The RADIUS server supports RFC2869-based accounting. When a user goes online, the NAS sends an accounting start packet to the RADIUS server which then starts accounting. When the user goes offline, the NAS sends an accounting end packet to the RADIUS server which then completes the accounting and generates a network fee accounting list. Different servers may perform accounting in different ways. Moreover, not all servers support accounting. Therefore, refer to the usage guide of the authentication server during actual deployment and accounting.

Accounting Start

After a user passes authentication, the accounting-enabled switch sends the RADIUS server an accounting start packet carrying user accounting attributes such as user name and accounting ID. After receiving the packet, the RADIUS server starts accounting.

Accounting Update

The NAS periodically sends Accounting Update packets to the RADIUS server, making the accounting more real-time. The accounting update interval can be provided by the RADIUS server or configured on the NAS.

Accounting End

After a user goes offline, the NAS sends the RADIUS server an accounting end packet carrying the online period and traffic of the user. The RADIUS server generates online records based on the information carried in this packet.

4.4 Configuration

Configuration	Description and Command	
Configuring 802.1X Basic Functions	(Mandatory) It is used to configure basic authentication and accounting.	
	aaa new-model	Enables AAA.
	aaa authentication dot1x	Configures an AAA authentication method list.
	aaa accounting network	Configures an AAA accounting method list.
	radius-server host	Configures the RADIUS server parameters.
	radius-server key	Configures the preshared key for communication between the NAS and the RADIUS server.
	dot1x port-control auto	Enables 802.1X authentication on a port.
Configuring 802.1X Parameters	(Optional) It is used to configure 802.1X parameters. Ensure that the 802.1X server timeout is longer than the RADIUS server timeout. Online FS client detection applies only to FS Supplicant.	
	dot1x re-authentication	Enables re-authentication.
	dot1x timeout re-authperiod	Configures the re-authentication interval.
	dot1x timeout tx-period	Configures the interval of EAP-Request/Identity packet retransmission.
	dot1x reauth-max	Configures the maximum times of EAP-Request/Identity packet retransmission.
	dot1x timeout supp-timeout	Configures the interval of EAP-Request/Challenge packet retransmission.
	dot1x max-req	Configures the maximum times of EAP-Request/Challenge packet retransmission.
	dot1x timeout server-timeout	Configures the authentication server timeout.
	dot1x timeout quiet-period	Configures the quiet period after authentication fails.
	dot1x auth-mode	Specifies the authentication mode (EAP/CHAP/PAP).
	dot1x client-probe enable	Enables online FS client detection.
	dot1x probe-timer interval	Configures the interval of online FS client detection.
	dot1x probe-timer alive	Configures the duration of online FS client detection.
Configuring Authorization	(Optional) It is used to configure authorization. FS Supplicant should be used to perform supplicant authorization in IP authorization mode.	
	aaa authorization ip-auth-mode	Specifies the IP authorization mode.
	dot1x private-supplicant-only	Filters non-FS clients.
	dot1x redirect	Enables Web Redirection for 2G FS Supplicant Deployment.

Configuring MAB	(Optional) It is used to configure MAC Authentication Bypass (MAB). 802.1X authentication takes priority over MAB. MAB does not support IP authorization. Single-user MAB and multi-user MAB cannot be enabled at the same time. MAB adopts the PAP authentication mode. Ensure correct server configurations during deployment.	
	dot1x mac-auth-bypass	Enables single-user MAB.
	dot1x mac-auth-bypass multi-user	Enables multi-user MAB.
	dot1x multi-mab quiet-period	Configures the quiet period after multi-user MAB fails.
	dot1x mac-auth-bypass timeout-activity	Configures the timeout of MAB users.
	dot1x mac-auth-bypass violation	Enables MAB violation mode.
	dot1x mac-auth-bypass vlan	Configures VLAN-based MAB.
Configuring IAB	(Optional) It is used to configure Inaccessible Authentication Bypass (IAB).	
	dot1x critical	Enables IAB.
	dot1x critical recovery action reinitialize	Enables IAB recovery.
	dot1x critical vlan	Configures the IAB VLAN.
Configuring Port Control	**dot1x port-control-mode mac-based**	Enables the MAC-based control mode.
	dot1x port-control-mode port-based	Enables the port-based control mode.
	dot1x port-control-mode port-based single-host	Enables the single-user port-based control mode.
	dot1x stationarity enable	Disables migration of dynamic users.
Configuring Extended Functions	(Optional) It is used to configure active authentication requests on a port. (Optional) It is used to configure the authenticated client list. (Optional) It is used to enable 802.1X packet sending with the pseudo source MAC address. (Optional) It is used to configure multiple accounts for the same MAC address.	
	dot1x auto-req	Enables active authentication.
	dot1x auto-req packet-num	Configures the number of active authentication requests.
	dot1x auto-req user-detect	Enables user detection for active authentication.
	dot1x auto-req req-interval	Configures the interval of active authentication request.
	dot1x auth-address-table address	Configures the authenticatable client list.
	dot1x pseudo source-mac	Enables 802.1X packets sending with the pseudo source MAC address.
	dot1x multi-account enable	Enables multi-account authentication with one MAC address.
	dot1x valid-ip-acct enable	Enables IP-triggered accounting.

	dot1x valid-ip-acct timeout	Configures the timeout of obtaining IP addresses after users get authenticated. If timeout is reached, they will be kicked off.
	dot1x auth-with-order	Eanbles 802.1X precedence over MAB.
	dot1x user-name compatible	Configures compatibility for H3C 802.1X authentication clients and authentication servers.

4.4.1 Configuring 802.1X Basic Functions

Configuration Effect

- Enable basic authentication and accounting services.
- On a wired network, run the **dot1x port-control auto** command in interface configuration mode to enable 802.1X authentication on a port.
- Run the **radius-server host** *ip-address* command to configure the IP address and port information of the RADIUS server and the **radius-server key** command to configure the RADIUS communication key between the NAS and the RADIUS server to ensure secure communication.
- Run the **aaa accounting update** command in global configuration mode to enable accounting update and the **aaa accounting update interval** command on the NAS to configure the accounting update interval. If the RADIUS server supports accounting update, you can also configure it on the RADIUS server. Prefer to use the parameters assigned by the authentication server than the parameters configured on the NAS.

Notes

- Configure accurate RADIUS parameters so that the basic RADIUS communication is proper.
- The 802.1X authentication method list and accounting method list must be configured in AAA. Otherwise, errors may occur during authentication and accounting.
- Due to chipset restriction on switches, if 802.1X is enabled on one port, all ports will send 802.1X packets to the CPU.
- If 802.1X is enabled on a port but the number of authenticated users exceeds the maximum number of users configured for port security, port security cannot be enabled.
- If port security and 802.1X are both enabled but the security address has aged, 802.1X users must re-initiate authentication requests to continue the communication.
- Users with IP addresses statically configured or compliant with IP-MAC binding can access the network without authentication.
- 802.1X uses the default method list by default. If the default method list is not configured for AAA, run the **dot1x authentication** and **dot1x accounting** commands to reconfigure the it.

Configuration Steps

Enabling AAA

- (Mandatory) 802.1X authentication and accounting take effect only after AAA is enabled.
- Enable AAA on the NAS that needs to control user access by 802.1X.

Command	**aaa new-model**

Parameter Description	N/A
Defaults	AAA is disabled by default.
Command Mode	Global configuration mode
Usage Guide	AAA is disabled by default. This command is mandatory for the deployment of 802.1X authentication.

Enabling an AAA Authentication Method List

- Mandatory.
- The AAA authentication method list must be consistent with the 802.1X authentication method list.
- Enable an AAA authentication method list after 802.1X authentication is enabled on the NAS.

Command	**aaa authentication dot1x** *list-name* **group radius**
Parameter Description	*list-name*: Indicates the 802.1X authentication method list of AAA.
Defaults	No AAA authentication method list is configured by default.
Command Mode	Global configuration mode
Usage Guide	AAA authentication modes are disabled by default. The AAA authentication mode must be consistent with the 802.1X authentication mode.

Configuring the RADIUS Server Parameters

- (Mandatory) The RADIUS server parameters must be configured to ensure proper communication between the NAS and the RADIUS server.
- Configure RADIUS server parameters after 802.1X authentication is enabled on the NAS.

Command	**radius-server host** *ip-address* [**auth-port** *port1*] [**acct-port** *port2*]
Parameter Description	*ip-address*: Indicates the IP address of the RADIUS server. *port1*: Indicates the authentication port. *port2*: Indicates the accounting port.
Defaults	No RADIUS server parameters are configured by default.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Preshared Key for Communication between the NAS and RADIUS Server

- (Mandatory) The preshared key for communication between the NAS and RADIUS server must be configured to ensure proper communication between the NAS and the RADIUS server.
- Configure the preshared key of the RADIUS server after 802.1X authentication is enabled on the NAS.

Command	**radius-server key** *string*
Parameter Description	*string*: Indicates the preshared key.
Defaults	No preshared key is configured for communication between the NAS and RADIUS server by default.

Command Mode	Global configuration mode
Usage Guide	The IP address of the NAS must be the same as that registered on the RADIUS server. The preshared key on the NAS must be the same as that on the RADIUS server. If the default RADIUS communication ports are changed on the RADIUS server, you need to change the communication ports on the NAS correspondingly.

Enabling 802.1X on a Port

- This command is mandatory for a wired network.
- Enable 802.1X on switches.

Command	**dot1x port-control auto**
Parameter Description	N/A
Defaults	802.1X is disabled on a port by default.
Command Mode	Interface configuration mode
Usage Guide	802.1X is disabled on a port by default. This command is mandatory for the deployment of 802.1X authentication. The default method list is used by default. If the 802.1X authentication method list in AAA is not the default one, the configured 802.1X authentication method list should match.

Verification

Start FS Supplicant, enter the correct account information, and initiate authentication. Then check whether the 802.1X and RADIUS configurations are correct.

Checking for 802.1X Authentication Entries

Command	**show dot1x summary**
Parameter Description	N/A
Command Mode	Privileged EXEC mode/Global configuration mode/Interface configuration mode
Usage Guide	Display entries of authenticated users to check the authentication status of users, for example, authenticating, authenticated, or quiet.
Command Display	FS#show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-state Port-Status User-Type Time --------- ---------- ---------- --------- ---- ---------- ---------- ---------- --------- ---------- 16777302 ts-user b048.7a7f.f9f3 wlan 1 1 Authenticated Idle Authed static 0days 0h 0m12s

Checking for AAA User Entries

Command	**show aaa user all**
Parameter Description	N/A
Command	Privileged EXEC mode/Global configuration mode/Interface configuration mode

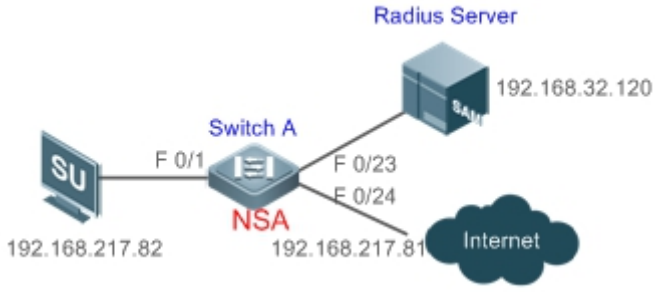
Mode	
Usage Guide	Display information of AAA users.
Command Display	FS#show aaa user all ---------- Id ----- Name 2345687901 wwxy ----------

- Check whether the RADIUS server responds to authentication based on the RADIUS packets between the NAS and the RADIUS server. If no, it means that the network is disconnected or parameter configurations are incorrect. If the RADIUS server directly returns a rejection reply, check the log file on the RADIUS server to identify the cause, e.g., of the authentication mode of the authentication server is incorrectly configured.

Configuration Example

In this example, FS-SAM acts as the authentication server.

Configuring 802.1X Authentication on a Switch

Scenario **Figure 4-5**	
Configuration Steps	● Register the IP address of the switch on the RADIUS server and configure the communication key between the switch and the RADIUS server. ● Create an account on the RADIUS server. ● Enable AAA on the switch. ● Configure RADIUS parameters on the switch. ● Enable 802.1X authentication on ports of the switch. Switch configurations are as follows. For detailed configuration on the RADIUS server, see the *Configuring RADIUS*.
	FS# configure terminal FS (config)# aaa new-model FS (config)# radius-server host 192.168.32.120 FS (config)# radius-server key FS FS (config)# interface FastEthernet 0/1 FS (config-if)# dot1x port-control auto
Verification	Check whether authentication is proper and network access behaviors change after authentication. ● The account is successfully created, such as **username:tests-user,password:test**. ● The user fails to ping 192.168.32.120 before authentication. ● After the user enters account information and click **Authenticate** on FS Supplicant, the authentication succeeds and the user can successfully ping 192.168.32.120.

	● Information of the authenticated user is displayed. FS# show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-State Port-Status User-Type Time --------- ---------- ---------- --------- ---- ---------- ---------- ---------- --------- ---------- 16778217 ts-user 0023.aeaa.4286 Fa0/1 2 Authenticated Idle Authed static 0days 0h 0m 7s

Common Errors

- RADIUS parameters are incorrectly configured.
- The RADIUS server has a special access policy, for example, the RADIUS packets must carry certain attributes.
- The AAA authentication mode list is different from the 802.1X authentication mode list, causing authentication failure.

4.4.2 Configuring 802.1X Parameters

Configuration Effect

- Adjust 802.1X parameter configurations based on the actual network situation. For example, if the authentication server has poor performance, you can raise the authentication server timeout.

Notes

- 802.1X and RADIUS have separate server timeouts. By default, the authentication server timeout of 802.1X is 5 seconds while that of RADIUS is 15 seconds. In actual situations, ensure that the former is greater than the latter. You can run the **dot1x timeout server-timeout** command to adjust the authentication server timeout of 802.1X. For detailed configuration about the RADIUS server timeout, see the *Configuring RADIUS*.
- Online client detection applies only to FS Supplicant.

Configuration Steps

Enabling Re-authentication

- (Optional) After re-authentication is enabled, the NAS can periodically re-authenticate online users.
- Enable re-authentication after 802.1X authentication is enabled on the NAS.

Command	**dot1x re-authentication**
Parameter Description	N/A
Defaults	Re-authentication is disabled by default.
Command Mode	Global configuration mode
Usage Guide	You can run this command to periodically re-authenticate users.

Configuring the Re-authentication Interval

- (Optional) You can configure the re-authentication interval for users.
- Configure the re-authentication interval after 802.1X authentication is enabled on the NAS. The re-authentication interval takes effect only after re-authentication is enabled.

Command	**dot1x timeout re-authperiod** *period*
Parameter Description	*period*: Indicates the re-authentication interval in the unit of seconds.
Defaults	The default value is 3,600 seconds.
Command Mode	Global configuration mode
Usage Guide	Adjust the re-authentication interval as required.

Configuring the Interval of EAP-Request/Identity Packet Retransmission

- (Optional) A larger value indicates a longer interval of packet retransmission.
- Configure the interval of EAP-Request/Identity packet retransmission after 802.1X authentication is enabled on the NAS.

Command	**dot1x timeout tx-period** *period*
Parameter Description	*period*: Indicates the interval of EAP-Request/Identity packet retransmission in the unit of seconds.
Defaults	The default value is 3 seconds.
Command Mode	Global configuration mode
Usage Guide	It is recommended to use the default value. Adjust the value based on how long the authentication client responds to the NAS's requests.

Configuring the Maximum Times of EAP-Request/Identity Packet Retransmission

- (Optional) A larger value indicates more frequent retransmissions.
- Configure the maximum times of EAP-Request/Identity packet retransmission after 802.1X authentication is enabled on the NAS.

Command	**dot1x reauth-max** *num*
Parameter Description	*num*: Indicates the maximum times of EAP-Request/Identity packet retransmission.
Defaults	The default value is 3 for switches
Command Mode	Global configuration mode
Usage Guide	It is recommended to use the default value. In the case of high-rate packet loss, increase this value so that the clients can easily receive packets from the NAS.

Configuring the Interval of EAP-Request/Challenge Packet Retransmission

- (Optional) A larger value indicates a longer retransmission interval.
- Configure the interval of EAP-Request/Challenge packet retransmission after 802.1X authentication is enabled on the NAS.

Command	**dot1x timeout supp-timeout** *time*
Parameter Description	*time*: Indicates the interval of EAP-Request/Challenge packet transmission in the unit of seconds.
Defaults	The default value is 3 seconds for switches
Command Mode	Global configuration mode

Usage Guide	It is recommended to use the default value. Increase this value in the case of high-rate packet loss.

Configuring the Maximum Times of EAP-Request/Challenge Packet Retransmission

- (Optional) A larger value indicates more frequent retransmissions.
- Configure the maximum times of EAP-Request/Challenge packet retransmission after 802.1X authentication is enabled on the NAS.

Command	**dot1x max-req** *num*
Parameter Description	*num*: Indicates the maximum times of EAP-Request/Challenge packet retransmission in the unit of seconds.
Defaults	The default value is 3.
Command Mode	Global configuration mode
Usage Guide	Optional. It is recommended to use the default value. Increase this value in the case of high-rate packet loss.

Configuring the Authentication Server Timeout

- (Optional) A larger value indicates a longer authentication server timeout.
- Configure the authentication server timeout after 802.1X authentication is enabled on the NAS.
- The server timeout of RADIUS must be greater than that of 802.1X.

Command	**dot1x timeout server-timeout** *time*
Parameter Description	*time*: Indicates the authentication server timeout in the unit of seconds.
Defaults	The default value is 5 seconds.
Command Mode	Global configuration mode
Usage Guide	It is recommended to use the default value. Increase this value if the communication between the NAS and RADIUS server is unstable.

Configuring the Quiet Period after Authentication Fails

- (Optional) A larger value indicates a longer quiet period.
- Configure the quiet period after 802.1X authentication is enabled on the NAS.

Command	**dot1x timeout quiet-period** *time*
Parameter Description	*time*: Indicates the quiet period after authentication fails. The unit is second.
Defaults	The default value is 10 seconds.
Command Mode	Global configuration mode
Usage Guide	It is recommended to use the default value. Increase this value to prevent users from frequently initiating authentication to the RADIUS server, thereby reducing the load of the authentication server.

Specifying the Authentication Mode

- (Optional) Configure the mode for 802.1X authentication.

- Configure the authentication mode after 802.1X authentication is enabled on the NAS.

Command	**dot1x auth-mode** {**eap** | **chap** | **pap**}
Parameter Description	**eap**: Indicates EAP authentication. **chap**: Indicates CHAP authentication. **pap**: Indicates PAP authentication.
Defaults	The default value is **eap**.
Command Mode	Global configuration mode
Usage Guide	Select the authentication mode supported by FS Supplicant and authentication server.

Enabling Online FS Client Detection

- (Optional) If online FS client detection is enabled, the NAS can find clients going offline in a timely manner to prevent incorrect accounting.
- This function applies only to FS 802.1X authentication clients.
- Enable online FS client detection after 802.1X authentication is enabled on the NAS.

Command	**dot1x client-probe enable**
Parameter Description	N/A
Defaults	Online FS client detection is disabled by default.
Command Mode	Global configuration mode
Usage Guide	It is recommended to enable this function when FS Supplicant is used.

Configuring the Interval of Online FS Client Detection

- (Optional) A larger value indicates a longer time interval at which FS clients send detection packets.
- Configure the interval of online FS client detection after 802.1X authentication is enabled on the NAS.

Command	**dot1x probe-timer interval** *time*
Parameter Description	*time*: Indicates the time interval at which FS Supplicant sends a heartbeat packet to the NAS. The unit is second.
Defaults	The default value is 20 seconds.
Command Mode	Global configuration mode
Usage Guide	It is recommended to use the default value.

Configuring the Duration of Online FS Client Detection

- (Optional) A larger value indicates a longer interval at which the NAS finds clients going offline.
- Configure the duration of online FS client detection after 802.1X authentication is enabled on the NAS.

Command	**dot1x probe-timer alive** *time*
Parameter Description	*time*: Indicates the duration of online FS client detection in the unit of seconds.
Defaults	The default value is 250 seconds.

Command Mode	Global configuration mode
Usage Guide	Optional. If the NAS does not receive any detection packets from an online client within the detection duration, it regards the client offline. It is recommended to use the default value.

Verification

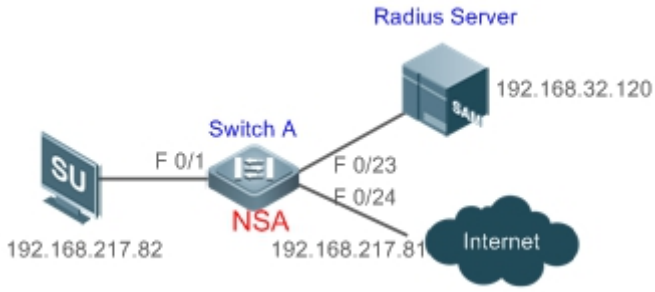
Run the **show dot1x** command to check whether parameter configurations take effect.

Configuration Example

Specifying the Authentication Mode

Scenario	The NAS is deployed in standalone mode.
Configuration Steps	Set the authentication mode to **chap**.
	FS(config)#dot1x auth-mode chap
Verification	Display the configurations. FS(config)#show dot1x 802.1X basic information: 802.1X Status ..... enable Authentication Mode ..... chap Authorization mode ..... disable Total User Number ..... 0 (exclude dynamic user) Authenticated User Number ..... 0 (exclude dynamic user) Dynamic User Number ..... 0 Re-authentication ..... disable Re-authentication Period ..... 3600 seconds Re-authentication max ..... 3 times Quiet Period ..... 10 seconds Tx Period ..... 30 seconds Supplicant Timeout ..... 3 seconds Server Timeout ..... 5 seconds Maximum Request ..... 3 times Client Online Probe ..... disable Eapol Tag ..... disable 802.1x redirect ..... disable Private supplicant only ..... disable

Enabling Online Client Detection

Scenario **Figure 4-6**	
Configuration Steps	Enable online client detection.
	FS(config)#dot1x client-probe enable
	● Users can remain online only when their FS Supplicant sends online detection packets as scheduled.
Verification	● Display the configurations. FS(config)#show dot1x 802.1X basic information: 802.1X Status ..... enable Authentication Mode ..... chap Authorization mode ..... disable Total User Number ..... 0 (exclude dynamic user) Authenticated User Number ..... 0 (exclude dynamic user) Dynamic User Number ..... 0 Re-authentication ..... disable Re-authentication Period ..... 3600 seconds Re-authentication max ..... 3 times Quiet Period ..... 10 seconds Tx Period ..... 30 seconds Supplicant Timeout ..... 3 seconds Server Timeout ..... 5 seconds Maximum Request ..... 3 times Client Online Probe ..... enable Eapol Tag ..... disable 802.1x redirect ..... disable

Common Errors

- The server timeout is shorter than the RADIUS timeout.
- Online client detection is enabled but the authentication program is not FS Supplicant.

4.4.3 Configuring Authorization

Configuration Effect

- In IP authorization, authenticated users have to use the specified IP addresses to access the network, preventing IP address fake. IP

authorization can be enabled in global configuration mode or interface configuration mode. IP authorization enabled in interface configuration mode takes priority over that configured in global configuration mode.

- Enable non-FS client filtering. If this function is enabled, users must use FS Supplicant for authentication so that they will enjoy services provided by FS Supplicant, such as anti-proxy or SMS.
- Enable Web redirection to support 2G FS Supplicant deployment. 2G FS Supplicant deployment means that a user needs to download FS Supplicant through the browser and then initiate authentication through FS Supplicant. 2G FS Supplicant deployment facilitates quick deployment of FS Supplicant in the case of massive users.

Notes

- If multiple authentication supplicants are used, disable this function.
- If the IP authorization mode is changed, all authenticated users will go offline and have to get re-authenticated before online again.
- In mixed authorization mode, IP authorization with a higher priority is used during user authentication. For example, if FS Supplicant provides an IP address for this RADIUS-authentication user during its re-authentication, this IP address will be used for authorization.
- 2G FS Supplicant deployment and Web authentication cannot be used at the same time.
- 2G FS Supplicant deployment requires the setting of the **redirect** parameter. For details, see the *Configuring Web Authentication*.

Configuration Steps

Specifying the Global IP Authorization Mode

- The **supplicant** mode only applies to FS Supplicant.
- In **radius-server** mode, the authentication server needs to assign IP addresses based on the **framed-ip** parameters.
- In **dhcp-server** mode, DHCP snooping must be enabled on the NAS.
- (Optional) Configure an IP-MAC binding.
- Configure the IP authorization mode after 802.1X authentication is enabled on the NAS.

Command	**aaa authorization ip-auth-mode** { **disable | supplicant | radius-server | dhcp-server | mixed** }
Parameter Description	**disable**: Disables IP authorization. **supplicant**: Indicates IP authorization by the supplicant. **radius-server**: Indicates IP authorization by the RADIUS server. **dhcp-server**: Indicates IP authorization by the DHCP server. **mixed**: Indicates IP authorization in a mixed manner.
Defaults	IP authorization is disabled by default.
Command Mode	Global configuration mode
Usage Guide	Select the IP authorization mode based on actual deployment.

Enabling Web Redirection for 2G FS Supplicant Deployment

- (Optional) If the redirection for 2G FS Supplicant deployment is enabled, users not having any 802.1X authentication clients on a controlled port can download and install an 802.1X authentication client through Web pages.

- Enable Web redirection for 2G FS Supplicant deployment after 802.1X authentication is enabled on the NAS.
- The **redirect** parameter must be configured. For details, see the *Configuring Web Authentication*.

Command	**dot1x redirect**
Parameter Description	N/A
Defaults	The redirection for 2G FS Supplicant deployment is disabled by default.
Command Mode	Global configuration mode
Usage Guide	The **redirect** parameter must be configured. For details, see the *Configuring Web Authentication*.

Enabling Non-FS Client Filtering

- (Optional) If this function is enabled, non-FS clients cannot perform authentication.
- Enable non-FS client filtering after 802.1X authentication is enabled on the NAS.

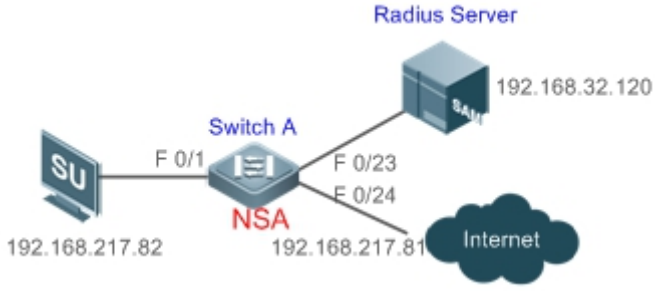
Command	**dot1x private-supplicant-only**
Parameter Description	N/A
Defaults	Non-FS client filtering is disabled by default.
Command Mode	Global configuration mode
Usage Guide	This function can be enabled only when FS Supplicant is used.

Verification

- After IP authorization is enabled, use the client to initiate authentication and go online, and then change the IP address. As a result, the client cannot access the network.
- Enable Web redirection for 2G FS Supplicant deployment. When you start the browser to visit a website, the system automatically redirects to the download Web page and downloads the authentication client. You can access the network only when authenticated by the client.

Configuration Example

Configuring the IP Authorization Mode

Scenario **Figure 4-7**	

Configuration Steps	● Enable AAA. ● Configure RADIUS. ● Enable 802.1X on a controlled port. ● Globally enable IP authorization in **supplicant** mode.
	FS(config)#aaa authorization ip-auth-mode supplicant
	● FS Supplicant initiates authentication and the authentication succeeds. ● FS Supplicant only uses 192.168.217.82 for communication.
Verification	● Display the configurations. FS(config)#show dot1x user name ts-user Supplicant information: MAC address ..... b048.7a7f.f9f3 Username ..... ts-user User ID ..... 16777303 Type ..... static VLAN ..... 1 Online duration ..... 0days 0h 0m21s Up average bandwidth ..... 0 kBps Down average bandwidth ..... 0 kBps Authorized VLAN ..... 1 Authorized session time ..... 20736000 seconds Authorized flux ..... unlimited Accounting ..... No Proxy user ..... Permit Dial user ..... Permit IP privilege ..... 0 Private supplicant ..... no Max user number on this port ..... 0 Authorization ip address 192.168.217.82

Common Errors

- There are multiple authentication clients on the network but non-FS client filtering is enabled, causing some users to fail authentication.
- The **redirect** parameter is incorrectly configured, causing abnormalities in redirection for 2G FS Supplicant downloading.

4.4.4 Configuring MAB

Configuration Effect

- If the MAC address of an access user is used as the authentication account, the user does not need to install any supplicants. This applies to some dumb users such as networking printers.
- Single-user MAB applies to two scenarios:
  - There is only one dumb user connected to a port.
  - Only one user needs to be authenticated. After this, all other users can access the network.
- Multi-user MAB applies to the scenario where multiple dumb users connected to a port. For example, multiple VoIP devices are

deployed in the network call center.

- Multi-user MAB can be used with 802.1X authentication. It applies to mixed access scenarios such as the PC-VoIP daisy-chain topology.

Notes

- A MAB-enabled port sends an authentication request packet as scheduled by **tx-period**. If the number of the sent packets exceeds the number specified by **reauth-max** but still no client responds, this port enters the MAB mode. Ports in MAB mode can learn the MAC addresses and use them as the account information for authentication.
- When using the MAC address as the user name and password on the authentication server, delete all delimiters. For example, if the MAC address of a user is 00-d0-f8-00-01-02, the user name and password should be set to 00d0f8000102 on the authentication server.
- 802.1X takes priority over MAB. Therefore, if a user having passed MBA authentication uses a client to initiate 802.1X authentication, MAB entries will be removed.
- MAB supports only PAP authentication. PAP authentication should be enabled also on the authentication server.
- Only when active authentication is enabled, can MAB detect whether the user can perform 802.1X authentication. Therefore, automatic authentication must be enabled for MAB deployment.

Configuration Steps

Enabling Single-User MAB

- Optional.
- Single-user MAB applies when only one user connected to a port needs to be authenticated.
- Enable single-user MAB on the 802.1X controlled port of the NAS.

Command	**dot1x mac-auth-bypass**
Parameter Description	N/A
Defaults	Single-user MAB is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	This command applies only to switches. Single-user MAB applies when only one dumb user connected to a port needs to be authenticated. If you want to restrict the number of users, enable the violation mode.

Configuring the Timeout of MAB Users

- Optional.
- After a MAC address in MAB mode is authenticated and goes online, the NAS regards the MAC address online unless re-authentication fails, the port goes down, or the MAC address goes offline due to management policies such as kickoff. You can configure the timeout of authenticated MAC addresses. The default value is 0, indicating always online.
- Configure the timeout of MAB users on the 802.1X controlled port of the NAS.

Command	**dot1x mac-auth-bypass timeout-activity** *value*
Parameter	*value*: Indicates the maximum online time of MAB users in the unit of seconds.

Description	
Defaults	The default value is 0, indicating no time restriction.
Command Mode	Interface configuration mode
Usage Guide	The MAB timeout applies to both single-user MAB and multi-user MAB.

Enabling the MAB Violation Mode

- Optional.
- Enable MAB violation on the 802.1X controlled port of the NAS.
- By default, after one MAC address passes MAB authentication, data of all switches connected to the port can be forwarded. However, for security purposes, the administrator may request one MAB port to support only one MAC address. In this case, you can enable MAB violation on the port. If more than one MAC address is found connected to a MAB violation-enabled port after the port enters MAB mode, the port will become a violation.

Command	**dot1x mac-auth-bypass violation**
Parameter Description	N/A
Defaults	MAB violation is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	This command applies only to switches. Configure this command only when only one dumb user is connected to the port. MAB violation applies only to single-user MAB.

Enabling Multi-user MAB

- Optional.
- Enable multi-user MAB on the 802.1X controlled port of the NAS.

Command	**dot1x mac-auth-bypass multi-user**
Parameter Description	N/A
Defaults	Multi-user MAB is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	This command applies only to switches. Configure this command when multiple dumb users connected to the port need to be authenticated.

Configuring the Quiet Period after Multi-user MAB Fails

- Optional.
- Configure the quite period of the multi-user MAB failure after multi-user MAB is enabled on the NAS.
- If multi-user MAB is enabled, you should prohibit unauthorized users from frequently initiating authentication to protect the NAS from attacks of these users and thereby reduce the load of the authentication server. Configure the quite period of the multi-user MAB failure in global configuration mode. That is, if a MAC address fails authentication, it needs to re-initiate authentication after

the quiet period. Configure this quiet period based on the actual situation. The default value is 0, indicating that a user can re-initiate authentication immediately after authentication fails.

Command	**dot1x multi-mab quiet-period** *value*
Parameter Description	*value*: Indicates the quiet period after authentication fails.
Defaults	The default value is 0s.
Command Mode	Global configuration mode
Usage Guide	This command applies only to switches. If too many dumb users connected to a port are authenticated, run this command to limit the authentication rate.

Configuring VLAN-based MAB

- Optional.
- Enable VLAN-based MAB after multi-user MAB is enabled on the NAS.
- If you configure VLANs as MAB VLANs, only users in these VLANs can perform MAB.

Command	**dot1x mac-auth-bypass vlan** *vlan-list*
Parameter Description	*vlan-list*: Indicates the VLANs supporting MAB.
Defaults	VLAN-based MAB is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	This command applies only to switches. Run this command when a port allows only users in specified VLANs to perform MAB.

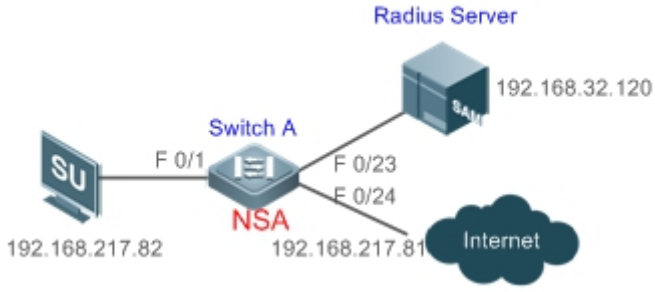
Verification

Check whether the dumb user can access the network. If yes, MAB takes effect. If no, MAB does not take effect.

- Check whether MAB functions are configured on the authentication server and NAS.
- Check whether dumb users with illegitimate MAC addresses cannot access the network.
- Check whether dumb users with illegitimate MAC addresses can access the network.

Configuration Example

Enabling Multi-user MAB on a Switch

Scenario **Figure 4-8**	

Configuration Steps	● Register the IP address of the Switch A on the RADIUS server and configure the communication key between Switch A and the RADIUS server. ● Create an account on the RADIUS server. ● Enable AAA on Switch A. ● Configure RADIUS parameters on Switch A. ● Enable 802.1X and multi-user MAB on a port of Switch A. Switch configurations are as follows. For detailed configuration on the RADIUS server, see the *Configuring RADIUS*.
	FS# configure terminal FS (config)# aaa new-model FS (config)# radius-server host 192.168.32.120 FS (config)# radius-server key FS FS (config)# interface FastEthernet 0/1 FS (config-if)# dot1x port-control auto FS (config-if)# dot1x mac-auth-bypass multi-user
Verification	Check whether authentication is proper and network access behaviors change after authentication. ● The account is successfully created, such as **username: 0023aeaa4286,password: 0023aeaa4286**. ● The user fails to ping 192.168.32.120 before authentication. ● The user connects to the switch, the authentication succeeds, and the user can successfully ping 192.168.32.120. ● Information of the authenticated user is displayed. FS# show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-State Port-Status User-Type Time --------- ---------- ---------- --------- ---- ---------- ---------- ---------- --------- ---------- 16778217 0023aea... 0023.aeaa.4286 Fa0/1 2 Authenticated Idle Authed static 0days 0h 5m 8s

4.4.5 Configuring MAB Auto Authentication

Configuration Effect

- When a STA accesses the network for the first time, Web authentication is performed. When the STA is disconnected from and then reconnects to the network, authentication is not required.

Notes

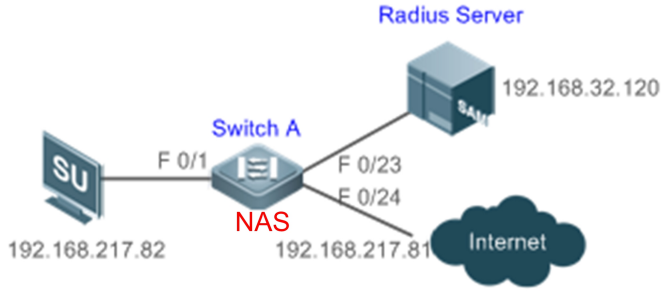
- Wireless MAB authentication is triggered by a STA advertisement. If a STA is already online, MAB authentication will not be triggered again. MAB authentication is triggered only after the STA is disconnected from and then reconnects to the network.
- When a STA accesses the network for the second time, a dialog box may be displayed for MAB authentication. When the STA accesses the network for the third time, the dialog box will not be displayed.
- If MAB authentication fails, a dialog box is displayed for Web authentication when the STA accesses the network next time.

Configuration Steps

For details about Web authentication configuration, see the Web authentication configuration document. For details about MAB authentication configuration, see section "Configuring MAB".

Configuration Example

Configuring MAB Auto Authentication

Scenario **Figure 4-9**	
Configuration Steps	● Register the IP address of the NAS on the RADIUS server and configure the communication key between the NAS and the RADIUS server. ● Create an account on the RADIUS server and bind it with a MAC address for imperceptible authentication. ● Enable AAA on the NAS. ● Configure RADIUS parameters on the NAS. ● Enable 802.1X authentication and MAB authentication on an interface of the NAS. ● Enable second-generation (or first-generation/embedded) Web authentication on an interface of the NAS and configure the Web authentication template globally. The following describes the NAS configurations. For detailed configuration on the RADIUS server, see the related configuration guide (The following describes configuration on the switch, which is similar to that on the AC/AP, except that the configuration on the switch is performed in interface configuration mode instead of WLAN RSNA configuration mode.)
	FS#configure terminal FS (config)#aaa new-model FS (config)#aaa authentication web-auth default group radius FS (config)#aaa authentication dot1x default group radius FS (config)aaa accounting net-work default start-stop group radius FS (config)#radius-server host 192.168.32.120 FS (config)#radius-server key FS FS (config)#web-auth template eportalv2 FS (config-tmplt-v2)#ip 192.158.32.9 FS (config-tmplt-v2)#url http://192.168.32.9:8080/eportal/index.jsp FS (config-tmplt-v2)#exit FS (config)#interface FastEthernet 0/1 FS (config-if)#dot1x port-control auto

	FS (config-if)#dot1x mac-auth-bypass multi-user FS (config-if)#web-auth enable eportalv2
Verification	Check whether authentication is normal and network access behaviors change after authentication. ● The account is successfully created, for example, the username is 0023aeaa4286 and the password is 0023aeaa4286. ● The STA fails to ping 192.168.32.120 before authentication. ● The STA connects to the NAS, a page indicating the authentication succeeds is displayed, and the STA can successfully ping 192.168.32.120. ● The STA is disconnected from and then reconnects to the network and can successfully ping 192.168.32.120. FS#show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-State Port-Status User-Type Time --------- ---------- ---------- --------- ---- ---------- ---------- ---------- --------- ---------- 16778217 0023aea... 0023.aeaa.4286 Fa0/1 2 Authenticated Idle Authed static 0days 0h 5m 8s

Common Errors

- The MAC account format is incorrect on the authentication server.

4.4.6 Configuring IAB

Configuration Effect

- Enable IAB. After IAB is enabled, newly authenticated users can access the network even when all RADIUS servers configured on the NAS are inaccessible.
- Enable IAB recovery. When RADIUS servers recover to their reachable status, re-verify the users authorized during inaccessibility.
- Configure IAB VLANs. When RADIUS servers are inaccessible and cannot authenticate users temporarily, you can add the ports connected with users to specified VLANs so that users can access only network resources of specified VLANs.

Notes

- Configure an account and standards for testing RADIUS server accessibility. For details, see the *Configuring RADIUS*.
- IAB takes effect only when only RADIUS authentication exists in the globally configured 802.1X authentication mode list and all RADIUS servers in the list are inaccessible. If other authentication modes (for example, local and none) exist in the list, IAB does not take effect.
- After multi-domain AAA is enabled, 802.1X authentication does not need the globally configured authentication mode list any more. If IAB detects that all RADIUS servers configured in the globally configured 802.1X authentication mode list are inaccessible, it directly returns an authentication success reply to users, with no need to enter the user name. Therefore, multi-domain AAA does not take effect on this port.
- Users authenticated in IAB mode do not need to initiate accounting requests to the accounting server.

- Authenticated users can properly access the network, not affected by server inaccessibility.
- In access authentication configuration mode, when 802.1X-based IP authotication is enabled globally, users on this port, except those habing been authenticated, cannot be authenticated in IAB mode. In gateway authentication mode, users are IP authorized if their IP addresses are obtained.
- Complete 802.1X authentication is required on such 802.1X authentication clients as those of Windows. It is possible that though these clients already pass the IAB authentication, there are prompts on the clients suggesting failed authentication.
- If the failed VLAN configured does not exist, a failed VLAN will be dynamically created when a port enters the failed VLAN and automatically removed when the port exits the failed VLAN.
- Failed VLANs cannot be private VLANs, remote VLANs, and super VLANs (including sub VLANs).

Configuration Steps

Enabling IAB

- (Optional) After IAB is enabled, the NAS authorizes newly authenticated users if the authentication server is faulty.
- Enable IAB after 802.1X authentication is enabled on the NAS.

Command	**dot1x critical**
Parameter Description	N/A
Defaults	IAB is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	This command applies to ports on which newly authenticated users need to be authorized when the authentication server is inaccessible.

Enabling IAB Recovery

- (Optional) After the authentication server is recovered, the NAS re-authenticates users that are authorized when the authentication server is inaccessible.
- Enable IAB recovery actions after 802.1X authentication is enabled on the NAS.

Command	**dot1x critical recovery action reinitialize**
Parameter Description	N/A
Defaults	IAB recovery is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	If IAB recovery is enabled on a port, properly authenticated users on the port can access the network without re-authentication after the authentication server is recovered. After the authentication server is recovered, the NAS initiates authentication only to users authenticated in IAB mode during server inaccessibility.

Configuring the IAB VLAN

- (Optional) Configure the VLAN on which newly authenticated users are authorized when the authentication server becomes inaccessible.
- Enable VLAN-based IAB after 802.1X authentication is enabled on the NAS.

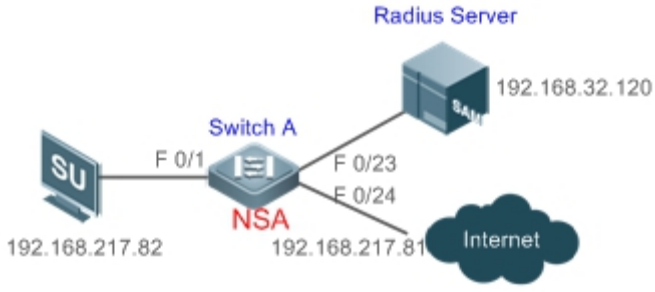
Command	**dot1x critical vlan** *vlan-id*
Parameter Description	*vlan-id*: Indicates the VLAN to redirect when the authentication server becomes inaccessible.
Defaults	The IAB VLAN is not configured by default.
Command Mode	Interface configuration mode
Usage Guide	Configure the IAB VLAN so that temporary network resources can be provided for users when servers are inaccessible.

Verification

- When the authentication server is accessible, check whether users can go online only by using the correct user name and password.
- When the authentication server is inaccessible, check whether new users can be authorized to access the network immediately after connecting to the NAS.

Configuration Example

Enabling IAB

Scenario **Figure 4-10**	
Configuration Steps	● Register the IP address of the NAS on the RADIUS server and configure the communication key between the NAS and the RADIUS server. ● Create an account on the RADIUS server. ● Enable AAA on the NAS. ● Configure RADIUS parameters and enable server accessibility probe on the NAS. ● Enable 802.1X and multi-user MAB on a port of the NAS. NAS configurations are as follows. For detailed configuration on the RADIUS server, see the *Configuring RADIUS*.
	FS# configure terminal FS (config)# aaa new-model FS (config)# radius-server host 192.168.32.120 FS (config)# radius-server key FS FS (config)# interface FastEthernet 0/1 FS (config-if)# dot1x port-control auto
Verification	Check whether authentication is proper and network access behaviors change after authentication. ● The account is successfully created, such as **username: test,password: test**. ● When the authentication server is accessible, the user fails to ping 192.168.32.120 before authentication.

	● When the authentication server becomes inaccessible, the user connects to the NAS, authentication succeeds, and the user can successfully ping 192.168.32.120. ● Information of the authenticated user is displayed. FS# show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-State Port-Status User-Type Time --------- ---------- ---------- --------- ---- ---------- ---------- ---------- --------- ---------- 16778217 test 0023.aeaa.4286 Fa0/1 2 Authenticated Idle Authed static 0days 0h10m20s

4.4.7 Configuring Port Control

Configuration Effect

- By default, the 802.1X controlled port is controlled based on the MAC address. That is, users using this MAC address can access the network only after authenticated.
- Configure the port-based control mode. As long as a user on a controlled port passes authentication, this port becomes authenticated and all users connected to this port can properly access the network.
- Configure the single-user control mode on a port. This port allows only a single user to pass authentication. If this port becomes authenticated, this user can properly access the network. At this time, if the NAS detects other users connected to this port, it will clear all users connected to this port and the user needs to re-initiate authentication.
- The port-based control mode allows or prohibits dynamic users migrating among different ports. By default, dynamic users can migrate among different ports.

Notes

- In port-based authentication mode, a controlled port supports only one authenticated user while all others are dynamic users.
- In single-user port-based authentication mode, only one user on a controlled port can pass authentication and access the network. This restriction remains even when a specified number of users is configured on this port.

Configuration Steps

Enabling the MAC-based Control Mode

- (Optional) After the MAC-based control mode is enabled, each user on an 802.1X controlled port must pass MAC-based authentication to access the network.
- Enable the MAC-based control mode after 802.1X authentication is enabled on the NAS.

Command	**dot1x port-control-mode mac-based**
Parameter Description	N/A
Defaults	The default port control mode is MAC-based control.
Command Mode	Interface configuration mode
Usage Guide	Configure the MAC-based control mode if all the users on a controlled port have to pass authentication to access the network.

Enabling the Port-based Control Mode

- (Optional) After a user on an 802.1X controlled port passes authentication, all other users on this port can access the network.
- Enable the port-based control mode after 802.1X authentication is enabled on the NAS.

Command	**dot1x port-control-mode port-based**
Parameter Description	N/A
Defaults	The default port control mode is MAC-based control.
Command Mode	Interface configuration mode
Usage Guide	You can configure the port-based control mode if the remaining users can access the network after a user on a controlled port passes authentication.

Enabling the Single-User Port-based Control Mode

- (Optional) Configure only one dynamic user to access the network in port-based authentication mode.
- Enable the single-user port-based control mode after 802.1X authentication is enabled on the NAS.

Command	**dot1x port-control-mode port-based single-host**
Parameter Description	N/A
Defaults	The single-user port-based control mode is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	Configure this command when only the authenticated user can act as a dynamic user in port-based control mode.

Disabling Migration of Dynamic Users

- (Optional) If this function is disabled, dynamic users on a controlled port cannot migrate to other ports until the port has aged.
- Disable this function after 802.1X authentication is enabled on the NAS.

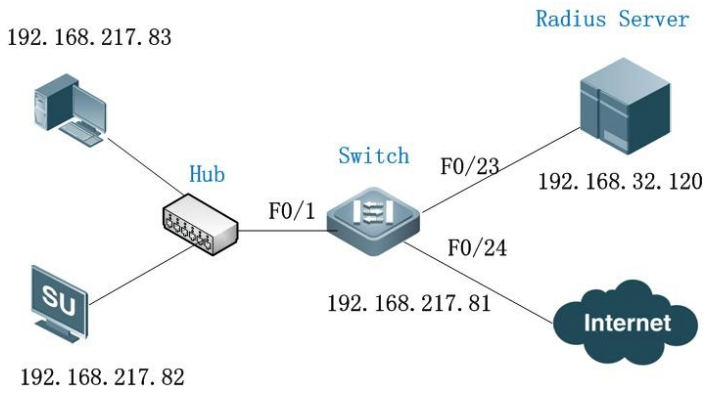
Command	**dot1x stationarity enable**
Parameter Description	N/A
Defaults	Dynamic users can migrate to other ports by default.
Command Mode	Global configuration mode
Usage Guide	Configure this command to prohibit dynamic users on a controlled port from migrating to other ports.

Verification

- In MAC-based control mode, each user on a controlled port can access the network only after authenticated.
- In port-based control mode, as long as a user on a controlled port passes authentication, other users can access the network without authentication.

Configuration Example

Enabling the Port-based Control Mode

Scenario **Figure 4- 11**	
Configuration Steps	● Register the IP address of the NAS on the RADIUS server and configure the communication key between the NAS and the RADIUS server. ● Create an account on the RADIUS server. ● Enable AAA on the NAS. ● Configure RADIUS parameters on the NAS. ● Enable 802.1X authentication on ports of the NAS. ● Enable port-based authentication on a controlled port. NAS configurations are as follows. For detailed configuration on the RADIUS server, see the *Configuring RADIUS*.
	FS# configure terminal FS (config)# aaa new-model FS (config)# radius-server host 192.168.32.120 FS (config)# radius-server key FS FS (config)# interface FastEthernet 0/1 FS (config-if)# dot1x port-control auto FS (config-if)# dot1x port-control-mode port-based

Verification	Check whether authentication is proper, network access behaviors change after authentication, and dynamic users can access the network. ● The account is successfully created, such as **username:tests-user,password:test**. ● The user fails to ping 192.168.32.120 before authentication. ● After the user enters account information and click **Authenticate** on FS Supplicant, the authentication succeeds and the user can successfully ping 192.168.32.120. ● After passing authentication, dynamic users can successfully ping 192.168.32.120. ● Information of the authenticated user is displayed. FS# show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-State Port-Status User-Type Time --------- ---------- ---------- --------- ---- ---------- ---------- ---------- --------- ---------- 16778217 ts-user 0023.aeaa.4286 Fa0/1 2 Authenticated Idle Authed static 0days 2h17m29s none N/A 0023.aeaa.4286 Fa0/1 2 Authenticated Idle Authed Dynamic N/A

4.4.8 Configuring Dynamic VLAN Assignment

Configuration Effect

- Enable 802.1X-based dynamic VLAN assignment for a port. If the authentication server assigns a VLAN to redirect after a user passes authentication, the NAS can add this user to the assigned VLAN to perform authorization on this user.
- Controlled ports on the VLAN to redirect fall in three types: Access, Trunk, and Hybrid (MAC VLAN is disabled). You can change native VLANs of these ports to realize 802.1X-based dynamic VLAN assignment.
- If controlled ports on the VLAN to redirect are Hybrid ports (and MAC VLAN is enabled), dynamically create MAC VLAN entries to add users to the assigned VLAN.

Notes

- The NAS can extend RADIUS attributes to assign VLANs. When assigning VLANs to the access switch based on extended attributes, the RADIUS server encapsulates these attributes in RADIUS Attribute 26, with the vendor ID of 0x00001311. The default type No. of the extended attribute is 4. You can run the **radius attribute 4 vendor-type type** command on the NAS to receive the VLAN of which the extended attribute type No. is set to **type**. For details about the command, see the *Configuring RADIUS*.
- The RADIUS server can assign VLANs based on the following RADIUS attributes:

  Attribute 64: Tunnel-Type, with the value being VLAN (13).

  Attribute 65: Tunnel-Medium-Type, with the value being 802 (6).

  Attribute 81: Tunnel-Private-Group-ID, which can be the VLAN ID or VLAN name.
- The NAS can perform 802.1X authentication on Access, Trunk, and Hybrid ports. If 802.1X-based dynamic VLAN assignment is enabled on other ports, authentication will fail.
- If the assigned VLAN is the VLAN name, the system checks whether the VLAN name exists on the access switch. If yes, the port of the user redirects to this VLAN. If no, the NAS identifies the assigned VLAN as the VLAN ID. If the VLAN ID is valid (in the VLAN ID range supported by the system), the port of the user redirects to this VLAN. If the VLAN ID is 0, no VLAN information is assigned. In

other cases, users fail authentication.

- Private VLANs, remote VLANs, or super VLANs (including sub VLANs) cannot be assigned for redirection.
- In dynamic VLAN assignment on an Access port, check whether any assigned VLAN is configured on the switch:
  - Yes: If the Access port can redirect to the assigned VLAN, the port will leave the configured VLAN and migrate to the assigned VLAN, and user authentication will succeed. Otherwise (see the related description below), user authentication will fail.
  - No: If the NAS identifies the assigned VLAN attribute as the VLAN ID, it will create a VLAN and enable the port to redirect to the new VLAN, and user authentication will succeed. If the NAS identifies the assigned VLAN attribute as the VLAN name, it will fail to find the corresponding VLAN ID, causing authentication failure.
- In dynamic VLAN assignment on a Trunk port, check whether any assigned VLAN is configured on the switch:
  - Yes: If the Trunk port can redirect to the assigned VLAN, the NAS will use the native VLAN of the port as the assigned VLAN, and user authentication will succeed. Otherwise (see the related description below), user authentication will fail.
  - No: If the NAS identifies the assigned VLAN attribute as the VLAN ID, it will use the native VLAN of the port, and user authentication will succeed. If the NAS identifies the assigned VLAN attribute as the VLAN name, it will fail to find the corresponding VLAN ID, causing authentication failure.
- If MAC VLAN is disabled on a Hybrid port, check whether any assigned VLAN is configured on the switch:
  - Yes: If the Hybrid port can redirect to the assigned VLAN or the assigned VLAN does not exist in the tagged VLAN list of the Hybrid port, the NAS will allow the assigned VLAN to pass through the Hybrid port without carrying any tags and uses the native VLAN as the assigned VLAN, and user authentication will succeed. Otherwise (see the related description below), user authentication will fail.
  - No: If the NAS identifies the assigned VLAN attribute as the VLAN ID, it will create a VLAN, allow the VLAN to pass through the Hybrid port without carrying any tags, and use the native VLAN as the assigned VLAN, and user authentication will succeed. If the NAS identifies the assigned VLAN attribute as the VLAN name, it will fail to find the corresponding VLAN ID, causing authentication failure.
- If MAC VLAN is enabled on a Hybrid port, VLAN assignment is as follows:
  If the VLAN assigned by the authentication server does not exist on the NAS (MAC VLAN requires VLANs to have static configurations), or has been added to the Hybrid port with tags, or is not supported by MAC VLAN (see the *Configuring MAC VLAN*), user authentication will fail. Otherwise, the NAS will dynamically create MAC VLAN entries based on the assigned VLAN and the MAC addresses of users, and user authentication will succeed. When users go offline, MAC VLAN entries will be dynamically removed.
- If MAC VLAN is disabled on a port, VLAN assignment changes only the native VLAN but not the **native vlan** command configurations of the port. The assigned VLAN takes priority over the VLAN configured in related commands. That is, the native VLAN effective after authentication acts as the assigned VLAN while the native VLAN configured in related commands takes effect only when users go offline.
- If MAC VLAN is enabled on a port and user authentication is based on the MAC address, VLAN assignment dynamically creates MAC VLAN entries without changing the native VLAN of the port.
- No matter MAC VLAN is enabled or not on a Hybrid port, if the assigned VLAN is added to the port with tags, VLAN assignment fails.
- If MAC VLAN is enabled on a port (see the *Configuring MAC VLAN*), VLAN assignment creates an MAC VLAN entry with an all-F mask. If the MAC address of an 802.1X user is overwritten by the MAC address specified by the new MAC VLAN entry, the assigned VLAN must be the same as the VLAN specified by the new MAC VLAN entry. Otherwise, errors will occur to 802.1X users in VLAN assignment. Errors are as follows (including but not limited to): User authentication succeeds but subsequent valid data packets are discarded, causing network access failure.
  When a user goes offline by sending an EAPOL-LOGOFF packet, the 802.1X authentication entry remains on the NAS and the user status on the authentication server is still online.

Configuration Steps

Enabling Dynamic VLAN Assignment on a Port

- (Optional) After dynamic VLAN assignment is enabled on a port, authenticated users on this port will enter the assigned VLAN.
- Enable dynamic VLAN assignment after 802.1X authentication is enabled on the NAS.

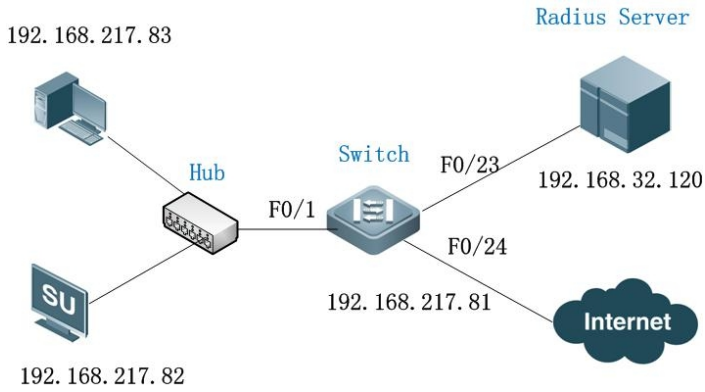
Command	**dot1x dynamic-vlan enable**
Parameter Description	N/A
Defaults	Dynamic VLAN assignment is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	Configure this command when authenticated users should be added to the VLAN assigned by the authentication server.

Verification

- Run the **show dot1x summary** command to display the VLAN of a user and the **show dot1x user id** command to display the VLAN assigned by the RADIUS server.
- Users with VLANs assigned can access the network in the assigned VLANs.

Configuration Example

Enabling Dynamic VLAN Assignment on a Port

Scenario **Figure 4- 12**	
Configuration Steps	● Register the IP address of the NAS on the RADIUS server and configure the communication key between the NAS and the RADIUS server. ● Create an account on the RADIUS server. ● Enable AAA on the NAS. ● Configure RADIUS parameters and enable VLAN delivery on the NAS. ● Enable 802.1X authentication on ports of the NAS. ● Enable dynamic VLAN assignment on a controlled port. NAS configurations are as follows. For detailed configuration on the RADIUS server, see the *Configuring RADIUS*.
	FS# configure terminal FS (config)# aaa new-model FS (config)# radius-server host 192.168.32.120

	FS (config)# radius-server key FS FS (config)# interface FastEthernet 0/1 FS (config-if)# dot1x port-control auto FS (config-if)# dot1x dynamic-vlan enable
Verification	Check whether authentication is proper, network access behaviors change after authentication, and dynamic users can access the network. ● The account is successfully created, such as **username:tests-user,password:test**. ● The user fails to ping 192.168.32.120 before authentication. ● After the user enters account information and click **Authenticate** on FS Supplicant, the authentication succeeds and the user can successfully ping 192.168.32.120. ● After passing authentication, dynamic users can successfully ping 192.168.32.120. ● Information of the authenticated user is displayed, showing that the user jumps from VLAN 2 to VLAN 3. FS# show dot1x summary ID Username MAC Interface VLAN Auth-State Backend-State Port-Status User-Type Time --- --- --- --- --- --- --- --- --- 16778217 ts-user 0023.aeaa.4286 Fa0/1 3 Authenticated Idle Authed static 0days 2h17m29s

Common Errors

- RADIUS attributes for VLAN assignment are incorrectly configured on the authentication server.
- RADIUS attribute support for VLAN assignment is disabled on the NAS.
- When MAC VLAN is enabled on a Hybrid port for dynamic VLAN assignment, the assigned VLAN has tags.

4.4.9 Configuring the Guest VLAN

Configuration Effect

- If no 802.1X authentication client is available on a controlled port, add the port to the guest VLAN so that users without any authentication clients can temporarily access the network in the guest VLAN.
- If the NAS receives an EAPOL packet after adding a port to a guest VLAN, it regards that this port has an 802.1X authentication client. Then this port is forced out of the guest VLAN to perform 802.1X authentication.

Notes

- A controlled port has no 802.1X authentication client if any one of the following conditions is met:
  1. The port sends three consecutive active authentication packets but does not receive any EAPOL replies within the specified period (**auto-req req-interval** x 3).
  2. The port does not receive any EAPOL replies within 90 seconds.
  3. MAB fails.
- 802.1X-based dynamic VLAN assignment must be enabled for a port.
- When the port status switches from up to down, the port exits from the guest VLAN. When the port status switches from down to up, the NAS re-checks whether to add this port to the guest VLAN.

- If failing to receive eapol packets after 90s, an interface enters the guest VLAN. Because of the increment mechanism of sending shcp discover packets, it may take a long time for a downlink terminal to initiate a dhcp request again. Therefore, the interface cannot obtain the ip address promptly.

Configuration Steps

Configuring the Guest VLAN

- (Optional) After the guest VLAN is configured on a port, check whether the port has 802.1X authentication clients. If no, add the port to the guest VLAN.
- Configure the guest VLAN after 802.1X authentication is enabled on the NAS.

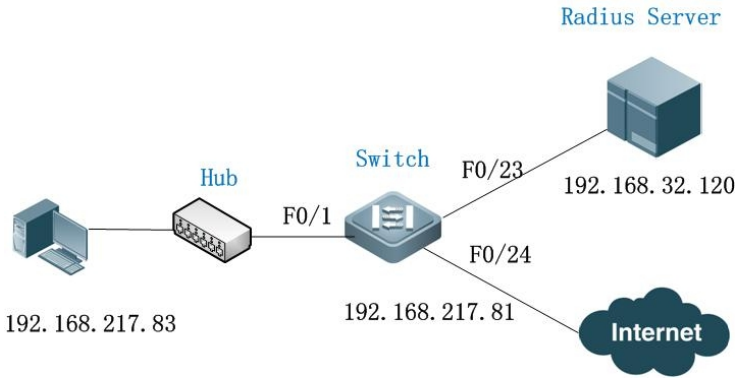
Command	**dot1x guest-vlan** *vid*
Parameter Description	*vid*: Indicates the guest VLAN to join.
Defaults	The guest VLAN is not configured by default.
Command Mode	Interface configuration mode
Usage Guide	Configure this command when a user connects to an 802.1X controlled port but has no authentication client. When guest VLAN is enabled on a port, do not configure Layer-2 attributes, and specially do not manually set the VLAN of the port.

Verification

- After a port switches to the guest VLAN, users connected to the port can communicate only in the guest VLAN.
- If a user connected to a port in the guest VLAN installs an 802.1X authentication client and initiates authentication, the port will exit the guest VLAN.

Configuration Example

Configuring Dynamic VLAN Assignment and Guest VLAN

Scenario **Figure 4- 13**	
Configuration Steps	● Enable 802.1X authentication on ports of the NAS. ● Enable dynamic VLAN assignment on a controlled port. ● Configure the guest VLAN on a controlled port. NAS configurations are as follows:
	FS (config)# interface FastEthernet 0/1

	FS (config-if)# dot1x port-control auto FS (config-if)# dot1x dynamic-vlan enable FS (config-if)# dot1x guest-vlan 3
Verification	Check whether network access behaviors change after a port joins a guest VLAN. ● Users cannot communicate before the port joins the guest VLAN while can communicate after that. The NAS prints the log as follows: %DOT1X-5-TRANS_DEFAULT_TO_GUEST: Transformed interface Fa0/1 from default-vlan 1 to guest-vlan 3 OK.

Common Errors

- A port receives an EAPOL packet, causing its failure to join the guest VLAN.

4.4.10 Configuring the Failed VLAN

Configuration Effect

- Configure the failed VLAN on an 802.1X controlled port. If a user fails authentication after failed VLAN is enabled, the port can be added to a failed VLAN so that the user can still access the network.
- Configure the maximum number of consecutive authentication failures. If this number is exceeded, the NAS adds the port to a failed VLAN.

Notes

- If the failed VLAN configured does not exist, a failed VLAN will be dynamically created when a port enters the failed VLAN and automatically removed when the port exits the failed VLAN.
- 802.1X-based dynamic VLAN assignment must be enabled for a port.
- If a port goes down, the port will automatically exit the failed VLAN.
- The failed VLAN and guest VLAN can be configured to the same VLAN.
- In port-based control mode, after a controlled port enters a failed VLAN, only users failing authentication can re-initiate authentication and other users' authentication requests will be discarded. This restriction does not exist in MAC-based control mode.
- Failed VLAN does not support private VLANs. That is, private VLANs cannot be configured as 802.1X failed VLANs.
- If GSN address binding is enabled on a port, users in a failed VLAN cannot access the network.

Configuration Steps

Configuring the Failed VLAN

- (Optional) If the failed VLAN is configured, the NAS adds users rejected by the authentication server to a failed VLAN.
- Configure the failed VLAN after 802.1X authentication is enabled on the NAS.

Command	**dot1x auth-fail vlan** *vid*
Parameter Description	*vid*: Indicates the failed VLAN to join.
Defaults	Failed VLAN is disabled by default.
Command Mode	Interface configuration mode

Usage Guide	Configure this command if users need to access the network even after authentication fails.

Configuring the Maximum Number of Failed VLAN Attempts

- (Optional) Configure the maximum number of times when a user is rejected by the authentication server. If this number is exceeded, the port can be added to a failed VLAN.
- Configure the maximum number of failed VLAN attempts after 802.1X authentication is enabled on the NAS.

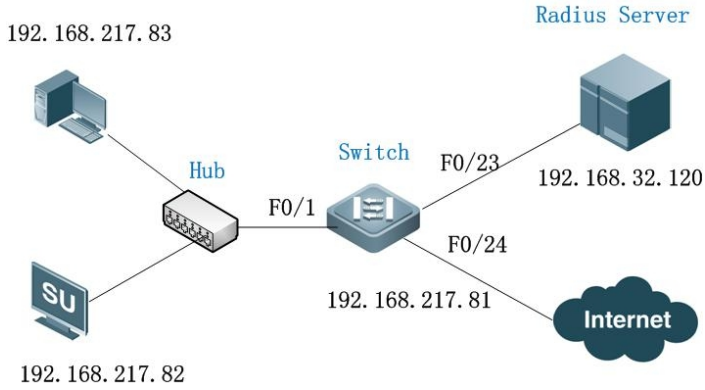
Command	**dot1x auth-fail max-attempt** *value*
Parameter Description	*value*: Indicates the maximum number of times when a user fails authentication.
Defaults	The default value is 3.
Command Mode	Interface configuration mode
Usage Guide	Configure this command when the maximum number of failed VLAN attempts needs to be adjusted.

Verification

- When a port switches to a failed VLAN, users connected to the port can communicate only in the failed VLAN.

Configuration Example

Configuring the Failed VLAN

Scenario **Figure 4- 14**	
Configuration Steps	● Register the IP address of the NAS on the RADIUS server and configure the communication key between the NAS and the RADIUS server. ● Create an account on the RADIUS server. ● Enable AAA on the NAS. ● Configure RADIUS parameters on the NAS. ● Enable 802.1X authentication on ports of the NAS. ● Enable port-based authentication on a controlled port. NAS configurations are as follows. For detailed configuration on the RADIUS server, see the *Configuring RADIUS*.

	FS# configure terminal FS (config)# aaa new-model FS (config)# radius-server host 192.168.32.120 FS (config)# radius-server key FS FS (config)# interface FastEthernet 0/1 FS (config-if)# dot1x port-control auto FS (config-if)# dot1x auth-fail vlan 3
Verification	Check whether authentication is proper, network access behaviors change after authentication, and dynamic users can access the network. ● The account is successfully created, such as **username:tests-user,password:test**. ● The user fails to ping 192.168.32.120 before authentication. ● Start FS Supplicant, enter incorrect account information, and click **Authenticate**. The authentication fails, the user can successfully ping the IP address of a failed VLAN. ● Information of the authenticated user is displayed. FS(config)#show dot1x user name ts-user Supplicant information: MAC address ..... b048.7a7f.f9f3 Username ..... ts-user User ID ..... 16777303 Type ..... static VLAN ..... 1 Port ..... wlan 1 Online duration ..... 0days 0h 0m21s Up average bandwidth ..... 0 kBps Down average bandwidth ..... 0 kBps Authorized VLAN ..... 1 Authorized session time ..... 20736000 seconds Authorized flux ..... unlimited Accounting ..... No Proxy user ..... Permit Dial user ..... Permit IP privilege ..... 0 Private supplicant ..... no Authorized by Auth-Fail-Vlan ..... 3 Max user number on this port 0

Common Errors

- If a user fails authentication not due to rejection of the authentication server, for example, due to installation failure as a result of hardware resource insufficiency, it cannot enter the failed VLAN.

4.4.11 Configuring Extended Functions

Configuration Effect

- Some users use authentication clients embedded in the operating system. These clients may not initiate authentication immediately after the users access the network, affecting user experience on network access. Enable active authentication to so that such users can initiate authentication immediately after accessing the network.
- Active authentication means that the NAS sends a request/id packet to trigger FS Supplicant to perform 802.1 authentication. Therefore, you can use this function to detect whether FS Supplicant is used. For example, this function is required for MAB deployment.
- Configure the authenticable host list to specify users that can be authenticated on the port, which restricts physical access points of users to enhance network security
- The multi-account function allows a user to switch its account upon re-authentication. In special scenarios such as Windows domain authentication, multiple authentications are required to access the domain and the user account changes during authentication. This function applies to these scenarios.
- By default, the NAS uses its own MAC address as the source MAC address of EAP packets during 802.1X authentication. Some versions of FS supplicants check whether the access switch is a FS switch based on the MAC address of EAP packets and implement some private features. When performing 802.1X authentication with these supplicants, you can enable the virtual source MAC address to use related private features.
- 802.1X allows users to obtain IP addresses before accounting. In this manner, the IP address is carried during user accounting, meeting service requirements. After a user is authenticated and goes online, the NAS can obtain the IP address of the user from the supplicant or through DHCP snooping, and then 802.1X server initiates an accounting request. To avoid the case in which the NAS does not initiate accounting for a long time due to failure to obtain the IP address of the authentication client, configure the IP detection timeout for this function. If the NAS does not obtain the IP address of the user within the configured time (5 minutes by default), it forces the user offline.
- If 802.1X authentication and MAB are enabled, you can specify the priorities of 802.1X authentication and MAB on the same port. By default, the one triggered first should be first performed and 802.1X authentication takes priority over MAB. That is, if the NAS receives an EAPOL packet from a user after the user performs MAB, the original MAB user goes offline and performs 802.1X authentication. At present, you can configure each user to first perform 802.1X authentication. If a user performs MAB following the 802.1X authentication failure and MAB takes priority over 802.1X authentication, the user can ignore 802.1X authentication after passing MAB.
- Based on a real scenario, H3C devices are deployed and the MAB authentication server configures the user name in xx-xx-xx-xx-xx-xx format. However, the default MAB authentication user name format of FS devices is xxxxxxxxxxxx. Therefore, a command needs to be added to control the user name format.

Notes

- The multi-account function must be disabled if accounting is enabled. Otherwise, accounting may be inaccurate.
- MAB requires active authentication. Therefore, active authentication must be enabled if MAB is enabled.
- IP-based accounting is not required in two situations:
  - IPv4 addresses and FS Supplicant are deployed. This function is not required because FS Supplicant can upload the IPv4 addresses of users.
  - Static IP addresses are deployed.

Configuration Steps

Enabling Active Authentication

- (Optional) If active authentication is enabled, the controlled port sends an authentication request actively after configuration. After

receiving this request, the authentication client initiates 802.1X authentication.

- Enable active authentication after 802.1X authentication is enabled on the NAS.

Command	**dot1x auto-req**
Parameter Description	N/A
Defaults	Apart from on N1800K switches, active authentication is enabled by default.
Command Mode	Global configuration mode
Usage Guide	The destination addresses of active authentication packets are the multicast address. If the connected clients may not initiate authentication automatically, configure this command to make the NAS actively initiate authentication. When controlled ports are Trunk ports, enable active authentication so that authentication requests can be sent based on each VLAN of trunk ports.

Configuring the Number of Active Authentication Requests

- (Optional) Configure the number of active authentication requests sent by the NAS.
- Configure the number of active authentication requests after 802.1X authentication is enabled on the NAS.

Command	**dot1x auto-req packet-num***num*
Parameter Description	*num*: Indicates the number of active authentication requests.
Defaults	The number of active authentication request is not configured by default.
Command Mode	Global configuration mode
Usage Guide	If active authentication is enabled, configure this command to restrict the number of active authentication packets sent by a port and thereby avoid sending excessive packets.

Enabling User Detection for Active Authentication

- (Optional) Configure the NAS not to send authentication requests actively if there are authenticated users on a controlled port.
- Enable user detection for active authentication after 802.1X authentication is enabled on the NAS.

Command	**dot1x auto-req user-detect**
Parameter Description	N/A
Defaults	User detection for active authentication is enabled by default.
Command Mode	Global configuration mode
Usage Guide	After this command is configured, the NAS does not send authentication packets actively if there are authenticated users on controlled Access ports. On Trunk ports, the NAS checks for authenticated users based each VLAN. If there are authenticated users on a VLAN, the NAS does not send authentication packets automatically.

Configuring the Interval of Active Authentication Request

- (Optional) Configure the interval at which the NAS sends an authentication request actively.

- Enable the interval of active authentication request after 802.1X authentication is enabled on the NAS.

Command	**dot1x auto-req req-interval** *time*
Parameter Description	*Time*: Indicates the interval of active authentication request.
Defaults	The default value is 30s.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Authenticatable Client List

- (Optional) Configure the authenticable client list on a controlled port. Only clients on the list can perform 802.1X authentication.
- Configure the authenticable client list after 802.1X authentication is enabled on the NAS.

Command	**dot1x auth-address-table address** *mac-addr* **interface** *interface*
Parameter Description	*mac-addr*: Indicates the MAC address of the access user. *interface*: Indicates the port of the access user.
Defaults	All users can perform authentication.
Command Mode	Global configuration mode
Usage Guide	Configure this command when specified users should be able to perform authentication on a controlled port.

Enabling 802.1X Packets Sending with the Pseudo Source MAC Address

- (Optional) Configure the **dot1x pseudo source-mac** command when FS Supplicant fails to identify the NAS as a FS device based on the MAC address of the NAS.
- Configure the pseudo MAC address as the source MAC address for 802.1X authentication after 802.1X authentication is enabled on the NAS.

Command	**dot1x pseudo source-mac**
Parameter Description	N/A
Defaults	User detection for active authentication is enabled by default.
Command Mode	Global configuration mode
Usage Guide	Configure this command when FS Supplicants cannot identify the NAS as a FS device based on the source MAC address in the EAPOL packet sent by the NAS or implement private attributes during authentication. If this command is configured, the EAPOL packet sent by the NAS uses 00-1A-A9-17-FF-FF as the source MAC address so that these FS Supplicants can identify the NAS as a FS device.

Enabling Multi-account Authentication with One MAC Address

- (Optional) Run the **dot1x multi-account enable** command to allow the same MAC address to be used by multiple accounts.
- Enable multi-account authentication with one MAC address after 802.1X authentication is enabled on the NAS.

Command	**dot1x multi-account enable**
Parameter	N/A

Description	
Defaults	Multi-account authentication is disabled by default.
Command Mode	Global configuration mode
Usage Guide	Configure this command when multi-account authentication is required in 802.1X authentication, e.g. in the case of Windows domain authentication. In this case, the authentication client can directly use a new account to initiate authentication while the previous account is still online. Multi-account authentication is disabled by default.

Configuring the Maximum Number of Authenticated Users on a Port

- (Optional) You can restrict the number of online users on a controlled port, including static users and dynamic users.
- Configure the maximum number of authenticated users on a port after 802.1X authentication is enabled on the NAS.

Command	**dot1x default-user-limit** *num*
Parameter Description	*num*: Indicates the maximum number of online users.
Defaults	There is no restriction on the number of users on a port by default.
Command Mode	Interface configuration mode
Usage Guide	Configure this command when there is a need to restrict the number of authenticated users on a port.

Enabling IP-triggered Accounting

- (Optional) If IP-triggered accounting is enabled, the NAS sends an accounting request to the authentication server after obtaining the IP address of the user.
- Enable IP-triggered accounting after 802.1X authentication is enabled on the NAS.

Command	**dot1x valid-ip-acct enable**
Parameter Description	N/A
Defaults	IP-triggered accounting is disabled by default.
Command Mode	Global configuration mode
Usage Guide	If both accounting and IP-triggered accounting are enabled, the NAS initiates accounting only after obtaining the IP address of the authentication client, and forces the user offline if it fails to obtain the IP address. If accounting is disabled but IP-triggered accounting is enabled, the NAS does not initiate accounting after obtaining the IP address of the authentication client, and forces the user offline if it fails to obtain the IP address within the timeout.

Configuring the Timeout of Obtaining IP Addresses After Authentication

- (Optional) Configure the timeout of obtaining IP addresses if IP-triggered accounting is enabled.
- Configure the IP address obtaining timeout after 802.1X authentication is enabled on the NAS.

Command	**dot1x valid-ip-acct timeout** *time*
Parameter Description	*time:* Indicates the timeout in the unit of minutes.
Defaults	The default value is 5 minutes.
Command	Global configuration mode

Mode	
Usage Guide	It is recommended to use the default value. Configure this command when there is a need to change the IP address obtaining timeout after users pass authentication.

Enabling 802.1X Precedence over MAB

- (Optional) If 802.1X precedence over MAB is enabled, a user first performs 802.1X authentication. If 802.1X authentication fails, the user initiates MAB and MAB takes priority of 802.1 authentication.
- 802.1X authentication and MAB must be enabled on a port.

Command	**dot1x auth-with-order**
Parameter Description	N/A
Defaults	802.1X precedence over MAB is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	Configure this command when a user needs to first perform 802.1X authentication and initiate MAB if 802.1X authentication fails, and MAB takes priority of 802.1 authentication.

Configuring the Compatibility Function for H3C 802.1X Authentication Clients and Authentication Servers

- (Optional) This function is effective to 802.1X-authenticated and MAB-authenticated users.

Command	**dot1x user-name compatible**
Parameter Description	-
Defaults	This function is disabled by default.
Command Mode	Global configuration mode
Usage Guide	Enable this function when the H3C authentication client and authentication server are used for 802.1X authentication or the H3C authentication server is used for MAB authentication.

Disabling Global 802.1X

- (Optional) This function is effective to both 802.1x and MAB-authenticated users.

Command	**dot1x system disable**
Parameter Description	-
Defaults	By default, global 802.1x is enabled.
Command Mode	Global configuration mode
Usage Guide	When the server is unreachable, disable global 802.1x, so users can access the Internet without authentication. After the server resumes reachability, enable global 802.1x, and users have to pass authentication before accessing the Internet.

4.5 Monitoring

Clearing

Authentication user information can be cleared after 802.1X is disabled.

Description	Command
Clears 802.1X user information.	**no do1x port-control auto**
Clears 802.1X user information.	**clear dot1x user**
Restores the default 802.1X configuration.	**dot1x default**

Notes

- The **dot1x default** command is used to restore global configurations.

Description	Command
Restore the default value of status machine timeout duration.	**dot1x timeout quiet-period** **dot1x timeout server-timeout** **dot1x timeout supp-timeout** **dot1x timeout tx-period**
Restore default values of configurations related to re-authentication.	**dot1x re-authentication** **dot1x timeout re-authperiod** **dot1x reauth-max**
Restore default values of configurations related to proactive requests.	**dot1x auto-req** **dot1x auto-req user-detect** **dot1x auto-req req-interval** **dot1x auto-req packet-num**
Restores the default value of the number of retransmission times.	**dot1x mac-req**
Restores the default value of the authentication mode.	**dot1x auth-mode**
Restore the default values of configurations related to client probing.	**dot1x client-probe enable** **dot1x probe-timer alive** **dot1x probe-timer interval**
Restores the default value of the function of supporting only the private client.	**dot1x private-supplicant-only**
Restores the default value of the pseudo source MAC address function.	**dot1x pseudo source-mac**
Restores the default value of the number of VLAN redirection times upon authentication failures.	**dot1x auth-fail max-attempt**
Restores the default value of the function of one MAC address for multiple accounts.	**dot1x multiaccount enable**
Restores the default value of the dot1x redirection function.	**dot1x redirect**
Restores the default value of the silent timeout duration.	**dot1x multi-mab quiet-period**
Restore the default values of functions related to accounting after obtaining the IP address.	**dot1x valid-ip-acct enable** **dot1x valid-ip-acct timeout**

Displaying

Description	Command
Displays the parameters and status of the RADIUS server.	**show radius server**
Displays 802.1X status and parameters.	**show dot1x**
Displays the authenticable host list.	**show dot1x auth-address-table**
Displays the active authentication status.	**show dot1x auto-req**
Displays the port control status.	**show dot1x port-control**
Displays the status and parameters of host probe.	**show dot1x probe-timer**
Displays of the information of authenticated users.	**show dot1x summary**
Displays the maximum times of EAP-Request/Challenge packet retransmission.	**show dot1x max-req**
Displays the information of controlled ports.	**show dot1x port-control**
Displays the non-FS client filtering information.	**show dot1x private-supplicant-only**
Displays the re-authentication status.	**show dot1x re-authentication**
Displays the maximum times of EAP-Request/Identity packet retransmission.	**show dot1x reauth-max**
Displays the quiet period after authentication fails.	**show dot1x timeout quiet-period**
Displays the re-authentication interval.	**show dot1x timeout re-authperiod**
Displays the authentication server timeout.	**show dot1x timeout server-timeou**t
Displays the supplicant timeout.	**show dot1x timeout supptimeout**
Displays the interval of EAP-Request/Identity packet retransmission.	**show dot1x timeout tx-period**
Displays user information based on the user ID.	**show dot1x user id**
Displays user information based on the MAC address.	**show dot1x user mac**
Displays user information based on the user name.	**show dot1x user name**

Debugging

System resources are occupied when debugging information is output. Therefore, disable the debugging switch immediately after use.

Description	Command

Debugs AAA. (For details, see the *Configuring AAA*.)	**debug aaa**
Debugs RADIUS. (For details, see the *Configuring RADIUS*.)	**debug radius**
Debugs 802.1X events.	**debug dot1x event**
Debugs 802.1X packets.	**debug dot1x packet**
Debugs 802.1X state machine (STM).	**debug dot1x stm**
Debugs 802.1X internal communication.	**debug dot1x com**
Debugs 802.1X errors.	**debug dot1x error**

5 Configuring Web Authentication

5.1 Overview

5.1.1 Web Authentication

Web authentication controls user access to networks. It requires no authentication software on clients. Instead, users can perform authentication on common browsers.

When unauthenticated clients attempt to access the Internet using browsers, the network access server (NAS) forcibly redirects the browsers to a specified site pointing to a Web authentication server, also called a portal server. Users can access the services on the portal server before being authenticated, such as downloading security patches and reading notices. If a user wants to access network resources beyond the portal server, the user must get authenticated by the portal server through a browser.

Besides providing convenient authentication, the portal server performs Webpage interaction with browsers, providing personalized services, such as advertisements, notices, and business links on the authentication page.

FS Web Authentication Versions

There are three versions of FS Web authentication, including FS First-Generation Web Authentication and FS Second-Generation Web Authentication. The Web authentication process varies with authentication versions. For details, see Section 5.3 "Features".

The three versions of Web authentication are highly divergent in features and configurations. It is recommended to read through the relevant chapters carefully before configuration.

Both FS Second-Generation Web Authentication supports local account authentication on the NAS. Because Remote Authentication Dial In User Service (RADIUS) authentication is more commonly used in reality, it is used as an example in the chapter "Applications".

The concept of "interface" varies with product types. For example, the interfaces on a layer-2 switch are physical ports. This document uses the unified term "interface" to include them. In application, recognize the real meaning based on specific products and functions.

Web authentication supports user online traffic detection. For details, see the Configuring SCC.

Web authentication supports the authentication of domain names. That is, accounts can be authenticated in the format of user name@domain name. This requires enabling the domain-name-based authentication, authorization and accounting (AAA) service. For details, see the Configuring AAA.

Protocols and Standards

- HTTP: RFC1945 and RFC2068
- HTTPS: RFC2818
- SNMP: RFC1157 and RFC 2578
- RADIUS: RFC2865, RFC2866, and RFC3576

5.2 Applications

Application	Description
Basic Scenario of Web Authentication	Basic layer-2 authentication scenario, where a NAS, portal server, and RADIUS server constitute an authentication system which connects a client with the NAS through the layer-2 network.

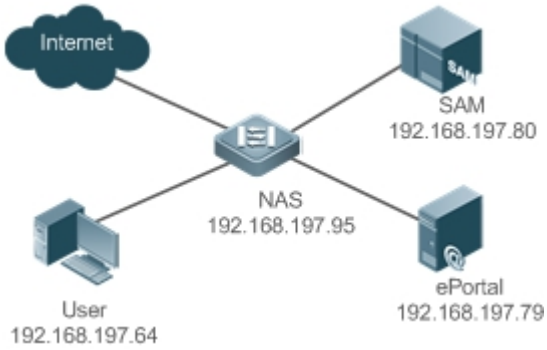
5.2.1 Basic Scenario of Web Authentication

Scenario

See Figure 5- 1.

- Deploy a Web authentication scheme on the NAS.
- The client connected to the NAS needs to pass Web authentication before accessing the Internet.

Figure 5- 1 Networking Topology of Web Authentication



Remarks	Web authentication is applicable to both layer-2 and layer-3 networks. At layer 3, the source MAC address and VID of a packet are changed after it is routed, but the source IP address remains the same as the only identifier of a client. Therefore, the binding policy of Web authentication on layer-3 devices must adopt the IP-only binding mode. Here, layer-2 NAS is used as an example. FS-SAM program is installed on the RADIUS server. FS-ePortal program is installed on the portal server.

Deployment

- Enable Web authentication on the client-accessed interface or globally on the NAS.
- Configure the ePortal server and the communication key on the NAS (for only FS First-Generation and Second-Generation Web Authentication).
- Configure the Simple Network Management Protocol (SNMP) communication parameters of the ePortal server on the NAS (for only FS First-Generation and Second-Generation Web Authentication).
- Configure the consistent communication parameters on the ePortal server and SAM server (for only FS First-Generation Web Authentication).
- Create user accounts on the SAM server.
- Configure AAA and method lists on the NAS (for only FS Second-Generation).
- Configure the IP address of the SAM server on the NAS (for only FS Second-Generation).
- Configure the names of the Web authentication method lists on the NAS (for only FS Second-Generation).

5.3 Features

Basic Concepts

↘ FS First-Generation Web Authentication

FS First-Generation Web Authentication should cooperate with the FS-ePortal software. The server installed with FS-ePortal provides a

login page to submit user authentication information, and initiates an authentication request to the RADIUS server directly. After authentication succeeds, the NAS gets user information delivered through the SNMP protocol, and thereby controls user access permissions. Communication during Web authentication of this version depends on private SNMP nodes. Moreover, the ePortal server takes the place of the NAS in authentication and accounting, which relieves the NAS from service burden.

FS Second-Generation Web Authentication

FS Second-Generation Web Authentication complies with the *CMCC WLAN Service Portal Specification*. The portal server is responsible only for Webpage interaction with users. The NAS interacts with the RADIUS server to implement authentication. The interaction between the portal server and the NAS complies with the *CMCC WLAN Service Portal Specification*. The portal server provides a login page for users to submit their information, and informs the NAS of user information through the portal protocols. The NAS completes authentication by interacting with the RADIUS server based on the user information, assigns access permissions to authenticated clients, and returns authentication results to the portal server.

The implementation process of FS Second-Generation Web Authentication is mainly completed on the NAS. This raises a higher demand on the NAS's capability to handle heavy tasks. Meanwhile, the portal server is simplified. The standard *CMCC WLAN Service Portal Specification*, which gains highly industry support, enables various vendors to develop compatible products.

Version Comparison

Authentication roles:

- Client: Its functions are the same among the three types of Web authentication.
- NAS: In FS First-Generation Web Authentication, the NAS implements only URL redirection and exchanges user login/logout notifications with the portal server. In FS Second-Generation Web Authentication, the NAS is responsible for redirecting and authenticating users as well as notifying the portal server of authentication results.
- Portal server: In FS First-Generation Web Authentication, the portal server is responsible for interaction with clients through Webpages, authenticating users, and notifying the NAS of authentication results. In FS Second-Generation Web Authentication, the portal server is responsible for interacting with clients through Webpages, notifying the NAS of users' authentication information, and receiving authentication results from the NAS.
- RADIUS server: Its functions are the same among the three types of Web authentication.

Authentication process:

- In FS Second-Generation Web Authentication, the authentication and accounting functions are transferred from the portal server to the NAS.
- Because authentication proceeds on the NAS, the second-generation NAS does not need to wait for the authentication results notified by the portal server as the first generation.

Logout process:

- In FS First-Generation Web Authentication, a logout action may be triggered by a notification from the portal server, or traffic detection or port status detection performed by the NAS. In FS Second-Generation Web Authentication, a logout action may be triggered by a notification from the portal server, a kickout notification from the RADIUS server, or traffic detection or port status detection performed by the NAS.
- In FS First-Generation Web Authentication, Accounting Stop packets are sent by the portal server. In FS Second-Generation Web Authentication, Accounting Stop packets are sent by the NAS.

The selection of the Web authentication versions depends on the type of the portal server in use.

Command parameters in this document may be shared by the three Web authentication versions or not. Read through this document carefully to avoid parameter misconfiguration that will affect Web authentication.

Overview

Feature	Description
FS First-Generation Web Authentication	The portal server is deployed and supports only FS First-Generation Web Authentication.
FS Second-Generation Web Authentication	The portal server is deployed and complies with the *CMCC WLAN Service Portal Specification*.

5.3.1 FS First-Generation Web Authentication

HTTP Interception

HTTP interception means the NAS intercepts to-be-forwarded HTTP packets. Such HTTP packets are initiated by the browsers of the clients connected to the NAS, but they are not destined for the NAS. For example, when a client attempts to visit the website www.google.com using the Internet Explorer, the NAS is expected to forward the HTTP request packets to the gateway. If HTTP interception is enabled, these packets will not be forwarded.

After HTTP interception is successful, the NAS redirects the HTTP requests from the client to itself to establish a session between them. Then, the NAS pushes a Webpage to the client through HTTP redirection, which can be used for authentication, software downloading or other purposes.

You can specify the clients and destination interfaces to enable or disable HTTP interception for Web authentication. In general, HTTP requests from unauthenticated clients will be intercepted, and those from authenticated clients will not. HTTP interception is the foundation of Web authentication. Web authentication is automatically triggered once HTTP interception succeeds.

HTTP Redirection

According to HTTP protocols, after the NAS receives a HTTP GET or HEAD request packet from a client, a packet with 200 (Ok) status code is replied if it is able to provide the required resources, or a packet with 302 (Moved Temporarily) status code is returned if unable. Another URL is provided in the 302 packet. After receiving the packet, the client may resend a HTTP GET or HEAD request packet to the new URL for requesting resources. This process is called redirection.

HTTP redirection is an important procedure following HTTP interception in Web authentication. It takes the advantage of 302 status code defined in HTTP protocols. HTTP interception creates a session between the NAS and a client. The client sends HTTP GET or HEAD request packets (which should have been sent to another site) to the NAS. The NAS responds with a 302 packet with a specific redirection page. Thereby, the client resends the requests to the redirection page.

Because more and more application programs run HTTP protocols, the use of the 302 redirection packet may divert a large amount of HTTP traffic (not sent by browsers) to the portal server, which will affect network authentication. To address this problem, HTTP redirection technology on the NAS adopts noise reduction to replace the 302 packets with the **js** script.

Working Principle

Figure 5- 1 shows the networking topology of Web authentication.

First-generation Webauth roles:

- Authentication client: Is usually a browser running HTTP protocols. It sends HTTP requests for accessing the Internet.
- NAS: Is an access-layer device in a network. The NAS is directly connected to clients and must be enabled with Web authentication.

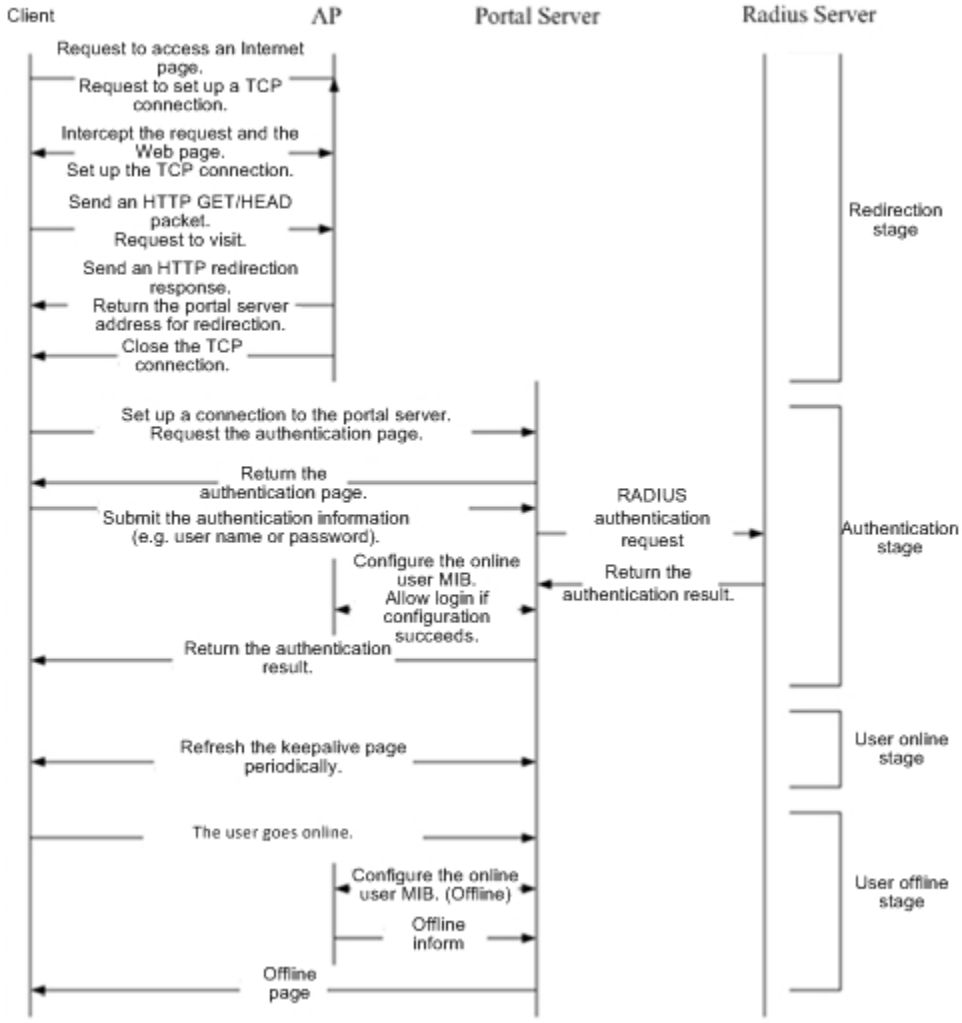
- Portal server: Provides a Web page for Web authentication and related operations. After receiving an HTTP authentication request from a client, the portal server extracts account information from the request, sends the information to the RADIUS server for authentication, and notifies the client and NAS of the authentication result. Figure 5- 1 shows FS ePortal server.
- RADIUS server: Provides the RADIUS-based authentication service to remote clients. The portal server extracts users' authentication account information from HTTP packets and initiates authentication requests to the RADIUS server through the RADIUS protocol. The RADIUS server returns the authentication result to the portal server through the RADIUS protocol.

First-generation Webauth process:

1. Before authentication, the NAS intercepts all HTTP requests from a client and redirects these requests to the iPortal server. Thereafter, an authentication page is displayed on the browser.
2. During authentication, the client enters information, for example, username, password, and verification code, on the Webauth URL to interact with the portal server and complete authentication.
3. After the user is authenticated, the portal server notifies the NAS that the client has passed authentication, and the NAS allows the client to access resources on the Internet.

Figure 5- 2 shows the flowchart of FS First-Generation Web Authentication by using an AP as the NAS.

Figure 5- 2 Flowchart of FS First-Generation Web Authentication



First-generation client logout process:

There are two scenarios of client logout. One scenario is detected by the NAS that a client gets offline for the maximum online time is out, the upper traffic limit is reached, or the link is disconnected. The other scenario is detected by the portal server that a client logs out by clicking the **Logout** button on the logout page or the keep-alive page is invalid.

1. Scenario 1: The NAS detects a client to logout and informs the portal server. Then the portal server deletes the user information on the NAS through SNMP and displays a logout page to the client.
2. Scenario 2: The portal server detects a client to logout and informs the NAS through SNMP and displays a logout page to the client.
3. In the two scenarios, the portal server sends an Accounting Stop request to the RADIUS server and notifies the RADIUS server that the client has logged out.

Related Configuration

Configuring the First-Generation Webauth Template

By default, the first-generation Webauth template is not configured.

Run the **web-auth template eportalv1** command in global configuration mode to create the first-generation Webauth template.

The template is used to implement Web authentication.

Configuring the IP Address of the Portal Server

By default, the IP address of the portal server is not configured.

Run the **ip** {*ip-address* } command in template configuration mode to configure the IP address of the portal server.

Any request packets to access the portal server will be filtered and rate-limited by the NAS.

Configuring the Webauth URL of the Portal Server

By default, the Webauth URL of the portal server is not configured.

Run the **url** {*url-string* } command in template configuration mode to configure the Webauth URL of the portal server.

The URL to which clients are redirected is the address of the Webauth URL provided by the portal server.

Specifying the Webauth Binding Mode

Run the **bindmode** command in template configuration mode to specify the Webauth binding mode.

In Web authentication on layer-3 networks, the source MAC address in a packet is changed after the packet is routed. In such case, configure the IP-only binding mode.

Configuring the Webauth Communication Key

By default, the Webauth communication key is not configured.

Run the **web-auth portal key** {*string* } command in global configuration mode to configure the Webauth communication key.

The communication key is used to encrypt URL parameters to avoid information disclosure.

Enabling FS First-Generation Web Authentication

By default, FS First-Generation Web Authentication is disabled.

Run the **web-auth enable** command in interface configuration mode to enable FS First-Generation Web Authentication on the client-connected ports.

After Web authentication is enabled, the unauthenticated clients connecting to a port will be redirected to the Webauth URL.

Configuring the SNMP-Server Host

By default, the SNMP-server host and community string are not configured.

Run the **snmp-server host** {*ip-address* }**version 2c** {*community-string* }**web-auth** command in global configuration mode to configure the SNMP-server host and community string for Web authentication.

The SNMP-server host is configured to receive Inform/Trap packets of user logout.

Configuring the SNMP-Server Community String

By default, the SNMP-server community string is not configured.

Run the **snmp-server community** {*community-string*} **rw** command in global configuration mode to configure the SNMP-server community string.

The SNMP-server community string is configured to read/write user information from/to the NAS.

Enabling the SNMP Trap/Inform Function

By default, the SNMP Trap/Inform function is disabled.

Run the **snmp-server enable traps web-auth** command in global configuration mode to enable the SNMP Trap/Inform function.

The SNMP Trap/Inform function is configured to enable the NAS to inform the portal server of user logout.

5.3.2 FS Second-Generation Web Authentication

HTTP Interception

Same as the HTTP interception technology of FS First-Generation Web Authentication.

HTTP Redirection

Same as the HTTP redirection technology of FS First-Generation Web Authentication.

Working Principle

Figure 5- 1 shows the networking topology of Web authentication.

Second-generation Webauth roles:

1. Authentication client: Is usually a browser running HTTP protocols. It sends HTTP requests for accessing the Internet.
2. NAS: Is an access-layer device in a network. The NAS is directly connected to clients and must be enabled with Web authentication. The NAS receives user authentication information from the portal server, sends authentication requests to the RADIUS server, determines whether users can access the Internet according to authentication results, and returns the authentication results to the portal server.
3. Portal server: Provides a Web page for Web authentication and related operations. After receiving an HTTP authentication request from a client, the portal server extracts account information from the request, transfers the information to the NAS, and displays the authentication result returned by the NAS to the user on a page. Figure 5- 1 shows FS ePortal server.
4. RADIUS server: Provides the RADIUS-based authentication service to remote clients. Figure 5- 1 shows the RADIUS server installed with the FS-SAM program.

Second-generation Webauth process:

1. Before authentication, the NAS intercepts all HTTP requests from a client and redirects these requests to the iPortal server.

Thereafter, an authentication page is displayed on the browser.

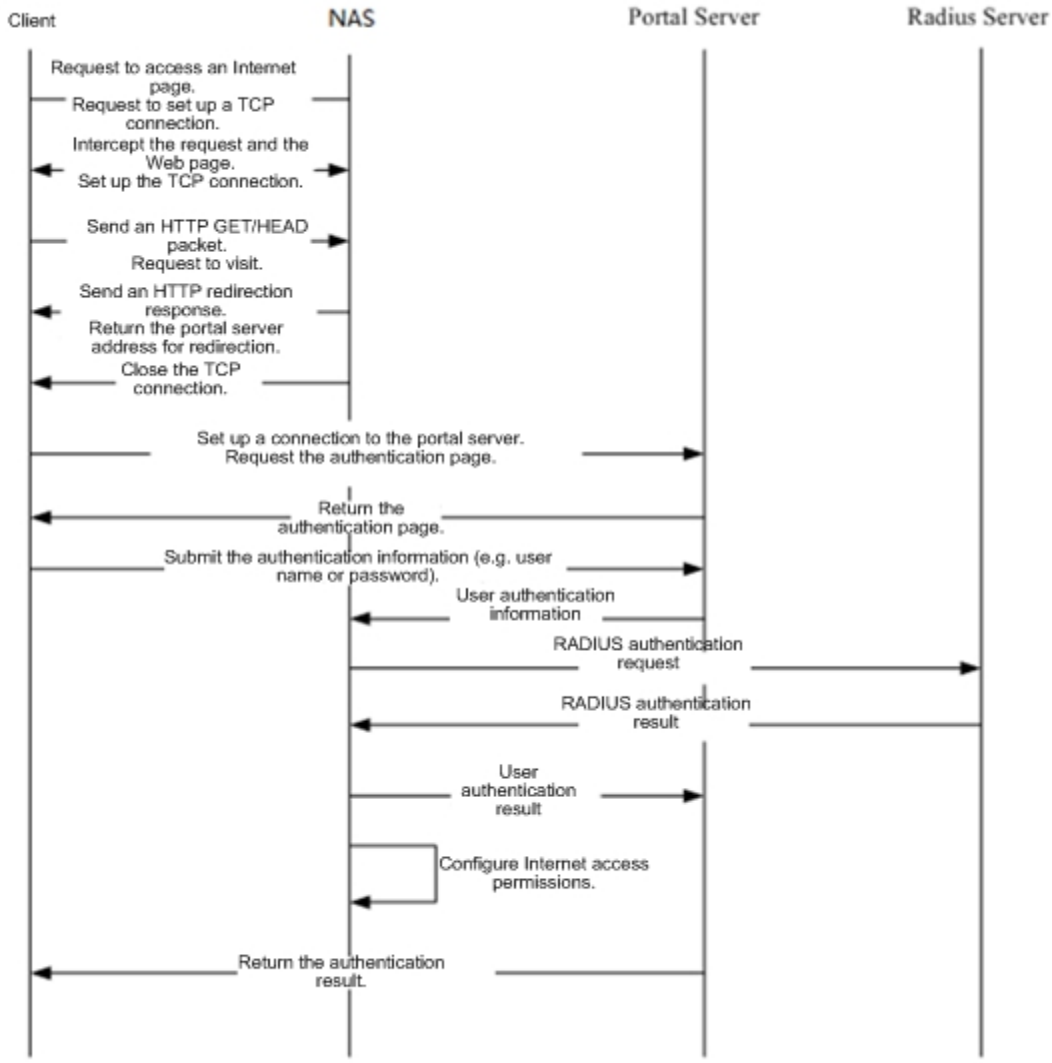
2. During authentication, the client enters information, for example, username, password, and verification code, on the Webauth URL to interact with the portal server.

3. The portal server sends the user authentication information to the NAS.

4. The NAS initiates authentication to the RADIUS server and returns the authentication result to the portal server.

5. The portal server displays the authentication result (success or failure) to the user on a page.

Figure 5- 3 Flowchart of FS Second-Generation Web Authentication



Second-generation client logout process:

There are two scenarios of client logout. One scenario is detected by the NAS that a client gets offline for the maximum online time is out, the upper traffic limit is reached, or the link is disconnected. The other scenario is detected by the portal server that a client logs out by clicking the **Logout** button on the logout page or the keep-alive page is invalid.

1. When a user clicks the **Logout** button on the online page, the portal server notifies the NAS to get the user offline.

2. The NAS gets a client offline with traffic lower than the threshold based on the parameters of user online traffic detection.

3. When the RADIUS server plans to force a client offline based on a certain policy, the NAS notifies the portal server to push a logout

page to the client.

Related Configuration

Configuring the Second-Generation Webauth Template

By default, the second-generation Webauth template is not configured.

Run the **web-auth template{eportalv2 |** *template-name* **v2}** command in global configuration mode to create a second-generation Webauth template.

The template is used to implement Web authentication.

Configuring the IP Address of the Portal Server

By default, the IP address of the portal server is not configured.

Run the **ip** { *ip-address* } command in template configuration mode to configure the IP address of the portal server.

Any request packets to access the portal server will be filtered and rate-limited by the NAS.

Configuring the Webauth URL of the Portal Server

By default, the Webauth URL of the portal server is not configured.

Run the **url** {*url-string* } command in template configuration mode to configure the Webauth URL of the portal server.

The URL to which clients are redirected is the address of the Webauth URL provided by the portal server.

Specifying the Webauth Binding Mode

Run the **bindmode** command in template configuration mode to specify the Webauth binding mode.

In Web authentication on layer-3 networks, the source MAC address in a packet is changed after the packet is routed. In such case, configure the IP-only binding mode.

Configuring the Webauth Communication Key

By default, the Webauth communication key is not configured.

Run the **web-auth portal key** { *string* } command in global configuration mode to configure the Webauth communication key.

The communication key is used to encrypt URL parameters to avoid information disclosure.

Enabling FS Second-Generation Web Authentication

By default, FS Second-Generation Web Authentication is disabled.

Run the **web-auth enable {eportalv2 |** *template-name* **v2}** command in interface configuration mode to enable FS Second-Generation Web Authentication on the client-connected ports.

After Web authentication is enabled, the unauthenticated clients connecting to a port will be redirected to the Webauth URL.

Enabling AAA

By default, AAA is disabled.

Run the **aaa new-model** command in global configuration mode to enable AAA.

FS Second-Generation Web Authentication relies on AAA. Enable AAA before you implement the former.

Configuring the RADIUS-Server Host and Communication Key

By default, the RADIUS-server host and communication key are not configured.

Run the **radius-server host** command in global configuration mode to configure the RADIUS-server host and communication key.

The RADIUS-server host is responsible for authenticating users.

Configuring an AAA Method List for FS Second-Generation Web Authentication

By default, no AAA method list is configured for FS Second-Generation Web Authentication.

Run the **aaa authentication web-auth** command in global configuration mode to configure an AAA method list for FS Second-Generation Web Authentication.

The AAA authentication method list is used for interaction during the Webauth process.

Configuring an AAA Method List for FS Second-Generation Web Accounting

By default, no AAA method list is configured for FS Second-Generation Web Accounting.

Run the **aaa accounting network** command in global configuration mode to configure an AAA method list for FS Second-Generation Web Accounting.

The AAA method list for Web accounting is used for accounting interaction during the Webauth process.

Specifying an AAA Method List

The default AAA method list is used if no list is specified.

Run the **authentication** command in template configuration mode to specify an AAA method list.

The AAA method list is specified to send authentication requests to AAA.

Specifying an AAA Accounting Method List

The default AAA accounting method list is used if no list is specified.

Run the **accounting** command in template configuration mode to specify an AAA accounting method list.

The AAA accounting method list is specified to send accounting requests to AAA.

Specifying the UDP Port of the Portal Server

By default, UDP Port 50100 is used.

Run the **port** command in template configuration mode to specify the UDP port of the portal server.

The UDP port is specified for the portal server to communicate with the NAS.

5.4 Configuration

Configuration	Description and Command	
Configuring FS First-Generation Web Authentication	(Mandatory) It is used to set the basic parameters of FS First-Generation Web Authentication.	
	web-auth template eportalv1	Configures the first-generation Webauth template.
	ip {*ip-address*}	Configures the IP address of the portal server.
	url { *url-string* }	Configures the Webauth URL of the portal server.

Configuration	Description and Command	
	web-auth portal key { *key-string* }	Configures the Webauth communication key.
	snmp-server community { *community-string* } **rw**	Configures the SNMP-server community string.
	snmp-server host { *ip-address* } **inform version** 2c { *community-string* } **web-auth**	Configures the SNMP-server host.
	snmp-server enable traps web-auth	Enables the SNMP-server Trap/Inform function.
	web-auth enable	Enables FS First-Generation Web Authentication on an interface.
Configuring FS Second-Generation Web Authentication	(Mandatory) It is used to set the basic parameters of FS Second-Generation Web Authentication.	
	aaa new-model	Enables AAA.
	radius-server host {*ip-address*}[**auth-port** *port-number*] [**acct-port** *port-number*] **key** {*string*}	Configures the RADIUS-server host and communication key.
	aaa authentication web-auth { **default** | *list-name* } *method1* [*method2*...]	Configures an AAA method list for Web authentication. (RADIUS authentication is implemented.)
	aaa accounting network { **default** | *list-name* } **start-stop** *method1* [*method2*...]	Configures an AAA method list for Web Accounting. (RADIUS accounting is implemented.)
	web-auth template{**eportalv2** |*portal-name***v2**}	Configures a second-generation Webauth template.
	ip {*ip-address* }	Configures the IP address of the portal server.
	url { *url-string* }	Configures the Webauth URL of the portal server.
	web-auth portal key { *key-string* }	Configures the Webauth communication key.
	web-auth enable	Enables FS Second-Generation Web Authentication on an interface.
Specifying an Authentication Method List	(Optional) It is used to specify an AAA authentication method list in template configuration mode. The name of the method list must be correctly specified.	
	authentication { *mlist-name* }	Specifies an AAA authentication method list(only for FS Second-Generation Web Authentication)
Specifying an Accounting Method List	(Optional) It is used to specify an AAA accounting method in template configuration mode. The name of the method list must be correctly specified.	
	accounting { *mlist-name* }	Specifies an AAA accounting method list(only for FS Second-Generation Web Authentication)
Configuring the Communication Port of the Portal Server	(Optional) It is used to specify the UDP port of the portal server in template configuration mode. The configured port number must be consistent with that on the RADIUS server.	
	port { *port-num* }	Configures the communication port of the portal server.
Specifying the Webauth Binding Mode	(Optional) It is used to specify the entry binding mode in template configuration mode.	
	bindmode {ip-mac-mode | ip-only-mode}	Specifies the template binding mode.

Configuration	Description and Command	
Configuring the Format of the Webauth URL	(Optional) It is used to configure the redirection URL format for a template.	
	fmt custom encry { [**none | md5 | des | des_ecb | des_ecb3**] } { [**user-ip** *userip-str*] [**user-mac** *usermac-str* **mac-format [dot | line | none]**] [**user-vid** *uservid-str*] [**user-id userid-str**] [**nas-ip** *nasip-str*] [**nas-id** *nasid-str*][**nas-id2***nasid2-str*][**ac-name** *acname-str*] [**ap-mac** *apmac-str* **mac-format [dot | line | none]**] [**url** *url-str*] [**ssid** *ssid-str*] [**port** *port-str*] [**ac-serialno** *ac-sno-str*] [**ap-serialno** *ap-sno-str*] [**additional** *extern-str*] }	Configures the format of the Webauth URL.
Configuring the Redirection HTTP Port	(Optional) It is used to configure the TCP interception port for redirection, so that the packets on the specified port can be redirected when interception is enabled.	
	http redirect port { *port-num* }	Configures the redirection TCP port.
Configuring Rate Limit Webauth Logging	(Optional) It is used to configure the syslog function in Web authentication.	
	web-auth logging enable {*num* }	Configures the rate limit Webauth logging.
Configuring the Maximum Number of HTTP Sessions for Unauthenticated Clients	(Optional) It is used to adjust the HTTP session limit. The limit value needs to be increased when there are many sessions in the background.	
	http redirect session-limit { *session-num* } [**port** { *port-session-num* }]	Configures the maximum number of HTTP sessions for unauthenticated clients.
Configuring the HTTP Redirection Timeout	(Optional) It is used to modify the timeout period for redirection connections. The timeout needs to be increased to complete redirection when the network condition is bad.	
	http redirect timeout{ *seconds* }	Configures the HTTP redirection timeout.
Configuring the Straight-Through ARP Resource Range	(Optional) It is used to permit the ARP of the specified addresses to pass. The gateway ARP must be permitted to pass when ARP check is enabled.	
	http redirect direct-arp { *ip-address* [*ip-mask*] }	Configures the straight-through ARP resource.
Configuring an Authentication-Exempted Address Range	(Optional) It is used to exempt clients from authentication when accessing the Internet.	
	web-auth direct-host { *ip-address* [*ip-mask*] [**arp**] } [**port** *interface-name*]	Configures the range of the IP or MAC addresses of clients free from authentication.
Configuring the Interval for Updating Online User Information	(Optional) It is used to configure the interval for updating online user information.	
	web-auth update-interval { *seconds* }	Configures the interval for updating online user information.
Configuring Portal Detection	(Optional) It is used to detect the availability of the portal server. If it is not available, the services are switched to the standby portal server. This function must be used together with portal standby function.	

Configuration	Description and Command	
	web-auth portal-check [**interval** *intsec* [**timeout** *tosec]* [**retransmit** retries]	Configures the portal server detection interval, timeout period, and timeout retransmission times.
Configuring Portal Escape	(Optional) It is used to allow new clients to access the Internet without authentication when the portal server is not available.	
	web-auth portal-escape	Configures portal escape.
Enabling DHCP Address Check	(Optional) It is used to check whether the IP address of a client is allocated by the DHCP server. If not, the client's authentication request is denied.	
	web-auth dhcp-check	Checks whether the IP address of a client is assigned by the DHCP server.
Configuring a Whitelist and Blacklist	(Optional) It is used to configure a blacklist to prevent authenticated clients from accessing some network resources, and a whitelist to allow unauthenticated clients to access some network resources.	
	web-auth acl{**black-ip** *ip*|**black-port** *port*|**black-url** *name*|**white-url** *name*}	Configures a whitelist and blacklist.
Configuring the Portal Communication Port	(Optional) It is used to configure the port (source port) used for the communication between the NAS and portal server.	
	ip portal source-interface *interface-type interface-num*	Specifies the port used for the communication between the NAS and portal server.
Configuring VLAN-Based Authentication on a Port	(Optional) It is used to configure the VLAN in which only the STAs inside the configured VLAN cannot initiate Web authentication.	
	web-auth vlan-control *vlan-list*	Configures the VLAN-based authentication on a port.
Disabling ARP Detection	(Optional) It is used to disable ARP detection. After the function is disabled, it is not detected that the IP address or MAC address of a client migrates.	
	no web-auth station-move arp-detect	Disables ARP detection.

5.4.1 Configuring FS First-Generation Web Authentication

Configuration Effect

Redirect unauthenticated clients to the Webauth URL to perform authentication.

Notes

N/A

Configuration Steps

Configuring the Portal Server

- (Mandatory)To enable Web authentication successfully, you must configure and apply the portal server.

- When the NAS or convergence device finds an unauthenticated client attempting to access network resources through HTTP, it redirects the access request to the specified Webauth URL, where the client can initiate authentication to the portal server. If the IP address of the portal server is configured as a free network resource, unauthenticated clients can directly visit this IP address through HTTP.

Configuring the Communication Key Between the NAS and Portal Server

- (Mandatory) To enable Web authentication successfully, you must configure the key used for the communication between the NAS or convergence device and portal server.
- When the NAS finds an unauthenticated client attempting to access network resources, it redirects the client to the specified Webauth URL, where the client can initiate authentication to the portal server. During the authentication process, the communication key is used to encrypt some data exchanged between the NAS and portal server to improve security.

Setting the SNMP Parameters Between the NAS and Portal Server

- (Mandatory) To enable Web authentication successfully, you must set the SNMP network management parameters used for the communication between the NAS and portal server.
- The NAS or convergence device and portal server jointly manage authenticated clients through SNMP/MIB. A table of authenticated clients is managed by MIB on the NAS. The portal server is able to access the MIB to obtain client statistics so as to control client login and logout. When a client logs out, the NAS or convergence device will inform the portal server by Webauth Inform packets.

Enabling FS First-Generation Web Authentication on an Interface

- Mandatory.
- When FS First-Generation Web Authentication is enabled in interface configuration mode, Web authentication is not enabled on any port by default. The users connecting to the port do not need to perform Web authentication.

Verification

- Check whether unauthenticated clients are required to perform authentication.
- Check whether authenticated clients can access the Internet normally.

Related Commands

Configuring the First-Generation Webauth Template

Command	**web-auth template eportalv1**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	**eportalv1** is the default template of FS First-Generation Web Authentication.

Configuring the IP Address of the Portal Server

Command	**ip** {*ip-address*}
Parameter Description	Indicates the IP address of the portal server.

Command Mode	Webauth template configuration mode
Usage Guide	N/A

Configuring the Webauth URL of the Portal Server

Command	**url** {*url-string*}
Parameter Description	*url-string*: Indicates the Webauth URL of the portal server.
Command Mode	Webauth template configuration mode
Usage Guide	The URL starts with **http://** or **https://**.

Specifying the Webauth Binding Mode

Command	**bindmode** { **ip-mac-mode | ip-only-mode** }
Parameter Description	Indicates the Webauth binding mode.
Command Mode	Webauth template configuration mode
Usage Guide	N/A

Configuring the Webauth Communication Key

Command	**web-auth portal key** {*key-string*}
Parameter Description	*key-string:* Indicates the Webauth communication key used for the communication between the NAS and portal server. The key contains up to 255 characters.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the SNMP-Server Community String

Command	**snmp-server community** {*community-string*}**rw**
Parameter Description	*community-string*: Indicates the community string. **rw**: Must be set to **rw** to support the read and write operations as the Set operation on MIB is required.
Command Mode	Global configuration mode
Usage Guide	The SNMP-server community string is used by the portal server to manage the online clients on the NAS or convergence device.

Configuring the SNMP-Server Host

Command	**snmp-server host** {*ip-address*} **inform version 2c** {*community-string*} **web-auth**
Parameter Description	*ip-address*: Indicates the IP address of the SNMP-server host, that is, the portal server. *community-string*: Configures the community string used to send an SNMP Inform message.
Command Mode	Global configuration mode

Usage Guide	Configure the SNMP-server host to receive Webauth messages, including the type, version, community string, and other parameters. **inform:** Enables the SNMP Inform function. The NAS or convergence device will send a message to the portal server when a client logs out. The message type is set to Inform instead of Trap to avoid message loss. **version 2c**: Indicates SNMPv2 for SNMP Inform is not supported in all SNMP versions excluding SNMPv1. **web-auth**: Indicates the preceding parameters to be used for Web authentication. For details regarding SNMP configuration and others, see the *Configuring SNMP*. The SNMP parameter **version 2c**listed here is aimed at SNMPv2. SNMPv3 is recommended if higher security is required for the SNMP communication between the NAS and portal server. To use SNMPv3, change **SNMP Community** to **SNMP User**, **version 2c** to **SNMPv3**, and set SNMPv3-related security parameters. For details, see the *Configuring SNMP*.

Enabling the Webauth Trap/Inform Function

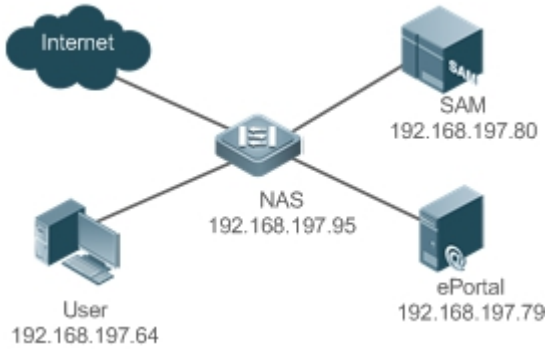
Command	**snmp-server enable traps web-auth**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure the NAS or convergence device to send Webauth Trap and Inform messages externally. web-auth: Indicates Web authentication messages.

Enabling FS First-Generation Web Authentication on an Interface

Command	**web-auth enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring FS First-Generation Web Authentication

Scenario **Figure 5-4**	
Configuration	● On the NAS, configure the IP address of the ePortal server and the key (FS) used for communicating with the

Steps	ePortal server. ● Configure the Webauth URL on the NAS. ● Set the SNMP network management parameters (community string: public) used for the communication between the NAS and ePortal server. ● Enable Web authentication on ports GigabitEthernet 0/2 and GigabitEthernet 0/3 on the NAS.
	FS# config Enter configuration commands, one per line. End with CNTL/Z. FS(config)#web-auth template eportalv1 FS(config.tmplt.eportalv1)#ip 192.168.197.79 FS(config.tmplt.eportalv1)#exit FS(config)# web-auth portal key FS
	FS(config)# web-auth template eportalv1 FS(config.tmplt.eportalv1)#url http://192.168.197.79:8080/eportal/index.jsp FS(config.tmplt.eportalv1)#exit
	FS(config)# snmp-server community public rw FS(config)# snmp-server enable traps web-auth FS(config)# snmp-server host 192.168.197.79 inform version 2c public web-auth FS(config)# exit
	FS(config)# interface range GigabitEthernet 0/2-3 FS(config-if-range)# web-auth enable FS(config-if-range)# exit
Verification	● Check whether Web authentication is configured successfully.

	FS(config)#show running-config
	…
	snmp-server host 192.168.197.79 inform version 2c public web-auth
	snmp-server enable traps web-auth
	snmp-server community public rw
	…
	web-auth template eportalv1
	  ip 192.168.197.79
	  url http://192.168.197.79:8080/eportal/index.jsp
	!
	web-auth portal key FS

	…
	interface GigabitEthernet 0/2
	  web-auth enable
	!
	interface GigabitEthernet 0/3
	  web-auth enable

	FS#show web-auth control
	  Port Control Server Name Online User Count
	  ---------- -------- ---------- ----------
	…
	GigabitEthernet 0/2On eportalv1 0
	GigabitEthernet 0/3On eportalv1 0
	…

	FS#show web-auth template
	Webauth Template Settings:
	---------- 	
	  Name: eportalv1
	  Url: http://17.17.1.21:8080/eportal/index.jsp
	  Ip: 17.17.1.21
	  BindMode: ip-mac-mode
	  Type: v1
	……

Common Errors

- The SNMP parameters used for the communication between the portal server and NAS are configured incorrectly, causing authentication failures.
- Specify the IP-MAC binding mode to deploy Web authentication on layer-3 networks, causing authentication failures.
- When Web authentication is used in conjunction with VRRP, run the snmp-server trap-source ip command to specify the VRRP address; otherwise, the portal server cannot process Trap packets correctly.

5.4.2 Configuring FS Second-Generation Web Authentication

Configuration Effect

Redirect unauthenticated clients to the Webauth URL to perform authentication. IPv6 is supported.

Notes

- FS Second-Generation Web Authentication complies with the CMCC WLAN Service Portal Specification. Furthermore, it is extended to support FS portal server. Perform compatible configuration based on the server performance in actual deployment. For details, see the subsequent chapter.
- The cmcc-normal and cmcc-ext1 parameters in the fmt command support only IPv4. If IPv6 is used, the configuration of the portal server is invalid.

Configuration Steps

Enabling AAA

- (Mandatory) To enable FS Second-Generation Web Authentication, you must enable AAA.
- The NAS is responsible for initiating authentication to the portal server through AAA in FS Second-Generation Web Authentication.

Configuring the RADIUS-Server Host and Communication Key

- (Mandatory) To enable FS Second-Generation Web Authentication, you must configure the RADIUS server.
- Clients' account information is stored on the RADIUS server. The NAS needs to connect to the RADIUS server to validate a client.

Configuring an AAA Method List for Web Authentication

- (Mandatory) To enable FS Second-Generation Web Authentication, you must configure an AAA authentication method list.
- An AAA authentication method list associates Web authentication requests with the RADIUS server. The NAS selects an authentication method and server based on the method list.

Configuring an AAA Method List for Web Accounting

- (Mandatory) To enable FS Second-Generation Web Authentication, you must configure an AAA method list for Web accounting.
- An accounting method list is used to associate an accounting method and server. In Web authentication, accounting is implemented to record client fees.

Configuring the Portal Server

- (Mandatory) To enable FS Second-Generation Web Authentication, you must configure and apply the portal server.
- When the NAS or convergence device finds an unauthenticated client attempting to access network resources through HTTP, it redirects the access request to the specified Webauth URL, where the client can initiate authentication to the portal server. If the IP address of the portal server is configured as a free network resource, unauthenticated clients can directly visit this IP address through HTTP.

Configuring the Communication Key Between the NAS and Portal Server

- (Mandatory) To enable FS Second-Generation Web Authentication, you must configure the key used for the communication between the NAS or convergence device and portal server.
- When the NAS finds an unauthenticated client attempting to access network resources, it redirects the client to the specified Webauth URL, where the client can initiate authentication to the portal server. During the authentication process, the communication key is used to encrypt some data exchanged between the NAS and portal server to improve security.

Configuring the Portal Server in Global or Interface Configuration Mode

- (Mandatory) To enable FS Second-Generation Web Authentication, you must specify the use of the second generation portal server in global or interface configuration mode.

- The NAS first selects the portal server in interface configuration mode. If such a portal server does not exist, the NAS selects the portal server in global configuration mode. If such a portal server does not exist, eportalv1 is used by default. The NAS redirects users to the selected portal server.

Enabling FS Second-Generation Web Authentication on an Interface

- Mandatory.
- When FS Second-Generation Web Authentication is enabled in interface configuration mode, Web authentication is not enabled on any port by default. The users connecting to the port do not need to perform Web authentication.

Verification

- Check whether unauthenticated clients are required to perform authentication.
- Check whether authenticated clients can access the Internet normally.

Related Commands

Enabling AAA

Command	**aaa new-model**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	You can configure the AAA authentication and accounting method lists only after AAA is enabled.

Configuring the RADIUS-Server Host and Communication Key

Command	**radius-server host** {*ip-address*} [**auth-port***port-number1*] [**acct-port** *port-number 2*] **key** {*string*}
Parameter Description	*ip-address*: Indicates the IP address of the RADIUS server host. *port-number1:* Indicates the authentication port. *port-number2:* Indicates the accounting port. *string*: Indicates the key string.
Command Mode	Global configuration mode
Usage Guide	By default, the authentication port number is 1812, and the accounting port number is 1813.

Configuring an AAA Method List for Web Authentication

Command	**aaa authentication web-auth** { **default** | *list-name* } *method1* [*method2*...]
Parameter Description	*list-name*: Creates a method list. *method1*: Configures method 1. *method2*: Configures method 2.
Command Mode	Global configuration mode
Usage Guide	FS Second-Generation Web Authentication adopts the RADIUS authentication method.

Configuring an AAA Method List for Web Accounting

Command	**aaa accounting network** { **default** | *list-name* } **start-stop** *method1* [*method2*...]
Parameter Description	*list-name*: Creates a method list. *method1*: Configures method 1. *method2*: Configures method 2.
Command Mode	Global configuration mode
Usage Guide	FS Second-Generation Web Authentication adopts the RADIUS accounting method.

Configuring the Second-Generation Webauth Template

Command	**web-auth template**{**eportalv2 |** *portal-name* **v2**}
Parameter Description	*portal-name*: Indicates the customized portal server name.
Command Mode	Global configuration mode
Usage Guide	**eportalv2** indicates the default template of FS Second-Generation Web Authentication.

Configuring the IP Address of the Portal Server

Command	**ip** { *ip-address* }
Parameter Description	Indicates the IP address of the portal server.
Command Mode	Webauth template configuration mode
Usage Guide	N/A

Configuring the Webauth URL of the Portal Server

Command	**url** { *url-string* }
Parameter Description	Indicates the Webauth URL of the portal server.
Command Mode	Webauth template configuration mode
Usage Guide	The URL starts with **http://** or **https://**.

Configuring the Format of the Webauth URL

Command	**fmt** {**cmcc-ext1 | cmcc-ext2 | cmcc-mtx | cmcc-normal | ct-jc | cucc| FS | custom**}
Parameter Description	Indicates the format of the Webauth URL.
Command Mode	Webauth template configuration mode
Usage Guide	The **cmcc-normal** and **cmcc-ext1** parameters in the **fmt** command support only IPv4. If IPv6 is used, the configuration of the portal server is invalid. The **cmcc-ext2** is supported for Liaoning CMCC. When **fmt** is set to **cmcc-mtx**, the URL format of mobile AC vendors is supported. The **ct-jc** format is supported for Chine Telecom.

	The **cucc** format is supported for Shandong China Telecom. The **custom** format is defined by users.

Specifing the Template Binding Mode

Command	**bindmode {ip-mac-mode | ip-only-mode}**
Parameter Description	Indicates the template binding mode.
Command Mode	Webauth template configuration mode
Usage Guide	N/A

Configuring the Webauth Communication Key

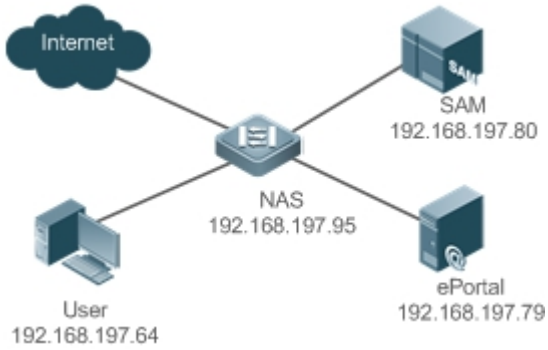
Command	**web-auth portal key** { *key-string* }
Parameter Description	*key-string*: Indicates the Webauth communication key used for the communication between the NAS and portal server. The key contains up to 255 characters.
Command Mode	Global configuration mode
Usage Guide	N/A

Enabling FS Second-Generation Web Authentication on an Interface

Command	**web-auth enable** {**eportalv2 |** *template-name*}
Parameter Description	Indicates a Webauth template.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring FS Second-Generation Web Authentication

Scenario **Figure 5-5**	
Configuration Steps	● Enable AAA on the NAS. ● Configure the RADIUS-server host and communication key on the NAS. ● Configure the default AAA method lists for Web authentication and accounting on the NAS. ● Configure the IP address of the portal server and the Webauth communication key (FS) used for communicating

	with the portal server on the NAS. ● Configure the Webauth URL on the NAS. ● Configure FS Second-Generation Web Authentication in global configuration mode on the NAS. ● Enable Web authentication on ports GigabitEthernet 0/2 and GigabitEthernet 0/3 on the NAS.
	FS#configure Enter configuration commands, one per line. End with CNTL/Z. FS(config)#aaa new-model
	FS(config)#radius-server host 192.168.197.79 key FS
	FS(config)#aaa authentication web-auth default group radius FS(config)#aaa accounting network default start-stop group radius
	FS(config)#web-auth template eportalv2 FS(config.tmplt.eportalv2)#ip 192.168.197.79 FS(config.tmplt.eportalv2)#exit FS(config)#web-auth portal key FS
	FS(config)# web-auth template eportalv2 FS(config.tmplt.eportalv2)#url http://192.168.197.79:8080/eportal/index.jsp FS(config.tmplt.eportalv2)#exit
	FS(config)# interface range GigabitEthernet 0/2-3 FS(config-if-range)# web-auth enable eportalv2 FS(config-if-range)# exit
Verification	● Check whether Web authentication is configured successfully.
	FS(config)#show running-config … aaa new-model aaa authentication web-auth default group radius aaa accounting network default start-stop group radius … radius-server host 192.168.197.79 key FS … web-auth template eportalv2 ip 192.168.197.79 url http://192.168.197.79:8080/eportal/index.jsp

	! web-auth portal key FS ! web-auth enable ! interface GigabitEthernet 0/2 interface GigabitEthernet 0/3
	FS#show web-auth control Port Control Server Name Online User Count ---------- -------- ---------- ---------- …Global On eportalv2 1 …
	FS#show web-auth template Webauth Template Settings: ---------- Name: eportalv2 Url: http://17.17.1.21:8080/eportal/index.jsp Ip: 17.17.1.21 BindMode: ip-mac-mode Type: v2 Port: 50100 State: Active Acctmlist: default Authmlist: default …

Common Errors

- The communication key between the portal server and NAS is configured incorrectly or only on the portal server or NAS, causing authentication errors.
- The communication parameters of the RADIUS server and NAS are set incorrectly, causing authentication errors.
- The portal server does not support the *CMCC WLAN Service Portal Specification*, causing compatibility failure.

5.4.3 Specifying an Authentication Method List

Configuration Effect

- The portal server sends an authentication request to the NAS when a user submits authentication information. The NAS resolves the authentication server information and other information based on the configured authentication method list name before initiating authentication.
- The NAS selects the authentication server based on the specified authentication method list.

Notes

- Before you configure an authentication method list name, ensure that the authentication methods in the list have been configured on the AAA module. The command used to configure authentication methods on the AAA module is **aaa authentication web-auth** { **default** | *list-name* }*method1 [method2*...].

Configuration Steps

- Optional.
- The default authentication method is used if no authentication method list is configured. Run the **authentication** { *mlist-name* } command to configure an authentication method list name when the authentication method list name on the AAA module needs to be modified or multiple method lists exist.

Verification

- Configure two authentication method lists on the AAA module. Apply list 1 to server 1 and list 2 to server 2.
- Create user a and configured a password for the user on server 1. Create user b on server 2.
- Configure the use of list 1.
- Perform authentication as user b and check that authentication fails.
- Perform authentication as user a and check that authentication is successful.

Related Commands

Specifying an Authentication Method List

Command	**authentication** {*mlist-name*}
Parameter Description	Indicates a method list name.
Command Mode	Webauth template configuration mode
Usage Guide	Ensure that the configured authentication method list name is consistent with that on the AAA module.

Configuration Example

Specifying an Authentication Method List

Configuration Steps	• Specify the authentication method list **mlist1**.
	FS(config.tmplt.iportal)#authentication mlist1
Verification	• Check whether the configuration is successful.
	FS#show web-auth template

Configuration Steps	● Specify the authentication method list **mlist1**.
	FS(config.tmplt.iportal)#authentication mlist1
Verification	● Check whether the configuration is successful.
	Webauth Template Settings: ---------- Name: eportalv2 Url: http://17.17.1.21:8080/eportal/index.jsp Ip: 17.17.1.21 BindMode: ip-only-mode Type: v2 Port: 50100 State: Active Acctmlist: default Authmlist: mlist1

5.4.4 Specifying an Accounting Method List

Configuration Effect

- The NAS sends an accounting request when a user passes authentication. The recipient of the request depends on the configuration of the accounting method list and is usually the portal server.
- Specify an accounting method list for the NAS to perform accounting.

Notes

- Ensure that the accounting method list has been configured on the AAA module. The command used to configure accounting methods on the AAA module is **aaa accounting network** {**default** | *list-name* }start-stop method1 [method2...].

Configuration Steps

- Optional.
- The default accounting method is used if no accounting method list is configured. Run the **accounting** {*mlist-name* } command to configure an accounting method list name when the accounting method list name on the AAA module needs to be modified or multiple method list names exist.

Verification

- Configure two accounting method lists on the AAA module. Apply list 1 to server 1 and list 2 to server 2.
- Configure the use of list 1.
- Use a valid account to perform authentication to access the Internet.
- View user accounting information on server1 and server2. Check that the user accounting information exists only on server1.

Related Commands

Specifying an Accounting Method List

Command	**accounting**{*mlist-name*}
Parameter Description	Indicates a method list name.
Command Mode	Webauth template configuration mode
Usage Guide	Ensure that the configured accounting method list name is consistent with that on the AAA module.

Configuration Example

Specifying an Accounting Method List

Configuration Steps	● Specify the accounting method list **mlist1**.
	FS(config.tmplt.eportalv2)#accounting mlist1
Verification	● Check whether the configuration is successful.
	FS#show web-auth template Webauth Template Settings: ---------- Name: eportalv2 Url: http://17.17.1.21:8080/eportal/index.jsp Ip: 17.17.1.21 BindMode: ip-mac-mode Type: v2 Port: 50100 State: Active Acctmlist: mlist1 Authmlist: mlist1

5.4.5 Configuring the Communication Port of the Portal Server

Configuration Effect

- When the NAS detects that a user logs out, it notifies the portal server. The NAS interacts with the portal server through the portal specification, which specifies the port number used to listen to and send/receive packets.
- When the listening port of the portal server is changed, the communication port of the portal server must be modified on the NAS to enable the NAS to interact with the portal server.
- If it is built-in Portal Web authentication, this function is used to configure the http port that the local machine listens on, the

default is 8081.

Notes

- The configured port number must be consistent with the port actually used by the portal server.
- Applicable to second-generation Web authentication and built-in Portal Web authentication, and the default port numbers of the two types of authentication are not the same. The second-generation Web authentication port number is used for the device and the Portal server to interact with the Portal protocol, and the port number for built-in Portal Web authentication is used for local message listening.

Configuration Steps

- Optional.
- Run the **port** *port-num* command to maintain port configuration consistency when the portal server does not use the default port number or the listening port of the NAS conflicts with other port and needs to be adjusted.

Verification

- Configure FS Second-Generation Web Authentication.
- Change the listening port of the server to 10000.
- Run the **port** *port-num* command to configure the port number 10000.
- Simulate the scenario where a user performs authentication to access the Internet.
- Force the user offline on the NAS, refresh the online page, and check that a user logout notification is displayed.

Related Commands

Configuring the Communication Port of the Portal Server

Command	**port** *port-num*
Parameter Description	*port-num*: Indicates the port number.
Command Mode	Webauth template configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Communication Port of the Portal Server

Configuration Steps	• Configure the communication port of the portal server as port 10000.
	FS(config.tmplt.eportalv2)#port 10000
Verification	• Check whether the configuration is successful.
	FS#show web-auth template Webauth Template Settings:

Configuration Steps	● Configure the communication port of the portal server as port 10000.
	FS(config.tmplt.eportalv2)#port 10000
Verification	● Check whether the configuration is successful.
	---------- Name: eportalv2 Url: http://17.17.1.21:8080/eportal/index.jsp Ip: 17.17.1.21 BindMode: ip-only-mode Type: v2 Port: 10000 Acctmlist: Authmlist:

5.4.6 Specifying the Webauth Binding Mode

Configuration Effect

- When a user goes online, the user's entry needs to be written to a forwarding rule. The forwarding rule mapping method can be modified by specifying different binding modes, which further affects the Internet access rules applied to users. In IP-only mode, all the packets carrying the specified IP address are permitted to pass, and the STAs who send the packets can access the Internet. In IP+MAC mode, only the packets carrying both the specified IP address and MAC address are permitted to pass, and the STAs who send the packets can access the Internet.

Notes

- In Layer-3 authentication, the MAC addresses visible to the NAS are the gateway addresses of STAs. Because these MAC addresses are not accurate, the IP-only mode should be used.

Configuration Steps

- (Optional) The default Webauth binding mode is IP+MAC.
- Determine a binding mode based on the accuracy of user information obtained by the NAS. When the IP and MAC addresses of STAs are accurate (in L2 authentication, for example), IP+MAC is recommended. When the IP and MAC addresses are not accurate, select IP-only.

Verification

- Change the binding mode to IP-only.
- Simulate the scenario where a user performs authentication to access the Internet.
- Modify the MAC address of the user, or use a client with the same IP address but a different MAC address to access the Internet.
- Check that the user accesses the Internet normally.

Related Commands

Specifying the Webauth Binding Mode

Command	**bindmode {ip-mac-mode | ip-only-mode}**
Parameter Description	**ip-mac-mode**: Indicates IP-MAC binding mode. **ip-only-mode**: Indicates IP-only binding mode.
Command Mode	Webauth template configuration mode
Usage Guide	N/A

Configuration Example

Specifying the Webauth Binding Mode

Configuration Steps	● Set the binding mode to IP-only.
	FS(config.tmplt.eportalv2)#bindmode ip-only-mode
Verification	● Check whether the configuration is successful.
	FS#show web-auth template Webauth Template Settings: ---------- Name: eportalv2 Url: http://17.17.1.21:8080/eportal/index.jsp Ip: 17.17.1.21 BindMode: ip-only-mode Type: v2 Port: 10000 Acctmlist: Authmlist:

5.4.7 Configuring the Redirection HTTP Port

Configuration Effect

- When an STA accesses network resources (for example, the user accesses the Internet using a browser), the STA sends HTTP packets. The NAS or convergence device intercepts these HTTP packets to determine whether the STA is accessing network resources. If the NAS or convergence device detects that the STA is not authenticated, it prevents the STA from accessing network resources and displays an authentication page to the STA. By default, the NAS intercepts the HTTP packets that STAs send to port 80 to determine whether STAs are accessing network resources.
- After a redirection HTTP port is configured, the HTTP requests that STAs send to the specified destination port can be redirected.

Notes

- The commonly used management ports on the NAS or convergence device, such as ports 22, 23 and 53, and ports reserved by the system are not allowed to be configured as the redirection port. All ports except port 80 with numbers smaller than 1000 are seldom used by the HTTP protocol. To avoid a conflict with the well-known TCP port, do not configure a port with a small number as the redirection port unless necessary.

Configuration Steps

- Optional.
- When you configure automatic client acquisition, if you need to enable the NAS to intercept the HTTP packets that STAs send to the specified destination port, configure a redirection HTTP port.

Verification

- Configure an interception port.
- Open the browser of a PC and access the Internet through the port without performing authentication.
- Check whether the access requests are redirected to an authentication page.

Related Commands

Configuring the Redirection HTTP Port

Command	**http redirect port** *port-num*
Parameter Description	*port-num*: Indicates the port number.
Command Mode	Global configuration mode
Usage Guide	A maximum of 10 different destination port numbers can be configured, not including default ports 80 and 443.

Configuration Example

Configuring the Redirection HTTP Port

Configuration Steps	● Configure port 8080 as the redirection HTTP port.
	FS(config)#http redirect port 8080
Verification	● Check whether the configuration is successful.
	FS(config)#show web-auth rdport Rd-Port: 80 443 8080

5.4.8 Configuring Rate Limit Webauth Logging

Configuration Effect

- The Web authentication module sends syslog messages to the administrator to display the information and relevant events of

users who perform login/logout. By default, syslog messages are shielded.

- After syslog output rate limiting is configured, syslog messages are sent at a certain rate.

Notes

- When the login/logout rate is high, syslog messages are output frequently, which affects device performance and results in spamming.

Configuration Steps

- Optional.
- Configure syslog output rate limiting when you need to view the syslog messages about user login/logout.

Verification

- Configure logging rate limiting.
- Check whether users log in and out at a certain rate.
- Check that syslog messages are printed out at the limit rate.

Related Commands

Configuring Rate Limit Webauth Logging

Command	**web-auth logging enable** *num*
Parameter Description	*num*: Indicates the syslog output rate (entry/second).
Command Mode	Global configuration mode
Usage Guide	When the syslog output rate is set to **0**, syslog messages are output without limit. The output of syslog messages of the critical level and syslog messages indicating errors is not limited.

Configuration Example

Configuring Rate Limit Webauth Logging

Configuration Steps	• Disable rate limit Webauth Logging.
	FS(config)#web-auth logging enable 0
Verification	• Check whether the configuration is successful.
	FS(config)#show running-config … web-auth logging enable 0 …

5.4.9 Configuring the Maximum Number of HTTP Sessions for Unauthenticated Clients

Configuration Effect

- When an unauthenticated user accesses network resources, the user's PC sends requests for HTTP session connection. The NAS or convergence device intercepts the HTTP packets and redirects the user to a Web authentication page. To prevent an unauthenticated user from initiating too many HTTP connection requests and save resources on the NAS, it is necessary to limit the maximum number of HTTP sessions that the unauthenticated user can initiate on the NAS.
- A user occupies an HTTP session when performing authentication, and the other application programs of the user may also occupy HTTP sessions. For this reason, it is recommended that the maximum number of HTTP sessions for an unauthenticated user be not set to 1. By default, each unauthenticated user can initiate 255 HTTP sessions globally, and each port supports up to 300 HTTP sessions initiated by unauthenticated clients.

Notes

- If the authentication page fails to be displayed during Web authentication, the maximum number of HTTP sessions may be reached. When this happens, the user can close the application programs that may occupy HTTP sessions and then perform Web authentication again.

Configuration Steps

- Optional.
- Perform this configuration when you need to change the maximum number of HTTP sessions that each unauthenticated user can initiate and the maximum number of HTTP sessions that unauthenticated clients can initiate on each port.
- Perform this configuration when you configure automatic SU client acquisition.

Verification

- Modify the maximum number of HTTP sessions that an unauthenticated user can initiate.
- Simulate the scenario where an unauthenticated user constructs identical sessions to connect to the NAS continuously.
- Simulate the scenario where the unauthenticated user accesses the Internet using a browser. Check whether the access requests are redirected and the NAS notifies the user that the maximum number of sessions is reached.

Related Commands

Configuring the Maximum Number of HTTP Sessions for Unauthenticated Clients

Command	**http redirect session-limit** { *session-num* }[**port** { *port-session-num* }]
Parameter Description	*session-num*: Indicates the maximum number of HTTP sessions for unauthenticated clients. The value range is 1 to 255. The default value is 255. *port-session-num*: Indicates the maximum number of HTTP sessions on each port for authenticated clients. The value range is 1 to 65,535. The default value is 300.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Maximum Number of HTTP Sessions for Unauthenticated Clients

Configuration Steps	● Set the maximum number of HTTP sessions for unauthenticated clients to 3.
	FS(config)#http redirect session-limit 3
Verification	● Check whether the configuration is successful.
	FS(config)#show web-auth parameter HTTP redirection setting: session-limit: 3 timeout: 3 FS(config)#

5.4.10 Configuring the HTTP Redirection Timeout

Configuration Effect

- Configure the HTTP redirection timeout to maintain redirection connections. When an unauthenticated user tries to access network resources through HTTP, the TCP connection requests sent by the user will be intercepted and re-established with the NAS or convergence device. Then, the NAS or convergence device waits for the HTTP GET/HEAD packets from the user and responds with HTTP redirection packets to close the connection. The redirection timeout is intended to prevent the user from occupying the TCP connection for a long time without sending GET/HEAD packets. By default, the timeout for maintaining a redirection connection is 3s.

Notes

N/A

Configuration Steps

- Optional.
- Perform this configuration to change the timeout for maintaining redirection connections.

Verification

- Change the timeout period.
- Use a network packet delivery tool to set up a TCP connection.
- View the status of the TCP connection on the NAS. Check whether the TCP connection is closed when the timeout is reached.

Related Commands

Configuring the HTTP Redirection Timeout

Command	**http redirect timeout** { *seconds* }
Parameter Description	*Seconds*: Indicates the timeout for maintaining redirection connections, in the unit of seconds. The value ranges from 1 to 10. The default value is 3s.
Command Mode	Global configuration mode

Usage Guide	N/A

Configuration Example

Configuring the HTTP Redirection Timeout

Configuration Steps	● Set the HTTP redirection timeout to 5s.
	FS(config)#http redirect timeout 5
Verification	● Check whether the configuration is successful.
	FS(config)#show web-auth parameter HTTP redirection setting: session-limit: 255 timeout: 5

5.4.11 Configuring the Straight-Through Network Resources

Configuration Effect

- After Web authentication or 802.1Xauthentication is enabled on a port, the users connecting to the port need to pass Web authentication or 802.1Xauthentication before accessing network resources.
- Perform this configuration to exempt users from authentication when accessing some network resources.
- If a website is configured as a network resource of authentication exemption, all users, including unauthenticated clients, can access the website. By default, authentication exemption is not configured, and unauthenticated clients are not allowed to access network resources.

Notes

- The maximum number of free resources and the maximum number of unauthenticated clients cannot exceed 1000 respectively. The actual number of available resources may be reduced because of other security modules. Therefore, it is recommended that network segments be configured if many addresses need to be set.
- **http redirect direct-site** is used to configure the straight-through URL address for users, and **http redirect** is used to configure the straight-through IP address of the Web authentication server. The addresses configured using the two commands can be accessed without authentication, but they have different usages. It is recommended not to configure the IP address of the Web authentication server by using **http redirect direct-site**.

Configuration Steps

- Optional.
- Run the **http redirect direct-site** command to enable unauthenticated clients to access network resources.

Verification

- Configure the straight-through network resources.
- Check whether unauthenticated clients can access the configured network resources using PCs.

Related Commands

Configuring the Straight-Through Network Resources

Command	**http redirect direct-site** { *ipv4-address* [*ip-mask*] [**arp**]| *mac-address* | range *starip-address endip-address* } [description *description-str*] [group *group-name*]
Parameter Description	*ipv4-address*: Indicates the IPv4 address of the network exempt from authentication. *ip-mask*: Indicates the mask of the IPv4 address of the network exempt from authentication. mac-address: MAC address of the authentication-free network destination Startip-address: the starting ip of the continuous authentication-free network Endip-address: the end ip of the continuous authentication-free network Group-name: the group to which the authentication-free network belongs Description-str: Description information of the authentication-free network
Command Mode	Global configuration mode
Usage Guide	To set authentication-exempted ARP resource, use the **http redirect direct-arp** command preferentially. After configuring the port field, the authentication-free function only takes effect when the user is under the configured interface, not on the rest of the interface

Configuration Example

Configuring the Straight-Through Network Resources

Configuration Steps	● Configure the straight-through network resources as 192.168.0.0/16.
	FS(config)#http redirect direct-site 192.168.0.0 255.255.0.0
Verification	● Check whether the configuration is successful.
	FS(config)#show web-auth direct-site Direct sites: Address Mask ARP Binding Group Description --- --- --- --- --- 192.168.0.0 255.255.0.0 Off N/A N/A FS(config)#

5.4.12 Configuring the Straight-Through ARP Resource Range

Configuration Effect

- When ARP check or similar functions are enabled, the ARP learning performed by clients is controlled. As a result, clients cannot learn the ARPs of the gateway and other devices, which affects user experience. You can configure the straight-through ARP resource range to permit the ARP learning packets destined for the specified address to pass.

Notes

- When ARP check is enabled, you need to configure the gateway of the PCs connecting to the Layer-2 access device as a

straight-through ARP resource. Note the following point when you perform the configuration:

- When ARP check is enabled, if the outbound addresses of the PCs connecting to the Layer-2 access device are not the gateway address, configure the outbound addresses as straight-through ARP resources. If multiple outbound addresses exist, configure these addresses as straight-through ARP resources.

Configuration Steps

- Optional.
- If ARP check is enabled on the NAS, you must configure the free resources and gateway address as straight-through ARP resources.

Verification

- Configure straight-through ARP resources.
- Clear the ARP cache of the PC of an unauthenticated user. (Run the **arp -d** command in the Windows operating system.)
- Run the **ping** command on the PC to access the straight-through ARP resources.
- View the ARP cache on the PC (run the **arp -a** command in the Windows operating system) and check whether the PC learns the ARP address of the straight-through ARP resources.

Related Commands

Configuring the Straight-Through ARP Resource Range

Command	**http redirect direct-arp** {*ip-address* [*ip-mask*] }
Parameter Description	*ip-address*: Indicates the IP address of free resources. *ip-mask*: Indicates the mask of free resources.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Straight-Through ARP Resource

Configuration Steps	● Configure the straight-through ARP resource as 192.168.0.0/16.
	FS(config)#http redirect direct-arp 192.168.0.0 255.255.0.0
Verification	● Check whether the configuration is successful.
	FS(config)#show web-auth direct-arp Direct arps: Address Mask ---------- ---------- 192.168.0.0 255.255.0.0 FS(config)#

5.4.13 Configuring an Authentication-Exempted Address Range

Configuration Effect

- Exempt users from Web authentication when accessing reachable network resources. By default, no authentication-exempted address range is configured. All users must pass Web authentication before accessing network resources.
- The authentication-exempted address range can be configured as an IP address range or MAC address range.

Notes

N/A

Configuration Steps

- Optional.
- Perform this configuration to allow unauthenticated clients to access network resources.

Verification

- Configure an authentication-exempted user.
- Check whether the user can access the Internet without authentication.

Related Commands

Configuring an Authentication-Exempted Address Range

Command	**web-auth direct-host** { *ipv4-address* [*ipv4-mask*] [**arp**] [**port** *interface-name*]
Parameter Description	*ipv4-address*: Indicates the IPv4 address of the user exempt from authentication. *ip-mask*: Indicates the mask of the IPv4 address of the user exempt from authentication. *interface-name*: Indicates the name of the interface on which authentication exemption is enabled.
Command Mode	Global configuration mode
Usage Guide	The **arp** field is used to assign pass permissions to ARP packets. This field must be set when ARP check is enabled.

Configuration Example

Configuring an Authentication-Exempted Address Range

Configuration Steps	● Configure an authentication-exempted address range.
	FS (config)# web-auth direct-host 192.168.197.64
	● Set the range of consecutive users exempt from authentication to 10.0.0.1-12.0.0.1. FS(config)# web-auth direct-host range 10.0.0.1 12.0.0.1
Verification	● Check whether the configuration is successful.
	FS(config)#show web-auth direct-host Direct hosts: Address Mask Port ARP Binding Group Description

	---------- ---------- ---------- ---------- ---------- ---------- 192.168.197.64 255.255.255.255 Off N/A N/A FS(config)#

5.4.14 Configuring the Interval for Updating Online User Information

Configuration Effect

- The NAS or convergence device maintains and periodically updates the information of online users, including users' online duration, to monitor the usage of network resources. When the online duration threshold is reached, users will be prevented from using network resources.

Notes

- The user information updating interval must be configured as 60 or multiple of 60; otherwise, the system will select the minimum multiple of 60 above and closest to the actual configuration as the interval.

Configuration Steps

- Optional.
- Perform this configuration to allow unauthenticated clients to access network resources.

Verification

- Configure the interval for updating online user information.
- View the information of online users after the update interval has elapsed.

Related Commands

Configuring the Interval for Updating Online User Information

Command	**web-auth update-interval** *{ seconds }*
Parameter Description	*seconds*: Indicates the interval for updating online user information, in the unit of seconds. The value ranges from 30 to 3,600. The default value is 180s.
Command Mode	Global configuration mode
Usage Guide	To restore the default updating interval, run the **no web-auth update-interval** command in global configuration mode.

Configuration Example

Configuring the Interval for Updating Online User Information

Configuration Steps	● Set the interval for updating online user information to 60s.
	FS (config)# web-auth update-interval 60
Verification	● Check whether the configuration is successful.
	FS(config)#show run | include web-auth update-interval

	web-auth update-interval 60

5.4.15 Configuring Portal Detection

Configuration Effect

- Detect the availability of the active portal server periodically. When the active portal server is unavailable, the standby portal server takes over the services.
- FS Second-Generation Web Authentication provides two detection methods. One is that the NAS constructs and sends portal packets to the portal server. If the portal server returns response packets, the NAS determines that the portal server is available. Another is the NAS sends ping packets to the portal server. If the portal server returns response packets, the NAS determines that the portal server is available. Because some servers or intermediate network segments filter ping packets, the first method is commonly used. The ping detection method is only used based on special requirements. In FS First-Generation Web Authentication, the NAS connects to a port of the portal server and checks whether the port is reachable. If the portal is reachable, the NAS determines that the portal server is available.
- For the first method in the second-generation authentication, the interval of server availability detection is specified by the **interval** parameter, and the maximum number of packets that can be sent during each time of detection is specified by the **retransmit** parameter. If the portal server does not respond, the NAS determines that the portal server is unavailable. The timeout period for each packet is specified by the **timeout** parameter. The parameter settings are also supported by FS First-Generation Web Authentication.
- Portal server detection takes effect for FS First- and Second-Generation Web Authentication.
- If multiple portal servers are configured, these servers are working in active/standby mode.

Notes

- Multiple portal servers must be configured to realize failover when an error is detected on one server.
- Only one of the two detection methods can be used at a time in case of collision. If both detection methods are configured, a detection algorithm conflict will occur or the detection results will be inaccurate.
- The system will automatically select a detection method based on whether FS First- or Second-Generation Web Authentication is used.

Configuration Steps

- Optional.
- Configure multiple portal server templates applicable to FS First- or Second-Generation Web Authentication.

Verification

- Configure two portal server templates for FS First- or Second-Generation Web Authentication. Make the first template point to an unavailable server and the second template point to an available server.
- When the Console displays a log indicating that the portal server is not available, simulate the scenario where a user opens a browser to perform login authentication. Check whether the user is redirected to the second portal server.

Related Commands

Configuring Portal Detection

Command	**web-auth portal-check** [**interval** *intsec* [**timeout** *tosec]* [**retransmit** retries]
Parameter Description	*intsec*: Indicates the detection interval. The default value is 10s. *tosec:* Indicates the packet timeout period. The default value is 5s. *intsec*: Indicates the timeout retransmission times. The default value is 3 (times).
Command Mode	Global configuration mode
Usage Guide	In many network environments, only one portal server is deployed, and portal server detection does not need to be configured. If multiple portal servers exist, it is recommended that the parameters of portal server detection be not set to small values; otherwise, the NAS will send many packets within a short time, affecting performance.

Configuration Example

Configuring Portal Detection

Configuration Steps	● Configure portal detection.
	FS(config)#web-auth portal-check interval 20 timeout 2 retransmit 2
Verification	● Check whether the configuration is successful.
	FS(config)#show running-config … web-auth portal-check interval 20 timeout 2 retransmit 2 …

5.4.16 Configuring Portal Escape

Configuration Effect

- Allow new users to access the Internet without authentication when the portal server is not available.

Notes

- To use the portal escape function, you must configure portal detection.
- If multiple portal servers are configured, the escape function takes effect only when all the portal servers are not available.
- The escape function is intended only for the portal server, instead of the RADIUS server.

Configuration Steps

- Optional.
- Configure portal detection.
- Configure portal escape.
- (Optional) Configure the nokick attribute.

Verification

- Configure a portal server and disable the server.
- Configure the portal detection and escape functions.
- When the NAS detects that the portal server is not available, check whether a client accesses the Internet without authentication.

Related Commands

Configuring Portal Escape

Command	**web-auth portal-escape [nokick]**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure portal escape if the continuity of some critical services on the network needs to be maintained when the portal server is faulty. You must configure portal detection when you use this function. If the nokick attribute is configured, the system does not force users offline when the escape function takes effect. If the nokick attribute is deleted, the system forces users offline.

Configuration Example

Configuring Portal Escape

Configuration Steps	● Configure portal escape.
	FS(config)#web-auth portal-escape
Verification	● Check whether the configuration is successful.
	FS(config)#show running-config … web-auth portal-escape …

5.4.17 Enabling DHCP Address Check

Configuration Effect

- Allow only the clients that are allocated with IP addresses through DHCP to perform authentication.

Notes

- To use the DHCP address check function, you must configure DHCP snooping.
- DHCP address check is supported only for IPv4.
- DHCP address check is applicable only to FS Second-Generation Web Authentication and iPortal Web Authentication.
- The requirement that users obtain IP addresses through DHCP must be specified during network deployment. Those users cannot also use static IP addresses; otherwise, the existing users that use static IP addresses will be affected.

- If a few users need to use static IP addresses, configure these IP addresses as straight-through addresses, and these users are exempt from authentication.
- If DHCP address check needs to be enabled only on some interfaces or some VLANs of interfaces, disable the global DHCP address check and configure the VLAN range in which DHCP address check needs to be enabled in each interface.

Configuration Steps

- Optional.
- Enable DHCP snooping.
- Enable DHCP address check.

Verification

- Enable DHCP address check.
- Configure a static IP address that is not allocated by the DHCP server on a client.
- Connect the client to the Internet and check whether the STA cannot perform authentication.

Related Commands

Enabling Global DHCP Address Check

Command	**web-auth dhcp-check**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Configure DHCP address check to allow only the users who obtain IP addresses through DHCP to access the Internet. This function helps prevent the users who configure IP addresses without authorization from performing authentication to access the Internet.

Enabling Interface-based DHCP Address Check

Command	**web-auth dhcp-check {vlan [**vlan-list**]}**
Parameter Description	vlan-list: Indicates the VLAN range in which DHCP address check needs to be enabled in interface configuration mode.
Command Mode	Interface configuration mode
Usage Guide	If DHCP address check needs to be enabled only on some interfaces or some VLANs of interfaces, disable the global DHCP address check and configure the VLAN range in which DHCP address check needs to be enabled in each interface.

Configuration Example

Enabling DHCP Address Check

Configuration Steps	• Enable global DHCP address check.
	FS(config)#web-auth dhcp-check

Configuration Steps	● Enable interface-based DHCP address check.
	FS(config-if-TenGigabitEthernet 3/1)# web-auth dhcp-check vlan 1,3-4
Verification	● Check whether the configuration is successful.
	FS(config)#show running-config … web-auth dhcp-check … interface TenGigabitEthernet 3/1 web-auth dhcp-check vlan 1,3-4 …

5.4.18 Configuring the Portal Communication Port

Configuration Effect

- Configure the port (source port) used for the communication between the NAS and portal server.

Notes

- Only one port can be configured for the communication between the NAS and portal server.

Configuration Steps

- Configure a port as the portal communication port.

Verification

- After Web authentication is enabled, capture a packet on the portal server during the authentication process and check whether the source IP address of the packet is the IP address of the specified port.

Related Commands

Configuring the Portal Communication Port

Command	**ip portal source-interface** *interface-type interface-num*
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring the Portal Communication Port

Configuration Steps	● Configure an aggregate port as the portal communication port.
	FS(config)#ip portal source-interface Aggregateport 1
Verification	● Check whether the configuration is successful.
	FS(config)#show running-config ip portal source-interface Aggregateport 1

5.4.19 Configuring VLAN-Based Authentication on a Port

Configuration Effect

- With this function enabled, clients in a VLAN configured on a port of the NAS can initiate authentication. Otherwise, the authentication will not start.

Notes

- This function supports configuration of multiple VLANs. If no VLAN is specified, Web authentication is implemented based on ports.

Configuration Steps

- Configure port-based Web authentication.
- Configure the VLAN for Web authentication.

Verification

- After Web authentication is enabled, specify the VLAN in which clients can initiate authentication. The HTTP packets sent outside the specified VLAN cannot be redirected.

Related Commands

Configuring VLAN-Based Authentication on a Port

Command	**web-auth vlan-control** *vlan-list*
Parameter Description	*vlan-list* : Indicats the VLAN list to be authenticated.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

Configuring VLAN-Based Authentication on a Port

Configuration Steps	● Specify VLAN1 as the VLAN in which users can initiate authentication.

	FS(config-if-GigabitEthernet 0/14)#web-auth vlan-control 1
Verification	● Check whether the configuration is successful.
	FS(config)#show running-config … web-auth vlan-control 1

5.4.20 Disabling ARP Detection

Configuration Effect

- Disable the detection of user migration after its IP address or MAC address migrates.

Notes

- Address migration attacks may occur after ARP detection is disabled. ARP detection can be disabled in global configuration mode or in interface configuration mode. The interface configuration has a higher priority than the global configuration, and ARP detection is enabled by default in both global configuration mode and interface configuration mode. When the **no web-auth station-move arp-detect** command is executed on an interface, ARP detection is not performed when users of the interface are migrated to other interfaces.

Configuration Steps

Disabling ARP Detection

- Disable ARP detection.

Command	**no web-auth station-move arp-detect**
Parameter Description	N/A
Command Mode	Global configuration mode and interface configuration mode
Usage Guide	ARP detection is enabled by default. The interface configuration has a higher priority than the global configuration. Disable ARP detection on an interface that does not need ARP detection to prevent transmission of ARP detection packets during migration.

Verification

- After ARP detection is disabled, configure MAC address migration notices on different authentication ports, and check whether the users that perform Web authentication can migrate.

Configuration Example

Disabling ARP Detection

Configuration Steps	● Disable ARP detection.

	FS(config)#no web-auth station-move arp-detect FS(config-if-GigabitEthernet 2/1)# no web-auth station-move arp-detect
Verification	● Check whether the configuration is successful.
	FS(config)#show running-config … no web-auth station-move arp-detect interface GigabitEthernet 2/1 no web-auth station-move arp-detect

5.4.21 Upgrade Compatibility

Configuration Effect

- Some configuration commands are optimized in the 11.X series software and the command formats are changed. For details, see the subsequent description.
- The 10.X series software supports smooth upgrade without function loss. However, some commands are displayed in new formats after upgrade.
- When you run the commands in earlier formats in the **no** form in the 11.X series software, a message is displayed, indicating the **no** form is not supported. You need to perform the **no** operation in new command formats.

Configuration Steps

- It is recommended that you run commands in new formats.

Verification

- Check that function loss does not occur when the 10.X series software is upgraded to the 11.X series software, and commands are displayed and stored in new formats.
- The commands in new formats have the same functions as the commands in earlier formats.

Related Commands

Configuring the IP Address of the Portal Server in FS First-Generation Web Authentication

Command	**http redirect** *ip-address*
Parameter Description	*ip-address* : Indicates the ip address of the ePortal server in FS First-Generation Web Authentication.
Command Mode	Global configuration mode
Usage Guide	In the 11.X version, the command is converted into an eportalv1 template, and the **ip** command in template configuration mode is executed to configure and display the IP address of the portal server. For details, see section 5.4.1"Configuring FS First-Generation Web Authentication."

Configuring the URL of the Portal Server in FS First-Generation Web Authentication

Command	**http redirect homepge** *url*
Parameter Description	*url*: Indicates the URL of the ePortal server in FS First-Generation Web Authentication.
Command Mode	Global configuration mode
Usage Guide	In the 11.X version, the command is converted into an eportalv1 template, and the **ip** command in template configuration mode is executed to configure and display the IP address of the portal server. For details, see section 5.4.1"Configuring FS First-Generation Web Authentication."

Configuring the Portal Server

Command	**portal-server [eportal1 | eportalv2]**
Parameter Description	**eportav1**: Indicates the information of the portal server used in FS First-Generation Web Authentication. **eportav2**: Indicates the information of the portal server used in FS Second-Generation Web Authentication.
Command Mode	Global configuration mode
Usage Guide	In the 11.X version, the command is converted into an eportalv1 or eportalv2 template, and relevant information is filled in. The main parameters of the portal server include the IP address and URL of the server. The original command will be replaced by the **ip** command and **url** command in the template.

Configuring Web Authentication Control on a Port

Command	**web-auth port-control**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	In the 11.X version, the command is converted into **web-auth enable <type>**, in which **type** specifies the type (first or second generation) of Web authentication. The default type is FS First-Generation Web Authentication.

Configuring the IP-Only Binding Mode

Command	**web-auth port-control ip-only-mode**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	In the 11.X version, the command is converted into an eportalv1 or eportalv2 template, depending on the actual configuration. The server binding mode is configured and displayed by using the **bindmode** command in template configuration mode. For details, see section 5.4.1 "Configuring FS First-Generation Web Authentication" and section 5.4.2 "Configuring FS Second-Generation Web Authentication."

Configuring VLAN-Based Web Authentication

Command	**web-auth allow-vlan** *list*
Parameter	*list*: Indicates the list of VLANs for which Web authentication is enabled.

Description	
Command Mode	Global configuration mode
Usage Guide	In the 11.X version, the command is converted into a command used to configure VLAN-based SCC authentication exemption.

Displaying the Configuration Information of FS First-Generation Web Authentication

Command	**show http redirect**
Parameter Description	N/A
Command Mode	Privileged mode
Usage Guide	In the 11.X version, the command is unavailable and changed to **show web-auth template**.

Displaying the Port Control Information

Command	**show web-auth port-control**
Parameter Description	N/A
Command Mode	Privileged mode
Usage Guide	In the 11.X version, the command is unavailable and changed to **show web-auth control**.

Configuration Example

Configuring FS First-Generation Web Authentication

Configuration Steps	● Check that the NAS runs on the 10.X version and is configured with the IP address of the portal server used by FS First-Generation Web Authentication.
	FS(config)# http redirect 192.168.197.64
	● Upgrade the NAS to 11.X.
Verification	● Run the **show running-config** command after the upgrade and check whether the new command formats are used.
	FS#sh running-config web-auth template eportalv1 Ip 192.168.197.64 !

5.4.22 Configuring the Authenticated User Logout Delay on a Port

Configuration Effect

- Configure the delay after which the authenticated clients connected to a port go offline when the port fails.

Configuration Steps

Configuring the Authenticated User Logout Delay on a Port

- Configure the authenticated user logout delay on a port in global configuration mode.

Command	**web-auth linkdown-timeout**
Parameter Description	**timeout**: Indicates the logout delay. The default value is 60s.
Command Mode	Global configuration mode
Usage Guide	N/A

Verification

- Check that the authenticated clients connected to the faulty port go offline after the configured time has elapsed.

Configuration Example

Configuring the Authenticated User Logout Delay on a Port

Configuration Steps	- Configure the logout delay. FS(config)#web-auth linkdown-timeout {timeout}
Verification	- Check whether the configuration is successful. FS(config)#show running-config

5.4.23 Enabling the Perception-Free Feature

Configuration Effect

- Enable the perception-free feature to maintain the online state of users after port migration.

Configuration Steps

Enabling the Perception-Free Feature in Global Configuration Mode

Command	**web-auth station-move auto**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Verification

- Check that users do not go offline after migration.

Configuration Example

Configuring the Authenticated User Logout Delay on a Port

Configuration Steps	● Enable the perception-free feature. FS(config)# web-auth station-move auto
Verification	● Check whether the configuration is successful. FS(config)#show running-config

5.5 Monitoring

Clearing

Description	Command
Forces users offline.	**clear web-auth user** { **all** | **ip** *ip-address* | **mac** *mac-address* | **name** *name-string* }
Clears all the straight-through network resources.	**clear web-auth direct-site**
Clears all the authentication-exempted users.	**clear web-auth direct-host**

Displaying

Description	Command
Displays the basic parameters of Web authentication.	**show web-auth parameter**
Displays the Webauth template configuration.	**show web-auth template**
Displays the authentication-exempted host range.	**show web-auth direct-host**
Displays the straight-through address range.	**show web-auth direct-site**
Displays the straight-through ARP range.	**show web-auth direct-arp**
Displays the TCP interception port.	**show web-auth rdport**
Displays the online information of all users or specified users.	**show web-auth user**{ **all** | **ip** *ip-address* | **mac** *mac-address* | **name** *name-string* }
Displays the Webauth mapping information.	**show web-auth ip-mapping**
Displays the Webauth portal check information.	**show web-auth portal-check**
Displays online and offline records about users.	**show web-auth syslog ip** *ip-address*

Debugging

System resources are occupied when debugging information is output. Disable the debugging switch immediately after use.

Description	Command
Debugs Web authentication.	**debug web-auth all**

6 Configuring SCC

6.1 Overview

The Security Control Center (SCC) provides common configuration methods and policy integration for various access control and network security services, so that these access control and network security services can coexist on one device to meet diversified access and security control requirements in various scenarios.

Typical access control services are dot1x, Web authentication, Address Resolution Protocol (ARP) check, and IP Source Guard. The network security services include Access Control List (ACL), Network Foundation Protection Policy (NFPP), and anti-ARP gateway spoofing. When two or more access control or network security services are simultaneously enabled on the device, or when both access control and network security services are simultaneously enabled on the device, the SCC coordinates the coexistence of these services according to relevant policies.

For details about the access control and network security services, see the related configuration guide. This document describes the SCC only.

Protocol and Standards

N/A

6.2 Application

Typical Application	Scenario
Access Control of Extended Layer 2 Campus Networks	Students on a campus network can access the Internet based on dot1x client authentication or Web authentication. ARP spoofing between the students should be prevented. In addition, terminal devices in some departments (such as the headmaster's office) can access the Internet without authentication.

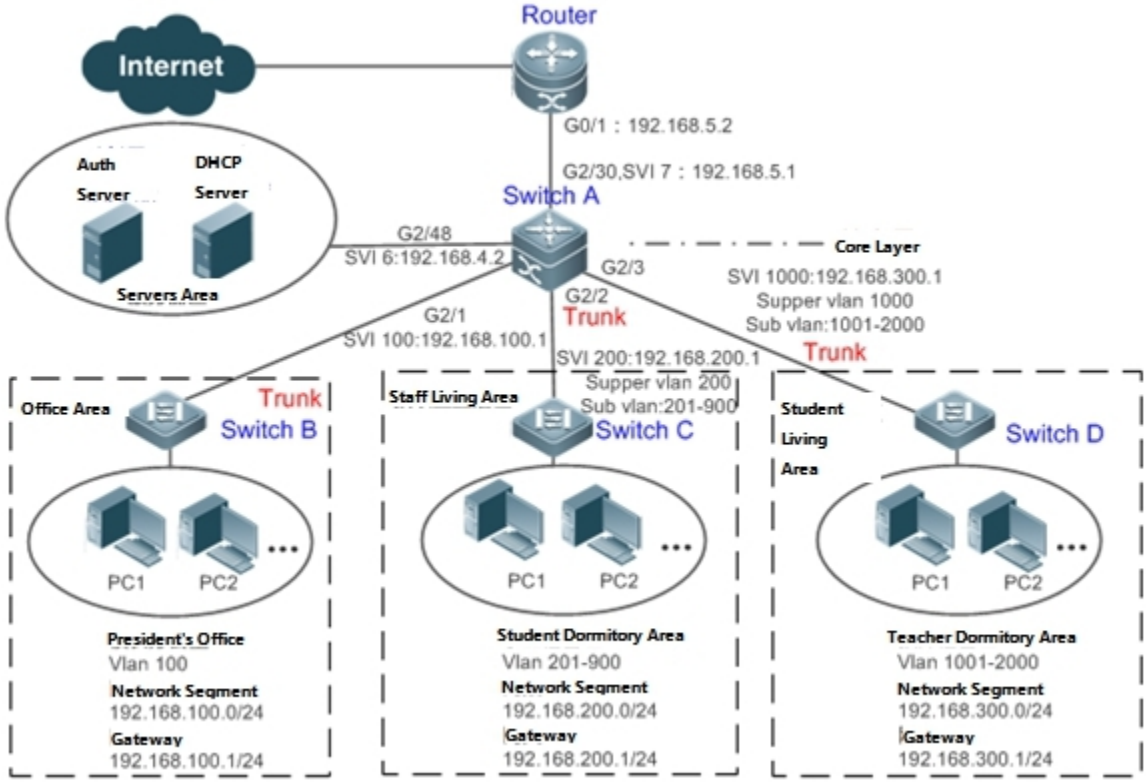
6.2.1 Access Control of Extended Layer 2 Campus Networks

Scenario

Students on a campus network of a university usually need to be authenticated through the dot1x client or Web before accessing the Internet, so as to facilitate accounting and guarantee the benefits of the university.

- The students can access the Internet through dot1x client authentication or Web authentication.
- ARP spoofing between the students is prevented, so as to guarantee the stability of the network.
- Terminal devices in some departments (such as the headmaster's office) can access the Internet without authentication.

Figure 6-1



Remarks	A traditional campus network is hierarchically designed, which consists of an access layer, a convergence layer and a core layer, where the access layer performs user access control. On an extended Layer 2 campus network, however, user access control is performed by a core switch, below which access switches exist without involving any convergence device in between. The ports between the core switch and the access switches (such as switches B, C, and D in Figure 6-1) are all trunk ports. The user access switches B, C, and D connect to PCs in various departments via access ports, and VLANs correspond to sub VLANs configured on the downlink ports of the core switch, so that access users are in different VLANs to prevent ARP spoofing. The core switch A connects to various servers, such as the authentication server and the DHCP server. Super VLANs and sub VLANs are configured on the downlink ports. One super VLAN correspond to multiple sub VLANs, and each sub VLAN represents an access user.

Deployment

- On the core switch, different access users are identified by VLAN and port numbers. Each access user (or a group of access users) corresponds to one VLAN. The ports on each access switch that connect to downstream users are configured as access ports, and one user VLAN is assigned to each access user according to VLAN planning. The core switch does not forward ARP requests. The core switch replies to the ARP requests from authenticated users only, so as to prevent ARP spoofing. On the core switch A, user VLANs are regarded as sub VLANs, super VLANs are configured, and SVIs corresponding to the super VLANs are configured as user gateways.
- On the downlink ports of the core switch (switch A in this example) that connect to the teachers' living area and the students' living area, both dot1x authentication and Web authentication are enabled, so that users can freely select either authentication mode for Internet access.
- Any special department (such as the headmaster's office in this example) can be allocated to a particular VLAN, and this VLAN can be configured as an authentication-exemption VLAN so that users in this department can access the Internet without

authentication.

6.3 Basic Concepts

Authentication-Exemption VLAN

Some special departments may be allocated to authentication-exemption VLANs to simplify network management, so that users in these departments can access network resources without authentication. For example, the headmaster's office can be divided into the authentication-exemption VLANs on the campus network, so that users in the headmaster's office can access the Internet without authentication.

IPv4 User Capacity

The number of IPv4 access users can be restricted to protect the access stability of online users on the Internet and improve the operational stability of the device.

The number of IPv4 access users is not restricted by default; that is, a large number of users can get online after being authenticated, till reaching the maximum hardware capacity of the device.

IPv4 access users include IP users (such as IP authenticated users) based on dot1x authentication, users based on Web authentication, and IP users manually bound (using IP source guard, ARP check, or other means).

Authenticated-User Migration

Online-user migration means that an online user can get authenticated again from different physical locations to access the network. On the campus network, however, for ease of management, students are usually requested to get authenticated from a specified location before accessing the Internet, but cannot get authenticated on other access ports. This means that the users cannot migrate. In another case, some users have the mobile office requirement and can get authenticated from different access locations. Then the users can migrate.

User Online-Status Detection

For a chargeable user, accounting starts immediately after the user passes the authentication and gets online. The accounting process does not end until the user actively gets offline. Some users, however, forget to get offline when leaving their PCs, or cannot get offline because of terminal problems. Then the users suffer certain economical losses as the accounting process continues. To more precisely determine whether a user is really online, we can preset a traffic value, so that the user is considered as not accessing the Internet and therefore directly brought offline when the user's traffic is lower than the preset value in a period of time or there is not traffic of the user at all in a period of time.

Features

Feature	Function
Authentication-Exemption VLAN	Users in a specified VLAN can be configured as authentication-exemption users.
IPv4 User Capacity	The IPv4 user capacity of a specified interface can be restricted to guarantee the access stability of users on the Internet.
Authenticated-User Migration	You can specify whether the authenticated can migrate.

User Online-Status Detection	You can specify whether to detect the traffic of online users, so that a user is forced offline when the traffic of the user is lower than a preset value in a period of time.

6.3.1 Authentication-Exemption VLAN

Authentication-exemption VLANs are used to accommodate departments with special access requirements, so that users in these departments can access the Internet without authentication such as dot1x or Web authentication.

Working Principle

Suppose the authentication-exemption VLAN feature is enabled on a device. When the device detects that a packet comes from an authentication-exemption VLAN, access control is not performed. In this way, users in the authentication-exemption VLAN can access the Internet without authentication. The authentication-exemption VLAN feature can be regarded as a kind of applications of secure channels.

- A maximum of 100 authentication-exemption VLANs can be configured.
- The authentication-exemption VLANs occupy hardware entries. When access control such as authentication is disabled, configuring authentication-exemption VLANs has the same effect as the case where no authentication-exemption VLANs are configured. Therefore, it is recommended that authentication-exemption VLANs be configured for users who need to access the Internet without authentication, only when the access control function has been enabled.
- Although packets from authentication-exemption VLANs are exempt from access control, they still need to be checked by a security ACL. If the packets of the users in an authentication-exemption VLAN are denied according to the security ACL, the users still cannot access the Internet.
- In gateway authentication mode, the device does not initiate any ARP request to a user in an authentication-exemption VLAN, and the ARP proxy will not work. Therefore, in gateway authentication mode, users in different authentication-exemption VLANs cannot access each other unless the users have been authenticated.

6.3.2 IPv4 User Capacity

To improve the operational stability of the device and guard against brutal force impacts from unauthorized users, you can restrict the total number of IPv4 access users on a certain port of the device.

Working Principle

If the total number of IPv4 access users is restricted, new users going beyond the total number cannot access the Internet.

- Only the switches support the restriction on the number of IPv4 access users.
- The number of IPv4 access users is not restricted on the device by default, but depends on the hardware capacity of the device.
- The number of IPv4 access users includes the IPv4 authenticated users based on dot1x authentication, IPv4 users based on Web authentication, and IPv4 users based on various binding functions. Because the number of IPv4 access users is configured in interface configuration mode, the restriction includes both the number of IPv4 users generated on the port and IPv4 users globally generated. For example, you can set the maximum number of IPv4 access users on the Gi 0/1 port to 2, run commands to bind an IPv4 user to the port, and then run commands to bind a global IPv4 user to the port. Actually there are already two access users on the port. If you attempt to bind another IPv4 user or another global IPv4 user to the port, the binding operation fails.

6.3.3 Authenticated-User Migration

On an actual network, users do not necessarily access the Internet from a fixed place. Instead, users may be transferred to another

department or office after getting authenticated at one place. They do not actively get offline but remove network cables and carry their mobile terminals to the new office to access the network. Then this brings about an issue about authenticated-user migration. If authenticated-user migration is not configured, a user who gets online at one place cannot get online at another place without getting offline first.

Working Principle

When authenticated-user migration is enabled, the dot1x or Web authentication module of the device detects that the port number or VLAN corresponding to a user's MAC address has changed. Then the user is forced offline and needs to be authenticated again before getting online.

Only the switches support authenticated-user migration. In addition, cross-switch migration is not supported. For example, authentication and migration are enabled on two N18000, and a user gets online after being authenticated on one of the two N18000. If the user attempts to migrate to the other N18000, the migration fails.

The authenticated-user migration function requires a check of users' MAC addresses, and is invalid for users who have IP addresses only.

The authenticated-user migration function enables a user who gets online at one place to get online at another place without getting offline first. If the user gets online at one place and then gets offline at that place, or if the user does not get online before moving to another place, the situation is beyond the control range of authenticated-user migration.

During migration, the system checks whether the VLAN ID or port number that corresponds to a user's MAC address has changed, so as to determine whether the user has migrated. If the VLAN ID or port number is the same, it indicates that the user does not migrate; otherwise, it indicates that the user has migrated. According to the preceding principle, if another user on the network uses the MAC address of an online user, the system will wrongly disconnect the online user unless extra judgment is made. To prevent such a problem, the dot1x or Web authentication will check whether a user has actually migrated. For a user who gets online through Web authentication or dot1x authentication with IP authorization, the dot1x or Web authentication sends an ARP request to the original place of the user if detecting that the same MAC address is online in another VLAN or on another port. If no response is received within the specified time, it indicates that the user's location has indeed changed and then the migration is allowed. If a response is received within the specified time, it indicates that the user actually does not migrate and a fraudulent user may exist on the network. In the latter case, the migration is not performed. The ARP request is sent once every second by default, and sent for a total of five times. This means that the migration cannot be confirmed until five seconds later. Timeout-related parameters, including the probe interval and probe times, can be changed using the **arp retry** times *times* and **arp retry** interval *interval* commands. For details about the specific configuration, see *ARP-SCG.doc*. It should be noted that the migration check requires the configuration of IP authorization for users based on dot1x authentication. In addition, the ARP probe is triggered only for user migration in gateway authentication mode but not triggered for user migration in access authentication mode.

6.3.4 User Online-Status Detection

After a user accesses the Internet, the user may forget to get offline or cannot actively get offline due to terminal faults. In this case, the user will keep being charged and therefore will suffer a certain economical loss. To protect the benefits of users on the Internet, the device provides a function to detect whether the users are really online. If the device considers that a user is not online, the device actively disconnects the user.

Working Principle

A specific detection interval is preset on the device. If a user's traffic is lower than a certain value in this interval, the device considers that the user is not using the network and therefore directly disconnects the user.

The switches support the user online-status detection function.

The user online-status detection function applies to only users who get online through dot1x or Web authentication.

Currently, the N18000 supports zero-traffic detection only.

Currently, due to hardware chip restrictions of the N18000, the time to disconnect a user without any traffic relates to the configured MAC address aging time. If the traffic detection interval is set to m minutes and the MAC address aging time is set to n minutes, the interval from the moment when an authenticated user leaves the network without actively getting offline to the moment when the user is disconnected upon detection of zero traffic is about [m, m+n] minutes. In other words, if an online user does not incur any Internet access traffic, the user is disconnected about [m, m+n] minutes later.

6.4 Configuration

Configuration Item	Suggestions and Related Commands	
Configuring Authentication-Exemption VLANs	Optional configuration, which is used to specify the users of which VLANs can access the Internet without authentication.	
	[no] direct-vlan	Configures authentication-exemption VLANs.
Configuring the IPv4 User Capacity	Optional configuration, which is used to specify the maximum number of users who are allowed to access a certain interface.	
	[no] nac-author-user maximum	Configures the number of IPv4 users who are allowed to access a certain interface.
Configuring Authenticated-User Migration	Optional configuration, which is used to specify whether online users with static MAC addresses can migrate.	
	[no] station-move permit	Configures whether authenticated users can migrate.
Configuring User Online-Status Detection	Optional configuration, which is used to specify whether to enable the user online-status detection function.	
	offline-detect interval threshold	Configures the parameters of the user online-status detection function.
	no offline-detect	Disables the user online-status detection function.
	default offline-detect	Restores the default user online-status detection mode.

6.4.1 Configuring Authentication-Exemption VLANs

Configuration Effect

Configure authentication-exemption VLANs, so that users in these VLANs can access the Internet without experiencing dot1x or Web authentication.

Configure authentication-exemption VLANs on a port, so that only users in specified VLANs on the port can access the Internet without experiencing authentication.

Precautions

Authentication-exemption VLANs only mean that users in these VLANs do not need to experience a check related to access authentication, but still need to experience a check based on a security ACL. If specified users or VLANs are denied according to the security ACL, corresponding users still cannot access the Internet. Therefore, during ACL configuration, you need to ensure that specified VLANs or specified users in the authentication-exemption VLANs are not blocked if you hope that users in the authentication-exemption VLANs can access the Internet without being authenticated.

Configuration Method

Configuring Authentication-Exemption VLANs

- Optional configuration. To spare all users in certain VLANs from dot1x or Web authentication, configure these VLANS as authentication-exemption VLANs.
- Perform this configuration on access, convergence, or core switches depending on user distribution.
- Authentication-exemption VLANs can be configured in interface configuration mode.

Command	**[no] direct-vlan** *vlanlist*
Parameter Description	**no**: If the command carries this parameter, it indicates that the authentication-exemption VLAN configuration will be deleted. *vlanlist*: This parameter indicates the list of authentication-exemption VLANs to be configured or deleted.
Defaults	No authentication-exemption VLAN has been configured.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure or delete authentication-exemption VLANs.

Verification

Check the authentication-exemption VLAN configuration using the following method:

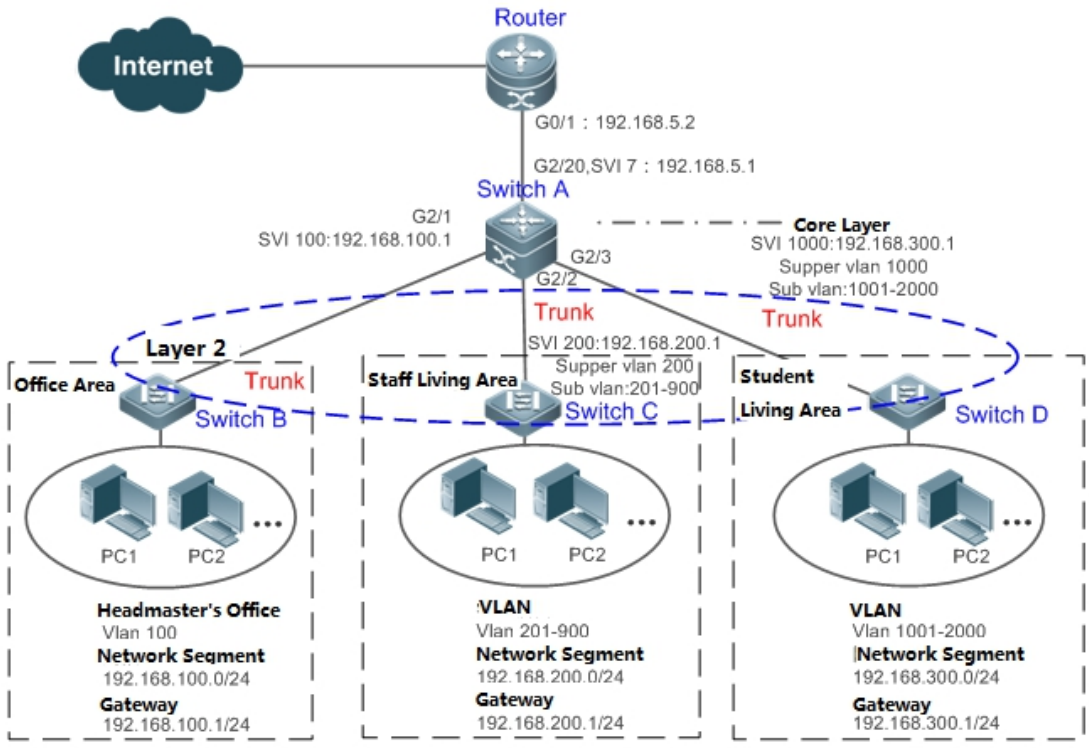
- Enable dot1x authentication on downlink ports that connect to user terminals, add the downlink ports that connect to the user terminals to a specific VLAN, and configure the VLAN as an authentication-exemption VLAN. Then open the Internet Explorer, and enter a valid extranet address (such as www.google.com). If the users can open the corresponding webpage on the Internet, it indicates that the authentication-exemption VLAN is valid; otherwise, the authentication-exemption VLAN does not take effect.
- Use the **show direct-vlan** command to check the authentication-exemption VLAN configuration on the device.

Command	**show direct-vlan**
Parameter Description	-
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	Global configuration mode
Usage Example	FS#show direct-vlan direct-vlan 100

Configuration Examples

The following configuration example describes SCC-related configuration only.

Configuring Authentication-exemption VLANs so that Specific Users Can Access the Internet Without Being Authenticated

Scenario **Figure 6-2**	
Configuration Steps	● On switch A (which is the core gateway device), set the GI 2/1 port as a trunk port, and enable dot1x authentication on this port. ● On switch A (which is the core gateway device), configure VLAN 100 to which the headmaster's office belongs as an authentication-exemption VLAN.
Switch A	SwitchA(config)#vlan 100 SwitchA(config-vlan)#exit SwitchA(config)#direct-vlan 100 SwitchA(config)#int GigabitEthernet 0/1 SwitchA(config-if-GigabitEthernet 0/1)#switchport mode trunk SwitchA(config-if-GigabitEthernet 0/1)#dot1x port-control auto *Oct 17 16:06:45: %DOT1X-6-ENABLE_DOT1X: Able to receive EAPOL packet and DOT1X authentication enabled.
Verification	● Open the Internet Explorer from any PC in the headmaster's office, enter a valid extranet address, and confirm that the corresponding webpage can be opened. ● Use the **show direct-vlan** command to check whether the authentication-exemption VLAN is valid.
Switch A	SwitchA(config)#show direct-vlan direct-vlan 100

6.4.2 Configuring the IPv4 User Capacity

Configuration Effect

Configure the IPv4 user capacity, so as to restrict the number of users who are allowed to access an access port.

Precautions

N/A

Configuration Method

Configuring the IPv4 User Capacity

- Optional configuration. To limit the maximum of users who are allowed to access an access port, configure the IPv4 user capacity. The access user capacity is not limited on an access port by default. Suppose the user capacity limit is configured on a specific interface. When the number of authenticated users on the interface reaches the maximum, new users cannot be authenticated on this interface and cannot get online, until existing authenticated users get offline on the interface.
- Perform this configuration on access switches, which may be access switches on the network edge or core gateway devices.

Command	**nac-author-user maximum** *max-user-num* **no nac-author-user maximum**
Parameter Description	**no**: If the command carries this parameter, it indicates that the limit on the IPv4 access user capacity will be removed from the port. *max-user-num*: This parameter indicates the maximum number of IPv4 users who allowed to access the port. The value range is from 1 to 1024.
Defaults	The number of IPv4 access users is not limited.
Command Mode	Interface configuration mode
Usage Guide	Use this command to limit the number of IPv4 access users on a specific access port.

Verification

Check the IPv4 user capacity configuration on a port using the following method:

- dot1x authentication: When the number of users who get online based on 1x client authentication on the port reaches the specified user capacity, no any new user can get online from this port.
- Web authentication: When the number of users who get online based on Web authentication on the port reaches the specified user capacity, no any new user can get online from this port.
- Use the **show nac-author-user** [**interface** *interface-name*] command to check the IPv4 user capacity configured on the device.

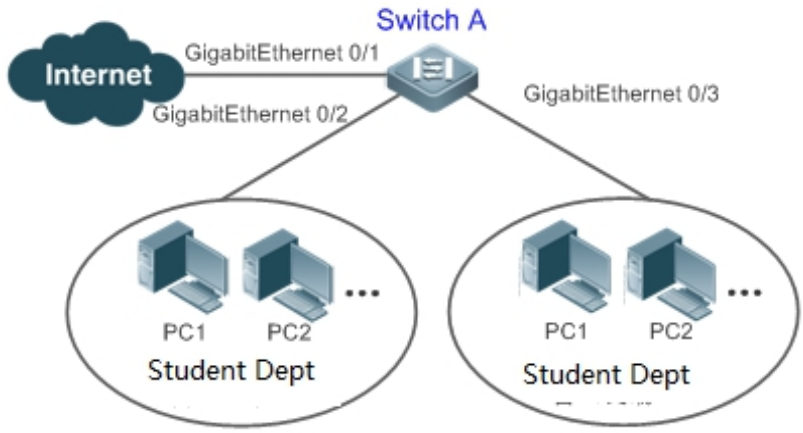
Command	**show nac-author-user** [**interface** *interface-name*]
Parameter Description	***interface-name***: This parameter indicates the interface name.
Command Mode	Privileged EXEC mode, global configuration mode, or interface configuration mode
Usage Guide	Global configuration mode
Usage Example	FS#show nac-author-user interface GigabitEthernet 0/1

	Port Cur_num Max_num -------- ------- ------- Gi0/1 0 **4**

Configuration Examples

The following configuration example describes SCC-related configuration only.

Restricting the Number of IP4 Users on a Port to Prevent Excessive Access Terminals from Impacting the Network

Scenario **Figure 6-3**	
Configuration Steps	● Assume that the dot1x authentication environment has been well configured on the access switch A, and dot1x authentication is enabled on the Gi 0/2 port. ● Set the maximum number of IPv4 access users on the Gi 0/2 port to 4.
Switch A	SwitchA(config)#int GigabitEthernet 0/2 SwitchA(config-if-GigabitEthernet 0/2)#nac-author-user maximum 4
Verification	● Perform dot1x authentication for all the four PCs in the dormitory, so that the PCs get online. Then take an additional terminal to access the network, and attempt to perform dot1x authentication for this terminal. Verify that the terminal cannot be successfully authenticated to get online. ● Use the **show nac-author-user** command to check whether the configuration has taken effect.
Switch A	SwitchA(config)#show nac-author-user Port Cur_num Max_num -------- ------- ------- Gi0/1 0 **4**

6.4.3 Configuring Authenticated-User Migration

Configuration Effect

By default, when a user gets online after passing dot1x or Web authentication at a physical location (which is represented by a specific access port plus the VLAN number) and quickly moves to another physical location without getting offline, the user cannot get online

through dot1x or Web authentication from the new physical location, unless the authenticated-user migration feature has been configured in advance.

Precautions

- If the authenticated-user migration feature is not yet configured, an online user cannot get online from the new physical location after quickly moving from one physical location to another physical location without getting offline first. However, if the user gets offline before changing the physical location or gets offline during the location change (for example, the user online-status detection function disconnects the user), the user can still normally get online after being authenticated at the new physical location, even if the authenticated-user migration feature is not configured.
- After moving to the new physical location, the online user needs to perform dot1x or Web authentication so as to get online.

Configuration Method

Configuring Authenticated-User Migration

- Optional configuration. To allow users to be authenticated and get online from different physical locations, enable the authenticated-user migration function.
- Perform this configuration on access, convergence, or core switches depending on user distribution.

Command	**[no] station-move permit**
Parameter Description	**no station-move permit**: Indicates that authenticated-user migration is not permitted. **station-move permit**: Indicates that authenticated-user migration is permitted.
Defaults	Authenticated-user migration is not permitted; that is, when a user getting online from one physical location on the network moves to another physical location and attempts to get online from the new physical location without getting offline first, the authentication fails and the user cannot get online from the new physical location.
Command Mode	Global configuration mode
Usage Guide	Use this command to configure authenticated-user migration.

Verification

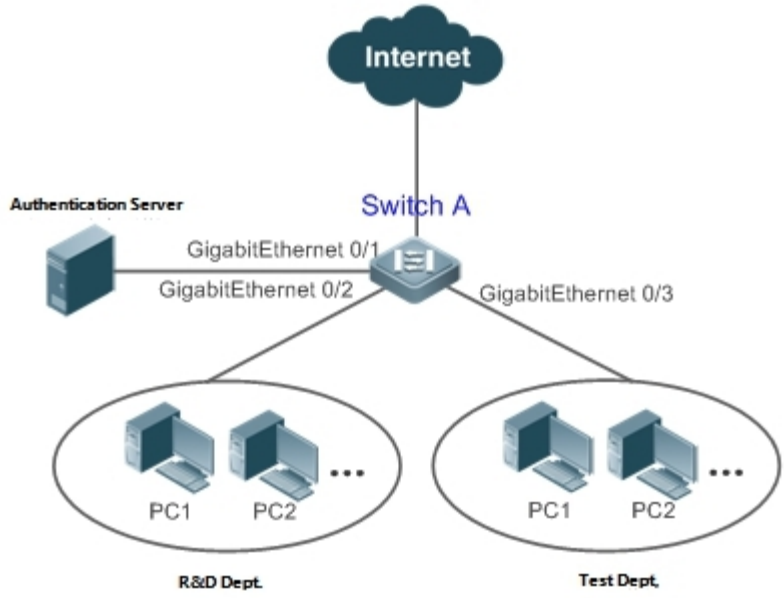
Check the authenticated-user migration configuration using the following method:

- A PC is authenticated and gets online from a dot1x-based port of the device using dot1x SU client, and does not actively get offline. Move the PC to another port of the device on which dot1x authentication is enabled, and perform dot1x authentication again. Check whether the PC can successfully get online.

Configuration Examples

The following configuration example describes SCC-related configuration only.

Configuring Online-User Migration so that an Online User Can Perform Authentication and Get Online from Different Ports Without Getting Offline First

Scenario **Figure 6-4**	
Configuration Steps	● Enable dot1x authentication on access ports Gi 0/2 and Gi 0/3, and configure authentication parameters. The authentication is MAC-based. ● Configure online-user migration.
Switch A	sw1(config)#station-move permit
Verification	● A lap-top PC in the R&D department performs authentication using dot1x SU client, and gets online. Remove the network cable from the PC, connect the PC to the LAN where the test department resides, and perform dot1x authentication for the PC again using dot1x SU client. Confirm that the PC can successfully get online.
Switch A	sw1(config)#show running-config | include station station-move permit

6.4.4 Configuring User Online-Status Detection

Configuration Effect

After the user online-status detection function is enabled, if a user's traffic is lower than a certain threshold within the specified period of time, the device automatically disconnects the user, so as to avoid the economical loss incurred by constant charging to the user.

Precautions

It should be noted that if disconnecting zero-traffic users is configured, generally software such as 360 Security Guard will run on a user terminal by default. Then such software will send packets time and again, and the device will disconnect the user only when the user's terminal is powered off.

Configuration Method

Configuring User Online-Status Detection

- Optional configuration. A user is disconnected if the user does not involve any traffic within eight hours by default.
- Perform this configuration on access, convergence, or core switches depending on user distribution. The configuration acts on only the configured device instead of other devices on the network.
- If the traffic threshold parameter threshold is set to 0, it indicates that zero-traffic detection will be performed.

Command	**offline-detect interval** *interval* **threshold** *threshold* **no offline-detect** **default offline-detect**
Parameter Description	*interval*: This parameter indicates the offline-detection interval. The value range is from 6 to 65535 in minutes on a switch or from 1 to 65535 in minutes on a non-switch device. The default value is 8 hours, that is, 480 minutes. *threshold*: This parameter indicates the traffic threshold. The range depends on the chip.The default value is 0, indicating that the user is disconnected when no traffic of the user is detected. **no offline-detect**: Disables the user online-status detection function. **default offline-detect**: Restores the default value. In other words, an online user will be disconnected when the device detects that the user does not have any traffic within eight hours.
Defaults	8 hours
Command Mode	Global configuration mode
Usage Guide	Use this command to configure user online-status detection, so that a user is disconnected when its traffic is lower than a specific threshold within a specific period of time. Use the **no offline-detect** command to disable the user online-status detection function, or use the **default offline-detect** command to restore the default detection mode.

Verification

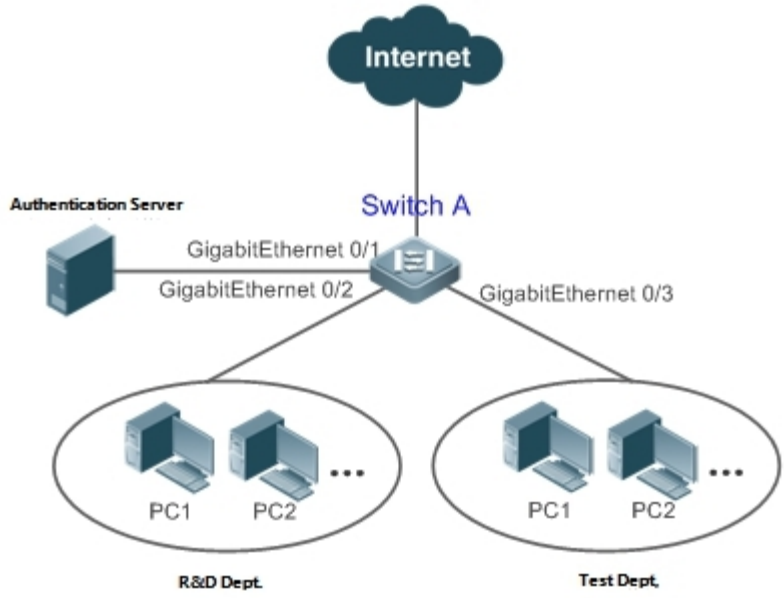
Check the user online-status detection configuration using the following method:

- After the user online-status detection function is enabled, power off the specified authenticated terminal after the corresponding user gets online. Then wait for the specified period of time, and run the online user query command associated with dot1x or Web authentication on the device to confirm that the user is already offline.

Configuration Examples

The following configuration example describes SCC-related configuration only.

Configuring User Online-Status Detection so that a User Is Disconnected if the User Does Not Have Traffic Within Five Minutes

Scenario **Figure 6-5**	
Configuration Steps	● Enable dot1x authentication on the access port Gi 0/2, and configure authentication parameters. The authentication is MAC-based. ● Configure user online-status detection so that a user is disconnected if the user does not have traffic within five minutes.
Switch A	sw1(config)# offline-detect interval 5 threshold 0
Verification	● Perform dot1x authentication using dot1x SU client for a PC in the R&D department, so that the PC gets online. Then power off the PC, wait for 6 minutes, and run the online user query command available with dot1x authentication on switch 1 to confirm that the user of the PC is already offline.
Switch A	sw1(config)#show running-config | include offline-detect offline-detect interval 5

6.5 Monitoring

Displaying

Command	Function
show direct-vlan	Displays the authentication-exemption VLAN configuration.
show nac-author-user [**interface** *interface-name*]	Displays information about IPv4 user entries on a specific interface.

Debugging

System resources are occupied when debugging information is output. Therefore, close the debugging switch immediately after use.

Command	Function
debug scc event	Debugs the SCC running process.

debug scc user [mac | author | mac]	Debugs SCC user entries.
debug scc acl-show summary	Debugs ACLs stored in the current SCC and delivered by various services.
debug scc acl-show all	Debugs all ALCs stored in the current SCC.

7 Configuring Global IP-MAC Binding

7.1 Overview

Enable the global IP-MAC binding function manually to verify the input packets. If a specified IP address is bound with a MAC address, the device receives only the IP packets containing matched IP address and MAC address. The other packets are discarded.

The address bounding feature is used to verify the input packets. Note that the address binding feature takes precedence over the 802.1X authentication, port security, and access control list (ACL).

7.2 Applications

Application	Description
Global IP-MAC Binding	Only hosts with the specified IP addresses can access the network, and the hosts connected to a device can move freely.

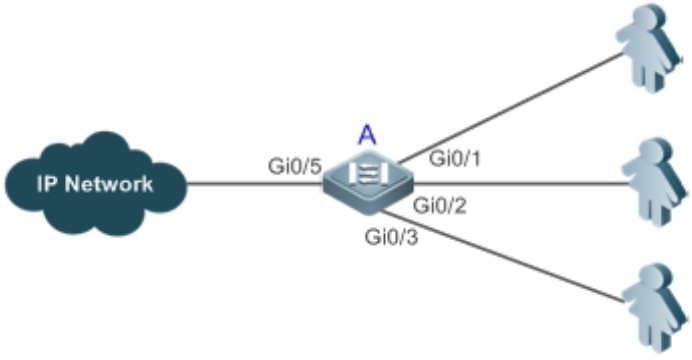
7.2.1 Global IP-MAC Binding

Scenario

The administrator assigns a fixed IP address for each host to facilitate management.

- Only hosts with the specified IP addresses can access the external network, which prevents IP address embezzlement by unauthorized hosts.
- Hosts can move freely under the same device.

Figure 7- 1



Remarks	A is an access device. A user is a host configured with a static IP address. IP Network is an external IP network.

Deployment

- Manually configure the global IP-MAC binding. (Take three users as an example.)

User	MAC Address	IP Address
User 1	00d0.3232.0001	192.168.1.10
User 2	00d0.3232.0002	192.168.1.20

User 3	00d0.3232.0003	192.168.1.30

- Enable the IP-MAC binding function globally.
- Configure the uplink port (Gi0/5 port in this example) of the device as the exclude port.

7.3 Features

Basic Concepts

IPv6 Address Binding Mode

IPv6 address binding modes include Compatible, Loose, and Strict. The default mode is Strict. If IPv4-MAC binding is not configured, the IPv6 address binding mode does not take effect, and all IPv4 and IPv6 packets are allowed to pass through. If IPv4-MAC binding is configured, the IPv6 address binding mode takes effect, and the device forwards IPv4 and IPv6 packets based on the forwarding rules described in the following table:

Mode	IPv4 Packet Forwarding Rule	IPv6 Packet Forwarding Rule
Strict	Packets matching the global IPv4-MAC binding are forwarded.	Packets matching the global IPv6-MAC binding are forwarded. (The binding is generated by other access security functions, such as port security and IPv6 Source Guard.)
Loose	Packets matching the global IPv4-MAC binding are forwarded.	If IPv6+MAC address binding is configured, packets matching the IPv6-MAC binding are forwarded. (The binding is generated by other access security functions, such as port security and IPv6 Source Guard.) If IPv6-MAC binding does not exist, all IPv6 packets are forwarded.
Compatible	Packets matching the global IPv4-MAC binding are forwarded.	If the IPv6 packets contain a MAC address matching the MAC address in the IPv4-MAC binding, the IPv6 packets are forwarded. Packets matching the global IPv6-MAC binding conditions are forwarded. (The binding is generated by other access security functions, such as port security and IPv6 Source Guard.)

Exclude Port

By default, the IP-MAC binding function takes effect on all ports of the device. You can configure exclude ports so that the address binding function does not take effect on these ports. In practice, the IP-MAC bindings of the input packets on the uplink port are not fixed. Generally, the uplink port of the device is configured as the exclude port so that the packets on the uplink port are not checked for IP-MAC binding.

Overview

Feature	Description
Configuring Global IP-MAC Binding	Control forwarding of IPv4 or IPv6 packets.
Configuring the IPv6 Address Binding Mode	Change the IPv6 packet forwarding rules.
Configuring the Exclude Port	Disable the global address binding function on the specified port.

7.3.1 Configuring Global IP-MAC Binding

Working Principle

Enable the global IP-MAC binding function manually to verify the input packets. If a specified IP address is bound with a MAC address, the device receives only the IP packets containing matched IP address and MAC address. The other packets are discarded.

Related Configuration

Configuring IP-MAC Binding

Run the **address-bind** command in global configuration mode to add or delete an IPv4-MAC binding.

Enabling the IP-MAC Binding Function

Run the **address-bind install** command in global configuration mode to enable the IP-MAC binding function. By default, this function is disabled.

7.3.2 Configuring the IPv6 Address Binding Mode

Working Principle

After the global IPv4-MAC binding is configured and enabled, IPv6 packets are forwarded based on the IPv6 address binding mode. IPv6 binding modes include Compatible, Loose, and Strict.

Related Configuration

Configuring the IPv6 Address Binding Mode

By default, the IPv6 address binding mode is Strict.

Run the **address-bind ipv6-mode** command to specify an IPv6 address binding mode.

7.3.3 Configuring the Exclude Port

Working Principle

Configure an exclude port so that the address binding function does not take effect on this port.

Related Configuration

Configuring the Exclude Port

Run the **address-bind uplink** command to configure an exclude port. By default, no port is the exclude port.

7.4 Configuration

Configuration	Description and Command	
Configuring Global IP-MAC Binding	(Mandatory) It is used to configure and enable address binding.	
	address-bind	Configures a global IPv4-MAC binding.
	address-bind install	Enables the address binding function.
	address-bind binding-filter logging	Configures a logging filter of global IPv4-MAC binding.
Configuring the IPv6 Address Binding Mode	(Optional) It is used to configure the IPv6 address binding mode.	
	address-bind ipv6-mode	Configures the IPv6 address binding mode.

Configuring the Exclude Port	(Optional) It is used to disable the address binding function on a specified port.	
	address-bind uplink	Configures an exclude port.

7.4.1 Configuring Global IP-MAC Binding

Configuration Effect

- Configure a global IPv4-MAC binding.
- Enable the address binding function to control forwarding of the IPv4 or IPv6 packets.

Notes

- If you run the **address-bind install** command without IP-MAC binding configured, IP-MAC binding does not take effect and all packets are allowed to pass through.

Configuration Steps

Configuring Global IP-MAC Binding

- (Mandatory) Perform this configuration in global configuration mode.

Enabling the Address Binding Function

- (Mandatory) Perform this configuration in global configuration mode.

Configuring a Logging Filter of Global IP-MAC Binding

- (Optional) Perform this configuration in global configuration mode.

Verification

Run the **show run** or **show address-bind** command to check whether the configuration takes effect.

Related Commands

Configuring Global IP-MAC Binding

Command	**address-bind** { *ip-address* | *ipv6-address* } *mac-address*
Parameter Description	*ip-address*: Indicates the bound IPv4 address. *ipv6-address*: Indicates the bound IPv6 address. *mac-address*: Indicates the bound MAC address.
Command Mode	Global configuration mode
Configuration Usage	Run this command to configure the binding relationship between an IPv4/IPv6 address and a MAC address.

Enabling the Address Binding Function

Command	**address-bind install**
Parameter Description	N/A
Command	Global configuration mode

Mode	
Configuration Usage	Run this command to enable the global IP-MAC binding function. This function is used to control forwarding of IPv4 or IPv6 packets.

Configuring a Logging Filter of Global IP-MAC Binding

Command	**address-bind binding-filter logging** [**rate-limit** *rate*]
Parameter Description	**rate-limit** *rate*: Indicates the printing rate of logging filter of global IPv4-MAC binding.
Command Mode	Global configuration mode
Configuration Usage	By default, the rate is 10 logs per minute. When a logging filter is configured, alert logs are printed if IP packets not containing matched IP address and MAC address are detected. When a logging filter is configured, the number of non-printed logs is prompted if the actual printing rate exceeds the set rate.

Configuration Example

Configuring Global IP-MAC Binding and Enabling Address Binding

Configuration Steps	● Configure a global IPv4-MAC binding. ● Enable the address binding function.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# address-bind 192.168.5.1 00d0.f800.0001 FS(config)# address-bind install
Verification	Display the global IP-MAC binding on the device.
	FS#show address-bind Total Bind Addresses in System : 1 IP Address Binding MAC Addr --- --- 192.168.5.1 00d0.f800.0001

7.4.2 Configuring the IPv6 Address Binding Mode

Configuration Effect

- Change the IPv6 address binding mode so as to change the forwarding rules for IPv6 packets.

Configuration Steps

Configuring the IPv6 Address Binding Mode

- (Optional) Perform this configuration when you want to change the forwarding rules for IPv6 packets.

Verification

- Run the **show run** command to check whether the configuration takes effect.

Related Commands

Configuring the IPv6 Address Binding Mode

Command	**address-bind ipv6-mode** { **compatible** | **loose** | **strict** }
Parameter Description	**compatible**: Indicates the Compatible mode. **loose**: Indicates the Loose mode. **strict**: Indicates the strict mode.
Command Mode	Global configuration mode
Configuration Usage	N/A

Configuration Example

Configuring the IPv6 Address Binding Mode

Configuration Steps	• Configure a global IP-MAC binding. • Enable the address binding function. • Set the IPv6 address binding mode to Compatible.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# address-bind 192.168.5.1 00d0.f800.0001 FS(config)# address-bind install FS(config)# address-bind ipv6-mode compatible
Verification	Run the **show run** command to display the configuration on the device.

7.4.3 Configuring the Exclude Port

Configuration Effect

- The address binding function is disabled on the exclude port, and all IP packets can be forwarded.

Notes

- The configuration can be performed only on a switching port or an L2 aggregate port.

Configuration Steps

Configuring the Exclude Port

- (Optional) Perform this configuration in global configuration mode when you want to disable the address binding function on a specified port.

Verification

Run the **show run** or **show address-bind uplink** command to check whether the configuration takes effect.

Related Commands

Configuring the Exclude Port

Command Syntax	**address-bind uplink** *interface-id*
Parameter Description	*interface-id*: Indicates the ID of a switching port or an L2 aggregate port.
Command Mode	Global configuration mode
Configuration Usage	N/A

Configuration Example

Configuring the Exclude Port

Configuration Steps	● Create a global IPv4-MAC binding. ● Enable the address binding function. ● Configure an exclude port.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# address-bind 192.168.5.1 00d0.f800.0001 FS(config)# address-bind install FS(config)# address-bind uplink GigabitEthernet 0/1
Verification	Display the global IP-MAC binding on the device.
	FS#show address-bind Total Bind Addresses in System : 1 IP Address Binding MAC Addr --- --- 192.168.5.1 00d0.f800.0001 FS#show address-bind uplink Port State --- --- Gi0/1 Enabled Default Disabled

7.5 Monitoring

Displaying

Description	Command
Displays the IP-MAC binding on the device.	**show address-bind**
Displays the exclude port.	**show address-bind uplink**

8 Configuring Password Policy

8.1 Overview

The Password Policy is a password security function provided for local authentication of the device. It is configured to control users' login passwords and login states.

The following sections introduce password policy only.

Protocols and Standards

N/A

8.2 Features

Basic Concepts

Minimum Password Length

Administrators can set a minimum length for user passwords according to system security requirements. If the password input by a user is shorter than the minimum password length, the system does not allow the user to set this password but displays a prompt, asking the user to specify another password of an appropriate length.

Strong Password Detection

The less complex a password is, the more likely it is to crack the password. For example, a password that is the same as the corresponding account or a simple password that contains only characters or digits may be easily cracked. For the sake of security, administrators can enable the strong password detection function to ensure that the passwords set by users are highly complex. After the strong password detection function is enabled, a prompt will be displayed for the following types of passwords:

1. Passwords that are the same as corresponding accounts;
2. Simple passwords that contain characters or digits only.

Mandatory Modification of Weak Passwords

By default, weak passwords are configurable. However, after this function is enabled, weak passwords have to be modified to strong ones, or the configuration fails. Weak passwords meet at least one of the following conditions:

1. Passwords that are the same as corresponding accounts;
2. Simple passwords that contain characters or digits only.

Password Life Cycle

The password life cycle defines the validity time of a user password. When the service time of a password exceeds the life cycle, the user needs to change the password.

If the user inputs a password that has already expired during login, the system will give a prompt, indicating that the password has expired and the user needs to reset the password. If the new password input during password resetting does not meet system requirements or the new passwords consecutively input twice are not the same, the system will ask the user to input the new password once again.

Guard Against Repeated Use of Passwords

When changing the password, the user will set a new password while the old password will be recorded as the user's history records. If the new password input by the user has been used previously, the system gives an error prompt and asks the user to specify another password.

The maximum number of password history records per user can be configured. When the number of password history records of a user is greater than the maximum number configured for this user, the new password history record will overwrite the user's oldest password history record.

Storage of Encrypted Passwords

Administrators can enable the storage of encrypted passwords for security consideration. When administrators run the **show running-config** command to display configuration or run the **write** command to save configuration files, various user-set passwords are displayed in the cipher text format. If administrators disable the storage of encrypted passwords next time, the passwords already in cipher text format will not be restored to plaintext passwords.

8.3 Configuration

Configuration	Description and Command	
Configuring the Password Security Policy	Optional configuration, which is used to configure a combination of parameters related to the password security policy.	
	password policy life-cycle	Configures the password life cycle.
	password policy min-size	Configures the minimum length of user passwords.
	password policy no-repeat-times	Sets the no-repeat times of latest password configuration, so that the passwords specified in these times of latest password configuration can no longer be used in future password configuration.
	password policy strong	Enables the strong password detection function.
	password policy forced-password-modify	Enables mandatory modification of weak passwords.
	service password-encryption	Sets the storage of encrypted passwords.

Networking Requirements

- Provide a password security policy for local authentication of the device. Users can configure different password security policies to implement password security management.

Notes

- The configured password security policy is valid for global passwords (configured using the commands **enable password** and **enable secret**) and local user passwords (configured using the **username** *name* **password** *password* command). It is invalid for passwords in Line mode.

Configuration Steps

Configuring the Password Life Cycle

- Optional

- Perform this configuration on each device that requires the configuration of a password life cycle unless otherwise stated.

Configuring the Minimum Length of User Passwords

- Optional
- Perform this configuration on each device that requires a limit on the minimum length of user passwords unless otherwise stated.

Setting the No-Repeat Times of Latest Password Configuration

- Optional
- Perform this configuration on each device that requires a limit on the no-repeat times of latest password configuration unless otherwise stated.

Enabling Mandatory Modification of Weak Passwords

- Optional
- Do not perform this configuration on each device unless otherwise stated.

Enabling the Strong Password Detection Function

- Optional
- Perform this configuration on each device that requires strong password detection unless otherwise stated.

Setting the Storage of Encrypted Passwords

- Optional
- Perform this configuration on each device that requires the storage of passwords in encrypted format unless otherwise stated.

Verification

Configure a local user on the device, and configure a valid password and an invalid password for the user.

- When you configure the valid password, the device correctly adds the password.
- When you configure the invalid password, the device displays a corresponding error log.

Related Commands

Configuring the Password Life Cycle

Command Syntax	**password policy life-cycle** *days*
Parameter Description	**life-cycle** *days*: Indicates the password life cycle in the unit of days. The value range is from 1 to 65535.
Command Mode	Global configuration mode
Usage Guide	The password life cycle is used to define the validity period of user passwords. If the user logs in with a password whose service time already exceeds the life cycle, a prompt is given, asking the user to change the password.

Configuring the Minimum Length of User Passwords

Command	**password policy min-size** *length*

Syntax	
Parameter Description	**min-size** *length*: Indicates the minimum length of passwords. The value range is from 1 to 31.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the minimum length of passwords. If the minimum length of passwords is not configured, users can input a password of any length.

Setting the No-Repeat Times of Latest Password Configuration

Command Syntax	**password policy no-repeat-times** *times*
Parameter Description	**no-repeat-times** *times*: Indicates the no-repeat times of latest password configuration. The value range is from 1 to 31.
Command Mode	Global configuration mode
Usage Guide	After this function is enabled, all old passwords used in the several times of latest password configuration will be recorded as the user's password history records. If the new password input by the user has been used previously, the system gives an error prompt and the password modification fails. You can configure the maximum number of password history records per user. When the number of password history records of a user is greater than the maximum number configured for the user, the new password history record will overwrite the user's oldest password history record.

Enabling Mandatory Modification of Weak Passwords

Command Syntax	**password policy forced-password-modify**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After mandatory modification of weak passwords is enabled, users have to change their passwords if the passwords are the same as corresponding accounts or contain characters or digits only.

Enabling the Strong Password Detection Function

Command Syntax	**password policy strong**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	After the strong password detection function is enabled, a prompt is displayed for the following types of passwords: 4. Passwords that are the same as corresponding accounts; 5. Simple passwords that contain characters or digits only.

Setting the Storage of Encrypted Passwords

Command Syntax	**service password-encryption**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	Before the storage of encrypted passwords is set, all passwords used in the configuration process will be displayed and stored in plaintext format, unless the passwords are configured in cipher text format. You can enable the storage of encrypted passwords for security consideration. When you run the **show running-config** command to display configuration or run the **write** command to save configuration files, various user-set passwords are displayed in the cipher text format. If you disable the storage of encrypted passwords next time, the passwords already in cipher text format will not be restored to plaintext passwords.

Checking User-Configured Password Security Policy Information

Command Syntax	**show password policy**
Parameter Description	-
Command Mode	Privileged EXEC mode/ Global configuration mode/ Interface configuration mode
Usage Guide	Use this command to display the password security policy configured on the device.

Checking Information Such as the Default Password Dictionary and Weak Passwords Manually Set

Command Syntax	**show password policy**
Parameter Description	-
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to display information such as the default password dictionary and weak passwords manually set on the device.

Configuration Examples

The following configuration example describes configuration related to a password security policy.

Configuring Password Security Check on the Device

Typical Application	Assume that the following password security requirements arise in a network environment: 1. The minimum length of passwords is 8 characters; 2. The password life cycle is 90 days; 3. Passwords are stored and transmitted in cipher text format; 4. The number of no-repeat times of password history records is 3; 5. Passwords shall not be the same as user names, and shall not contain simple characters or digits only.

	6. Mandatory modification of weak passwords is enabled.
Configuration Steps	● Set the minimum length of passwords to 8. ● Set the password life cycle to 90 days. ● Enable the storage of encrypted passwords. ● Set the no-repeat times of password history records to 3. ● Enable the strong password detection function. ● Enable mandatory modification of weak passwords.
	FS# configure terminal FS(config)# password policy min-size 8 FS(config)# password policy life-cycle 90 FS(config)# service password-encryption FS(config)# password policy no-repeat-times 3 FS(config)# password policy strong FS(config)# password policy forced-password-modify
Verification	When you create a user and the corresponding password after configuring the password security policy, the system will perform relevant detection according to the password security policy. ● Run the **show password policy** command to display user-configured password security policy information.
	FS# show password policy Global password policy configurations: Password encryption: Enabled Password strong-check: Enabled Password forced-password-modify Enabled Password secret-dictionary-check: Enabled Password min-size: Enabled (8 characters) Password life-cycle: Enabled (90 days) Password no-repeat-times: Enabled (max history record: 3)

Common Errors

- The time configured for giving a pre-warning notice about password expiry to the user is greater than the password life cycle.

8.4 Monitoring

Displaying

Command	Function
show password policy	Displays user-configured password security policy information.

9 Configuring Port Security

9.1 Overview

Port security is used to restrict access to a port. Source MAC addresses of packets can be used to restrict the packets that enter the ports of a switch. You can set the number of static MAC addresses or the number of MAC addresses that are dynamically learned to restrict the packets that can enter the port. Ports enabled with port security are called secure ports.

9.2 Applications

Application	Description
Allowing Only Specified Hosts to Use Ports	For network security, certain ports of a device can be used only by specified hosts.

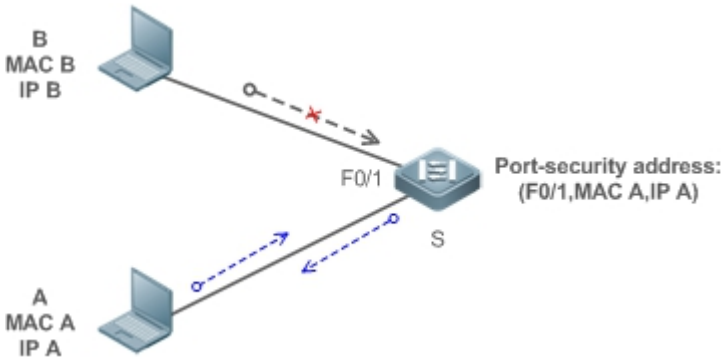
9.2.1 Allowing Only Specified Hosts to Use Ports

Scenario

In a scenario that has requirements for the network security, devices cannot be completely isolated physically. In this case, the devices need to be configured to restrict the PCs that connected to the ports of the devices.

- Only specified PCs can connect to the ports and normally use the network.
- Other PCs cannot use the network even if connected to the ports.
- After the configuration is complete, the administrator does not need to perform regular maintenance.

Figure 9- 1



Remarks	S is the access device. A is a PC that can use the port F0/1. B is an unknown PC.

Deployment

- Enable ARP Check for port F0/1 (omitted).
- Enable port security on access device S and set the violation handling mode to protect.
- Set the maximum number of secure addresses allowed by port F0/1 to 1.
- Configure a static port security address on the port F0/1.

9.3 Features

Basic Concepts

Secure Port

Ports configured with port security are called secure ports. At present, FS devices require that secure ports cannot be destination ports of mirroring.

Secure Addresses

Addresses bound to secure ports are called secure addresses. Secure addresses can be layer-2 addresses, namely MAC addresses, and can also be layer-3 addresses, namely, IP or IP+MAC addresses. When a secure address is bound to IP+MAC and a static secure MAC address is configured, the static secure MAC address must be the same as the MAC address bound to IP+MAC; otherwise, communication may fail due to inconsistency with the binding. Similarly, if only IP binding is set, only packets whose secure MAC addresses are statically configured or learned and whose source IP addresses are the bound IP address can enter the device.

Dynamic Binding

A method for a device to automatically learn addresses and convert learned addresses into secure addresses.

Static Binding

A command for manually binding secure addresses.

Aging of Secure Addresses

Regularly delete secure address records. Secure addresses for port security support aging configuration. You can specify only dynamically learned addresses for aging or specify both statically configured and dynamically learned secure addresses for aging.

Sticky MAC Address

Convert dynamically learned secure addresses into statically configured addresses. Addresses will not age. After the configurations are saved, dynamic secure addresses will not be learned again upon restart. If this function is not enabled, the secure MAC addresses dynamically learned must be learned again after device restart.

Security Violation Events

When the number of learned MAC addresses learned by a port exceeds the maximum number of secure addresses, security violation events will be triggered. You can configure the following modes for handing security violation events:

- protect: When security violation occurs, a corresponding secure port will stop learning MAC addresses and discard all packets of newly accessed users. This is the default mode for handling violation.
- restrict: When violation occurs, a port violation trap notification will be sent in addition to the behavior in the protect mode.
- shutdown: When violation occurs, the port will be disabled in addition to the behaviors in the preceding two modes.

Maximum Number of Secure Addresses

The maximum number of secure addresses indicates the total number of secure addresses statically configured and dynamically learned. When the number of secure addresses under a secure port does not reach the maximum number of secure addresses, the secure port can dynamically learn new dynamic secure addresses. When the number of secure addresses reaches the maximum number, the secure port will not learn dynamic secure addresses any longer. If new users access the secure port in this case, security violation events will occur.

Overview

Feature	Description
Enabling Port Security	Creates a secure address list for a port.
Filtering Layer-2 Users	Processes the packets received by a port from non-secure addresses.
Filtering Layer-3 Users	Checks the layer-2 and layer-3 addresses of packets passing a port.
Aging of Secure Addresses	Regularly deletes secure addresses.

9.3.1 Enabling Port Security

Enable port security for a port to restrict packets that access the network through the port.

Working Principle

When port security is enabled, the device security module will check the sources of received packets. Only packets from addresses in the secure address list can be normally forwarded; otherwise, the packets will be discarded or the port performs other violation handling behaviors.

When the port security and 802.1x are configured at the same time, packets can enter a switch only when the MAC addresses of the packets meet the static MAC address configurations of 802.1x or port security. If a port is configured with a secure channel or is bound to global IP+MAC, packets in compliance with the secure channel or bound to global IP+MAC can avoid checking of port security.

Related Configuration

Enabling Port Security for a Port

By default, port security is disabled.

You can run the **switchport port-security** command to enable or disable the port security function for a port.

You cannot enable this function for a destination port of SPAN.

Setting the Maximum Number of Secure Addresses for a Port

By default, the maximum number of secure addresses for a port is 128.

You can run the **switchport port-security maximum** command to adjust the maximum number of secure addresses for the port.

A smaller number of secure addresses mean fewer users that access the network through this port.

Setting the Mode for Handling Violation

By default, when the number of secure addresses reaches the maximum number, the secure port will discard packets from unknown addresses (none of the secure addresses of the port).

You can run the **switchport port-security violation** command to modify the violation handling mode.

Setting Secure Addresses That Can Be Dynamically Saved

By default, no secure address dynamically learned will be saved.

You can run the **switchport port-security mac-address sticky** command to save dynamically learned addresses to the configuration file. As long as the configuration file is saved, the device does not need to re-learn the secure addresses after the device is restarted.

9.3.2 Filtering Layer-2 Users

Set the secure addresses on a port to ensure that only devices whose MAC addresses are the same as the secure addresses can access the network through this port.

Working Principle

Add secure addresses for a secure port. When the number of secure addresses for a secure port does not reach the maximum number, the secure port can dynamically learn new dynamic secure addresses. When the number of secure addresses for the secure port reaches the maximum number, the secure port will not learn dynamic secure addresses any longer. The MAC addresses of users connecting to this port must be in the secure address list; otherwise, violation events will be triggered.

Related Configuration

Adding Secure Addresses for a Secure port

By default, a port dynamically learns secure addresses. If an administrator has special requirements, the administrator can manually configure secure addresses.

You can run the **switch portport-security interface** command to add or delete secure addresses for a device.

9.3.3 Filtering Layer-3 Users

Add binding of secure addresses and check layer-2 and layer-3 addresses of packets passing a port.

Working Principle

Layer-3 secure addresses support only IP binding and IP+MAC binding, and supports only static binding (not dynamic binding).

When a layer-3 secure port receives packets, layer-2 and layer-3 addresses need to be parsed. Only packets whose addresses are bound are valid packets. Other packets are considered as invalid packets and will be discarded, but no violation event will be triggered.

Related Configuration

Configuring Binding of Secure Addresses on Secure Ports

Binding of layer-3 secure addresses must be added manually.

You can run the **switchport port-security binding** command to add binding of secure addresses.

If only IP addresses are input, only IP addresses are bound. If IP addresses and MAC addresses are input, IP+MAC will be bound.

9.3.4 Aging of Secure Addresses

Regularly delete secure addresses. When this function is enabled, you need to set the maximum number of secure addresses. In this way, the device can automatically add and delete secure addresses on this port.

Working Principle

Enable the aging timer to regularly query and delete secure addresses whose aging time expires.

Related Configuration

Configuring Aging Time of Secure Addresses

By default, no secure address of a port will be aged.

You can run the **switchport port-security aging** command to enable aging time.

The **static** parameter can be used to age static addresses.

9.4 Configuration

Configuration	Description and Command	
Configuring Secure ports and Violation Handling Modes	(Mandatory) It is used to enable the port security service.	
	switchport port-security	Enables port security.
	switchport port-security maximum	Sets the maximum number of secure addresses for a port.
	switchport port-security violation	Configures the violation handling mode for port security.
	switchport port-security mac-address sticky	Configures automatic saving of dynamic addresses.
Configuring Secure Addresses on Secure Ports	(Optional) It is used to configure security filtering items.	
	switchport port-security mac-address	Configures the static secure addresses in the interface configuration mode.
	switchport port-security interface mac-address	Configures the static secure addresses in the global configuration mode.
	switchport port-security binding	Configures binding of secure addresses in the interface configuration mode.
	switchport port-security interface binding	Configures binding of secure addresses in the global configuration mode.
	switchport port-security aging	Configures aging time for all secure addresses on a port.
	switchport port-security binding-filter logging	Enables binding filter logging in the global configuration mode.

9.4.1 Configuring Secure ports and Violation Handling Modes

Configuration Effect

- Restrict the number of MAC addresses that can be learned from a port.
- Filter invalid packets based on MAC addresses, IP addresses or IP+MAC.

Notes

- A secure port cannot be the destination port of SPAN.
- The port security function cannot be configured for a DHCP Snooping trusted port.
- The port security function cannot be configured for excluded ports of global IP+MAC.
- The security function can be enabled only for wired switching ports and layer-2 AP ports in the interface configuration mode.
- The port security can work with other access control functions such as the 802.1x, global IP+MAC binding, and IP source guard. When these functions are used together, packets can enter a switch only when passing all security checks. If a security channel is configured for a port, packets in compliance with the security channel will avoid checking of the port security.

Configuration Steps

Enabling the Port Security Service

- Mandatory.
- If there is no special requirement, enable the port security service for a port on the access device.

Configuring the Maximum Number of Secure Addresses for a Port

- Optional. To adjust the maximum number of secure addresses running on a secure port, you can configure this item.
- Configure this item on a port enabled with port security.

Configuring Violation Handling Modes

- Optional. If you hope that other handling modes except discarding packets are implemented in case of violation, you can configure other handling modes.
- Configure this item on a port enabled with port security.

Saving Dynamically Learned Addresses

- Optional. If you hope that secure addresses are not re-learned after the device is restarted, you can configure this item.
- Configure this item on a port enabled with port security.

Verification

Run the command of the device for displaying the port security configurations to check whether the configurations take effect.

Related Commands

Setting Port Security

Command	**switchport port-security**
Parameter Description	-
Command Mode	Interface configuration mode
Usage Guide	By using the port security feature, you can strictly control the input of a port of a device by restricting the MAC addresses and IP addresses (optional) that access the port.

Setting the Maximum Number of Secure Addresses for a Port

Command	**switchport port-security maximum** *value*
Parameter Description	*value*: Indicates the number of secure addresses, ranging from 1 to 128.
Command Mode	Interface configuration mode
Usage Guide	If you set the maximum number to 1 and configure a secure address for this port, the workstation (whose address is the configured secure address) connected to this port will exclusively use all bandwidth of the port.

Configuring the Violation Handling Mode for Port Security

Command	**switchport port-security violation** { **protect** | **restrict** | **shutdown** }
Parameter Description	**protect**: Discards violated packets. **restrict**: Discards violated packets and send trap notifications. **shutdown**: Discards packets and disables the port.
Command Mode	Interface configuration mode
Usage Guide	-

Saving Dynamic Secure Addresses to a Configuration File

Command	**switchport port-security mac-address sticky** *mac-address* [**vlan** *vlan-id*]
Parameter Description	*mac-address*: Indicates a static secure address. *vlan-id*: Indicates the VID of a MAC address.
Command Mode	Interface configuration mode
Usage Guide	-

Configuration Example

Enabling Port Security for the Port gigabitethernet 0/3, Setting the Maximum Number of Addresses to 8, and Setting the Violation Handing Mode to protect

Configuration Steps	● Enable port security. ● Set the maximum number of secure addresses. ● Modify the violation handling mode.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface gigabitethernet 0/3 FS(config-if-GigabitEthernet 0/3)# switchport mode access FS(config-if-GigabitEthernet 0/3)# switchport port-security FS(config-if-GigabitEthernet 0/3)# switchport port-security maximum 8 FS(config-if-GigabitEthernet 0/3)# switchport port-security violation protect FS(config-if-GigabitEthernet 0/3)# switchport port-security mac-address sticky FS(config-if-GigabitEthernet 0/3)# end
Verification	Check the port security configuration on the device.
	FS# show port-security interface gigabitethernet 0/3 Interface : Gi0/3 Port Security: Enabled Port status : down Violation mode: Protect

	Maximum MAC Addresses:8 Total MAC Addresses:0 Configured MAC Addresses:0 Aging time : 0 mins SecureStatic address aging : Disabled

Common Errors

- Port security is enabled on a SPAN port.
- Port security is enabled on a DHCP trusted port.
- The configured maximum number of secure addresses is smaller than the number of existing secure addresses.

9.4.2 Configuring Secure Addresses on Secure Ports

Configuration Effect

- Allow specified users to use ports.
- Regularly update secure addresses of users.

Notes

- Sticky MAC addresses are special MAC addresses not affected by the aging mechanism. No matter dynamic or static aging is configured, sticky MAC addresses will not be aged.

Configuration Steps

Configuring Secure Addresses

- Optional. You need to manually add secure addresses for configuration.
- Configure this item on a port enabled with port security.

Configuring Binding of Secure Addresses

- Optional. You need to add layer-3 secure addresses for configuration.
- Configure this item on a port enabled with port security.

Configuring Aging Time

- Optional.
- Configure this item on a port enabled with port security.

Enabling Binding Filter Logging

- Optional.
- Enable binding filter logging in the global configuration mode.

Verification

- Run the command of the device for displaying the port security configurations to check whether the configurations take effect.

Related Commands

Adding Secure Addresses for Secure Ports in the Global Configuration Mode

Command	**switchport port-security interface** *interface-id* **mac-address** *mac-address* [**vlan** *vlan-id*]
Parameter Description	*interface-id*: Indicates the interface ID. *mac-address*: Indicates a static secure address. *vlan-id*: Indicates the VID of a MAC address.
Command Mode	Global configuration mode
Usage Guide	-

Adding Secure Addresses for Secure Ports in the Interface Configuration Mode

Command	**switchportport-security mac-address** *mac-address* [**vlan** *vlan_id*]
Parameter Description	*mac-address*: Indicates a static secure address. *vlan-id*: Indicates the VID of a MAC address.
Command Mode	Interface configuration mode
Usage Guide	-

Adding Binding of Secure Addresses for Secure Ports in the Global Configuration Mode

Command	**switchport port-security interface** *interface-id* **binding** [*mac-address* **vlan** *vlan_id*] { *ipv4-address* | *ipv6-address* }
Parameter Description	*interface-id*: Indicates the interface ID. *mac-address*: Indicates a bound source MAC address. *vlan_id*: Indicates the VID of a bound source MAC address. *ipv4-address*: Indicates a bound IPv4 address. *ipv6-address*: Indicates a bound IPv6 address.
Command Mode	Global configuration mode
Usage Guide	-

Adding Binding of Secure Addresses for Secure Ports in the Interface Configuration Mode

Command	**switchport port-security binding** [*mac-address* **vlan** *vlan_id*] { *ipv4-address* | *ipv6-address* }
Parameter Description	*mac-address*: Indicates a bound source MAC address. *vlan_id*: Indicates the VID of a bound source MAC address. *ipv4-address*: Indicates a bound IPv4 address. *ipv6-address*: Indicates a bound IPv6 address.
Command Mode	Interface configuration mode
Usage Guide	-

Configuring Aging Time for All Secure Addresses on a Port

Command	**switchport port-security aging** { **static** | **time** *time* }
Parameter	**static**: Indicates that the aging time will be applied to manually configured secure addresses and automatically learned

Description	addresses; otherwise, the aging time will be applied to only automatically learned addresses. **time** *time*: Indicates the aging time of the secure addresses on this port, ranging from 0 to 1440 minutes. If it is set to 0, it indicates that the aging function is disabled actually.
Command Mode	Interface configuration mode
Usage Guide	-

Enabling Binding Filter Logging

Command	**switchport port-security binding-filter logging** [**rate-limit** *rate*]
Parameter Description	**rate-limit** *rate*: Indicates the printing rate of binding filter logging.
Command Mode	Global configuration mode
Usage Guide	1. If you run the **switchport port-security binding-filter logging** command without configuring the *rate* parameter, binding filter logging is enabled and the default printing rate, 10logs/minute, is adopted. 2. After binding filter logging is enabled, for packets that do not comply with IP/IP-MAC binding, warmings are printed. 3. After binding filter logging is enabled, if the printing rate exceeds the configured rate, the number of suppressed packets is displayed.

Configuration Example

Configuring a Secure MAC Address 00d0.f800.073c for the Port gigabitethernet 0/3

Configuration Steps	● Enable port security. ● Add a secure address.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface gigabitethernet 0/3 FS(config-if-GigabitEthernet 0/3)# switchport mode access FS(config-if-GigabitEthernet 0/3)# switchport port-security FS(config-if-GigabitEthernet 0/3)# switchport port-security mac-address 00d0.f800.073c vlan 1 FS(config-if-GigabitEthernet 0/3)# end
Verification	Check the port security configuration on the device.
	FS# show port-security address Vlan Mac Address IP Address Type Port Remaining Age(mins) ---------- ---------- 1 00d0.f800.073c 192.168.12.202 Configured Gi0/3 8 1 00d0.f800.3cc9 192.168.12.5 Configured Gi0/1 7

Configuring a Security Binding of the IP Address 192.168.12.202 for the Port gigabitethernet 0/3

Configuration Steps	● Enable port security. ● Add a binding of the secure address.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface gigabitethernet 0/3 FS(config-if-GigabitEthernet 0/3)# switchport mode access FS(config-if-GigabitEthernet 0/3)# switchport port-security FS(config-if-GigabitEthernet 0/3)# switchport port-security binding 192.168.12.202 FS(config-if-GigabitEthernet 0/3)# end
Verification	Check the port security configuration on the device.
	FS# show port-security address Vlan Mac Address IP Address Type Port Remaining Age(mins) ---------- ---------- 1 00d0.f800.073c 192.168.12.202 Configured Gi0/3 8 1 00d0.f800.3cc9 192.168.12.5 Configured Gi0/1 7

Configuring a Secure MAC Address 00d0.f800.073c and a Security Binding of the IP Address 0000::313b:2413:955a:38f4 for the Port gigabitethernet 0/3

Configuration Steps	● Enable port security. ● Add a binding of the secure address.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface gigabitethernet 0/3 FS(config-if-GigabitEthernet 0/3)# switchport mode access FS(config-if-GigabitEthernet 0/3)# switchport port-security FS(config-if-GigabitEthernet 0/3)# switchport port-security binding 00d0.f800.073c vlan 1 0000::313b:2413:955a:38f4 FS(config-if)# end
Verification	Check the port security configuration on the device.
	FS# show port-security address Vlan Mac Address IP Address Type Port Remaining Age(mins) ---------- ---------- 1 00d0.f800.073c 192.168.12.202 Configured Gi0/3 8 1 00d0.f800.3cc9 192.168.12.5 Configured Gi0/1 7

Configuring the Aging Time of the Port gigabitethernet 0/3 to 8 Minutes, Which Is Also Applied to Statically Configured Secure Addresses

Configuration Steps	● Enable port security. ● Configure aging time.
	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface gigabitthernet 0/3 FS(config-if-GigabitEthernet 0/3)# switchport port-security aging time 8 FS(config-if-GigabitEthernet 0/3)# switchport port-security aging static FS(config-if-GigabitEthernet 0/3)# end
Verification	Check the port security configuration on the device.
	FS# show port-security gigabitethernet 0/3 Interface : Gi0/3 Port Security: Enabled Port status : down Violation mode:Shutdown Maximum MAC Addresses:8 Total MAC Addresses:0 Configured MAC Addresses:0 Aging time : 8 mins SecureStatic address aging : Enabled

9.5 Monitoring

Displaying

Description	Command
Displays all secure addresses or all secure addresses of a specified port.	**show port-security address** [**interface** *interface-id*]
Displays all bindings or all bindings of a specified port.	**show port-security binding** [**interface** *interface-id*]
Displays all valid secure addresses of ports and the security binding records of the ports.	**show port-security all**
Displays the port security configurations of an interface.	**show port-security interface** *interface-id*
Displays the statistics about port security.	**show port-security**

10 Configuring Storm Control

10.1 Overview

When a local area network (LAN) has excess broadcast data flows, multicast data flows, or unknown unicast data flows, the network speed will slow down and packet transmission will have an increased timeout probability. This situation is called a LAN storm. A storm may occur when topology protocol execution or network configuration is incorrect.

Storm control can be implemented to limit broadcast data flows, multicast data flows, or unknown unicast data flows. If the rate of data flows received by a device port is within the configured bandwidth threshold, packets-per-second threshold, or kilobits-per-second threshold, the data flows are permitted to pass through. If the rate exceeds the thresholds, excess data flows are discarded until the rate falls within the thresholds. This prevents flood data from entering the LAN causing a storm.

10.2 Applications

Application	Description
Network Attack Prevention	Enable storm control to prevent flooding.

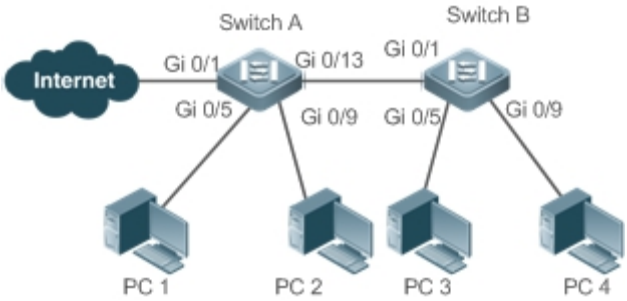
10.2.1 Network Attack Prevention

Scenario

The application requirements of network attack prevention are described as follows:

- Protect devices from flooding of broadcast packets, multicast packets, or unknown unicast packets.

Figure 10-1



Remarks	Switch A and Switch B are access devices. PC 1, PC 2, PC 3, and PC 4 are desktop computers.

Deployment

- Enable storm control on the ports of all access devices (Switch A and Switch B).

10.3 Features

Basic Concepts

Storm Control

If the rate of data flows (broadcast packets, multicast packets, or unknown unicast packets) received by a device port is within the configured bandwidth threshold, packets-per-second threshold, or kilobits-per-second threshold, the data flows are permitted to pass

through. If the rate exceeds the thresholds, excess data flows are discarded until the rate falls within the thresholds.

Storm Control Based on the Bandwidth Threshold

If the rate of data flows received by a device port is within the configured bandwidth threshold, the data flows are permitted to pass through. If the rate exceeds the threshold, excess data flows are discarded until the rate falls within the threshold.

Storm Control Based on the Packets-per-Second Threshold

If the rate of data flows received by a device port is within the configured packets-per-second threshold, the data flows are permitted to pass through. If the rate exceeds the threshold, excess data flows are discarded until the rate falls within the threshold.

Storm Control Based on the Kilobits-per-Second Threshold

If the rate of data flows received by a device port is within the configured kilobits-per-second threshold, the data flows are permitted to pass through. If the rate exceeds the threshold, excess data flows are discarded until the rate falls within the threshold.

Overview

Feature	Description
Unicast Packet Storm Control	Limits unknown unicast packets to prevent flooding.
Multicast Packet Storm Control	Limits multicast packets to prevent flooding.
Broadcast Packet Storm Control	Limits broadcast packets to prevent flooding.

10.3.1 Unicast Packet Storm Control

The unicast packet storm control feature monitors the rate of unknown unicast data flows received by a device port to limit LAN traffic and prevent flooding caused by excess data flows.

Working Principle

If the rate of unknown unicast data flows received by a device port is within the configured bandwidth threshold, packets-per-second threshold, or kilobits-per-second threshold, the data flows are permitted to pass through. If the rate exceeds the thresholds, excess data flows are discarded until the rate falls within the thresholds.

Related Configuration

Enabling Unicast Packet Storm Control on Ports

By default, unicast packet storm control is disabled on ports.

Run the **storm-control unicast** [{ **level** *percent* | **pps** *packets* | *rate-bps* }] command to enable unicast packet storm control on ports.

Run the **no storm-control unicast** or **default storm-control unicast** command to disable unicast packet storm control on ports.

The default command parameters are determined by related products.

10.3.2 Multicast Packet Storm Control

The multicast packet storm control feature monitors the rate of multicast data flows received by a device port to limit LAN traffic and prevent flooding caused by excess data flows.

Working Principle

If the rate of multicast data flows received by a device port is within the configured bandwidth threshold, packets-per-second threshold, or kilobits-per-second threshold, the data flows are permitted to pass through. If the rate exceeds the thresholds, excess data flows are discarded until the rate falls within the thresholds.

Related Configuration

Enabling Multicast Packet Storm Control on Ports

By default, multicast packet storm control is disabled on ports.

Run the **storm-control multicast** [{ **level** *percent* | **pps** *packets* | *rate-bps* }] command to enable multicast packet storm control on ports.

Run the **no storm-control multicast** or **default storm-control multicast** command to disable multicast packet storm control on ports.

The default command parameters are determined by related products.

10.3.3 Broadcast Packet Storm Control

The broadcast packet storm control feature monitors the rate of broadcast data flows received by a device port to limit LAN traffic and prevent flooding caused by excess data flows.

Working Principle

If the rate of broadcast data flows received by a device port is within the configured bandwidth threshold, packets-per-second threshold, or kilobits-per-second threshold, the data flows are permitted to pass through. If the rate exceeds the thresholds, excess data flows are discarded until the rate falls within the thresholds.

Related Configuration

Enabling Broadcast Packet Storm Control on Ports

By default, broadcast packet storm control is disabled on ports.

Run the **storm-control broadcast** [{ **level** *percent* | **pps** *packets* | *rate-bps* }] command to enable broadcast packet storm control on ports.

Run the **no storm-control broadcast** or **default storm-control broadcast** command to disable broadcast packet storm control on ports.

10.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of Storm Control	(Mandatory) It is used to enable storm control.	
	storm-control { **broadcast** | **multicast** | **unicast**} [{ **level** *percent* | **pps** *packets* | *rate-bps*}]	Enables storm control.

10.4.1 Configuring Basic Functions of Storm Control

Configuration Effect

- Prevent flooding caused by excess broadcast packets, multicast packets, and unknown unicast packets.

Notes

- When you run a command (for example, **storm-control unicas**t) to enable storm control, if you do not set the parameters, the default values are used.

Configuration Steps

Enabling Unicast Packet Storm Control

- Mandatory.
- Enable unicast packet storm control on every device unless otherwise specified.

Enabling Multicast Packet Storm Control

- Mandatory.
- Enable multicast packet storm control on every device unless otherwise specified.

Enabling Broadcast Packet Storm Control

- Mandatory.
- Enable broadcast packet storm control on every device unless otherwise specified.

Verification

- Run the **show storm-control** command to check whether the configuration is successful.

Related Commands

Enabling Unicast Packet Storm Control

Command	**storm-control unicast** [{ **level** *percent* | **pps** *packets* | *rate-bps*}]
Parameter Description	**level** *percent*: Indicates the bandwidth percentage. **pps** *packets*: Indicates the number of packets per second. *rate-bps*: Indicates the packet rate.
Command Mode	Interface configuration mode
Usage Guide	Storm control can be enabled only on switch ports.

Enabling Multicast Packet Storm Control

Command	**storm-control multicast** [{ **level** *percent* | **pps** *packets* | *rate-bps* }]
Parameter Description	**level** *percent*: Indicates the bandwidth percentage. **pps** *packets*: Indicates the number of packets per second. *rate-bps*: Indicates the packet rate.
Command Mode	Interface configuration mode
Usage Guide	Storm control can be enabled only on switch ports.

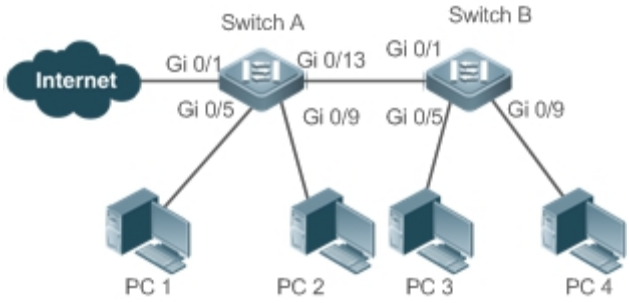
Enabling Broadcast Packet Storm Control

Command	**storm-control broadcast** [{ **level** *percent* | **pps** *packets* | *rate-bps* }]

Parameter Description	**level** *percent*: Indicates the bandwidth percentage. **pps** *packets*: Indicates the number of packets per second. *rate-bps*: Indicates the packet rate.
Command Mode	Interface configuration mode
Usage Guide	Storm control can be enabled only on switch ports.

Configuration Example

Enabling Storm Control on Devices

Scenario	
Figure 10-2	
Configuration Step	● Enable storm control on Switch A and Switch B.
Switch A	FS(config)#interface range gigabitEthernet 0/5,0/9,0/13 FS(config-if-range)#storm-control broadcast FS(config-if-range)#storm-control multicast FS(config-if-range)#storm-control unicast
Switch B	FS(config)#interface range gigabitEthernet 0/1,0/5,0/9 FS(config-if-range)#storm-control broadcast FS(config-if-range)#storm-control multicast FS(config-if-range)#storm-control unicast
Verification	Check whether storm control is enabled on Switch A and Switch B.
Switch A	FS# sho storm-control Interface Broadcast Control Multicast Control Unicast Control Action ---------- ---------- ---------- ---------- -------- GigabitEthernet 0/1 Disabled Disabled Disabled none GigabitEthernet 0/5 default default default none GigabitEthernet 0/9 default default default none GigabitEthernet 0/13 default default default none

Switch B	FS#sho storm-control Interface Broadcast Control Multicast Control Unicast Control Action ---------- ---------- ---------- ---------- -------- GigabitEthernet 0/1 default default default none GigabitEthernet 0/5 default default default none GigabitEthernet 0/9 default default default none

10.5 Monitoring

Displaying

Description	Command
Displays storm control information.	**show storm-control** [*interface-type interface-number*]

11 Configuring SSH

11.1 Overview

Secure Shell (SSH) connection is similar to a Telnet connection except that all data transmitted over SSH is encrypted. When a user in an insecure network environment logs into a device remotely, SSH helps ensure information security and powerful authentication, protecting the device against attacks such as IP address spoofing and plain-text password interception.

An SSH-capable device can be connected to multiple SSH clients. In addition, the device can also function as an SSH client, and allows users to set up an SSH connection with a SSH-server device. In this way, the local device can safely log in to a remote device through SSH to implement management.

Currently, a device can work as either the SSH server or an SSH client, supporting SSHv1 and SSHv2 versions. FS SSH service supports both IPv4 and IPv6.

Unless otherwise specified, SSH in this document refers to SSHv2.

Protocols and Standards

- RFC 4251: The Secure Shell (SSH) Protocol Architecture
- RFC 4252: The Secure Shell (SSH) Authentication Protocol
- RFC 4253: The Secure Shell (SSH) Transport Layer Protocol
- RFC 4254: The Secure Shell (SSH) Connection Protocol
- RFC 4419: Diffie-Hellman Group Exchange for the Secure Shell (SSH) Transport Layer Protocol
- RFC 4716: The Secure Shell (SSH) Public Key File Format
- RFC 4819: Secure Shell Public Key Subsystem
- RFC 3526: More Modular Exponential (MODP) Diffie-Hellman groups for Internet Key Exchange (IKE)
- RFC 2409: The Internet Key Exchange (IKE)
- RFC 1950: ZLIB Compressed Data Format Specification version 3.3
- draft-ietf-secsh-filexfer-05: SSH File Transfer Protocol
- draft-ylonen-ssh-protocol-00: The version of the SSH Remote Login Protocol is 1.5. Comware implements the SSH server functions, but not the SSH client functions.

11.2 Applications

Application	Description
SSH Device Management	Use SSH to manage devices.
SSH Local Line Authentication	Use the local line password authentication for SSH user authentication.
SSH AAA Authentication	Use the authentication, authorization and accounting (AAA) mode for SSH user authentication.
SSH Public Key Authentication	Use the public key authentication for SSH user authentication.
SSH File Transfer	Use the Secure Copy (SCP) commands on the client to exchange data with the SSH server.
SSH Client Application	Use the SSH client to safely log in to a remote device for management.

11.2.1 SSH Device Management

Scenario

You can use SSH to manage devices on the precondition that the SSH server function is enabled. By default, this function is disabled. The Telnet component that comes with the Windows system does not support SSH. Therefore, a third-party client software must be used. Currently, well-compatible software includes PuTTY, Linux, and SecureCRT. The following takes the PuTTY as an example to introduce the configurations of the SSH client. Figure 11- 1 shows the network topology.

Figure 11- 1 Networking Topology of SSH Device Management

Deployment

Configure the SSH client as follows:

- Start the PuTTY software.
- On the **Session** option tab of PuTTY, type in the host IP address of the SSH server and SSH port number **22**, and select the connection type **SSH**.
- On the **SSH** option tab of PuTTY, select the preferred SSH protocol version **2**.
- On the **SSH authentication** option tab of PuTTY, select the authentication method **Attempt "keyboard-interactive" auth**.
- Click **Open** to connect to the SSH server.
- Type in the correct user name and password to enter the terminal login interface.

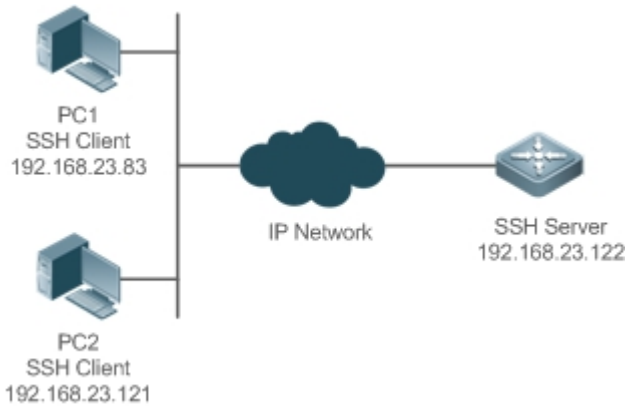
11.2.2 SSH Local Line Authentication

Scenario

SSH clients can use the local line password authentication mode, as shown in Figure 11- 2.To ensure security of data exchange, PC 1 and PC 2 function as the SSH clients, and use the SSH protocol to log in to the network device where the SSH server function is enabled. The requirements are as follows:

- SSH users use the local line password authentication mode.
- Five lines, including Line 0 to Line 4, are activated concurrently. The login password is "passzero" for Line 0 and "pass" for the remaining lines. Any user name can be used.

Figure 11- 2 Networking Topology of SSH Local Line Password Authentication



Deployment

- Configure the SSH server as follows:

1. Enable the SSH server function globally. By default, the SSH server supports two SSH versions: SSHv1 and SSHv2.
2. Configure the key. With this key, the SSH server decrypts the encrypted password received from the SSH clients, compares the decrypted plain text with the password stored on the server, and returns a message indicating the successful or unsuccessful authentication. SSHv1 uses an RSA key, whereas SSHv2 adopts an RSA or DSA key.
3. Configure the IP address of the FastEthernet 0/1 interface on the SSH server. The SSH client is connected to the SSH server using this IP address. The routes from the SSH clients to the SSH server are reachable.

- Configure the SSH client as follows:

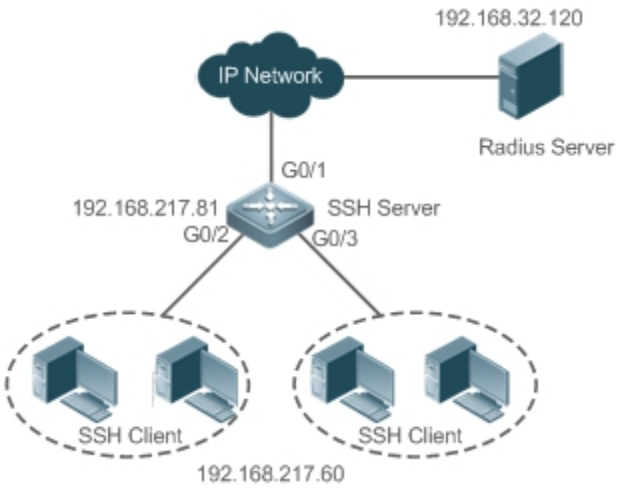
1. Diversified SSH client software is available, including PuTTY,Linux, and OpenSSH. This document takes PuTTY as an example to explain the method for configuring the SSH clients.
2. Open the PuTTY connection tab, and select SSHv1 for authenticated login. (The method is similar if SSHv2 is selected.)
3. Set the IP address and connected port ID of the SSH server. As shown in the network topology, the IP address of the server is 192.168.23.122, and the port ID is 22. Click **Open** to start the connection. As the current authentication mode does not require a user name, you can type in any user name, but cannot be null. (In this example, the user name is "anyname".)

11.2.3 SSH AAA Authentication

Scenario

SSH users can use the AAA authentication mode for user authentication, as shown in Figure 11- 3.To ensure security of data exchange, the PCs function as the SSH clients, and uses the SSH protocol to log in to the network device where the SSH server is enabled. To better perform security management, the AAA authentication mode is used for user login on the SSH clients. Two authentication methods, including Radius server authentication and local authentication, are provided in the AAA authentication method list to ensure reliability. The Radius server authentication method is preferred. If the Radius server does not respond, it turns to the local authentication.

Figure 11- 3 Networking Topology of SSH AAA Authentication



Deployment

- The routes from the SSH clients to the SSH server are reachable, and the route from the SSH server to the Radius server is also reachable.
- Configure the SSH server on the network device that functions as an SSH client.
- Configure the AAA parameters on the network device. When the AAA authentication mode is used, method lists are created to define the identity authentication and types, and applied to a specified service or interface.

11.2.4 SSH Public Key Authentication

Scenario

SSH clients can use the public keys for authentication, and the public key algorithm can be RSA or DSA, as shown in Figure 11- 4.SSH is configured on the client so that a secure connection is set up between the SSH client and the SSH server.

Figure 11- 4 Network Topology for Public Key Authentication of SSH Users

Deployment

- To implement public key authentication for the client, generate a key pair (RSA or DSA) on the client, configure the public key on the SSH server, and select the public key authentication mode.
- After the key is generated on the client, the SSH server will copy the file of the public key from the client to the flash and associates the file with the SSH user name. Each user can be associated with one RSA public key and one DSA public key.

11.2.5 SSH File Transfer

Scenario

The SCP service is enabled on the server, and SCP commands are used on the client to transfer data to the server, as shown in Figure 11- 5.

Figure 11- 5 Networking Topology of SSH File Transfer

Deployment

- Enable the SCP service on the server.
- On the client, use SCP commands to upload files to the server, or download files from the server.

11.2.6 SSH Client Application

Scenario

The SSH service is enabled on a remote SSH server, and the **ssh** command is used on the local client to set up an SSH connection with the server for secure data transmission, as shown in Figure 11- 6.

Figure 11- 6 Networking Topology of SSH Client Application

Deployment

- Enable the SSH service on the server.
- On the client, run the **ssh** command to set up an SSH connection with the server for secure data transmission.

11.3 Features

Basic Concepts

User Authentication Mechanism

- Password authentication

During the password authentication, a client sends a user authentication request and encrypted user name and password to the server. The server decrypts the received information, compares the decrypted information with those stored on the server, and then returns a message indicating the successful or unsuccessful authentication.

- Public key authentication

During the public key authentication, digital signature algorithms, such as RSA and DSA, are used to authenticate a client. The client sends a public key authentication request to the server. This request contains information including the user name, public key, and public key algorithm. On receiving the request, the server checks whether the public key is correct. If wrong, the server directly sends an authentication failure message. If right, the server performs digital signature authentication on the client, and returns a message indicating the successful or unsuccessful authentication.

Public key authentication is applicable only to the SSHv2 clients.

SSH Communication

To ensure secure communication, interaction between an SSH server and an SSH client undergoes the following seven stages:

- Connection setup

The server listens on Port 22 to the connection request from the client. After originating a socket initial connection request, the client sets up a TCP socket connection with the server.

- Version negotiation

If the connection is set up successfully, the server sends a version negotiation packet to the client. On receiving the packet, the client analyzes the packet and returns a selected protocol version to the server. The server analyzes the received information to determine whether version negotiation is successful.

- Key exchange and algorithm negotiation

If version negotiation is successful, key exchange and the algorithm negotiation are performed. The server and the client exchange the algorithm negotiation packet with each other, and determine the final algorithm based on their capacity. In addition, the server and the client work together to generate a session key and a session ID according to the key exchange algorithm and host key, which will be applied to subsequent user authentication, data encryption, and data decryption.

- User authentication

After the encrypted channel is set up, the client sends an authentication request to the server. The server repeatedly conducts authentication for the client until the authentication succeeds or the server shuts down the connection because the maximum number of authentication attempts is reached.

- Session request

After the successful authentication, the client sends a session request to the server. The server waits and processes the client request. After the session request is successfully processed, SSH enters the session interaction stage.

- Session interaction

After the session request is successfully processed, SSH enters the session interaction stage. Encrypted data can be transmitted and processed in both directions. The client sends a command to be executed to the client. The server decrypts, analyzes, and processes the received command, and then sends the encrypted execution result to the client. The client decrypts the execution result.

- Session ending

When the interaction between the server and the client is terminated, the socket connection disconnects, and the session ends.

Overview

Feature	Description
SSH Server	Enable the SSH server function on a network device, and you can set up a secure connection with the network device through the SSH client.
SCP Service	After the SCP service is enabled, you can directly download files from the network device and upload local files to the network device. In addition, all interactive data is encrypted, featuring authentication and security.
SSH Client	You can use the SSH client on the device to set up a secure connection with the SSH server on a network device.

11.3.1 SSH Server

Enable the SSH server function on a network device, and you can set up a secure connection with the network device through the SSH client. You can also shut down the SSH server function to disconnect from all SSH clients.

Working Principle

For details about the working principle of the SSH server, see the "SSH Communication" in "Basic Concepts." In practice, after enabling the SSH server function, you can configure the following parameters according to the application requirements:

- Version: Configure the SSH version as SSHv1 orSSHv2 to connect SSH clients.
- Authentication timeout: The SSH server starts the timer after receiving a user connection request. The SSH server is disconnected from the client either when the authentication succeeds or when the authentication timeout is reached.
- Maximum number of authentication retries: The SSH server starts authenticating the client after receiving its connection request. If authentication does not succeed when the maximum number of user authentication retries is reached, a message is sent, indicating the authentication failure.
- Public key authentication: The public key algorithm can be RSA or DSA. It provides a secure connection between the client and the server. The public key file on the client is associated with the user name. In addition, the public key authentication mode is configured on the client, and the corresponding private key file is specified. In this way, when the client attempts to log in to the server, public key authentication can be implemented to set up a secure connection.

Related Configuration

Enabling the SSH Server

By default, the SSH server is disabled.

In global configuration mode, run the [**no**] **enable service ssh-server** command to enable or disable the SSH server.

To generate the SSH key, you also need to enable the SSH server.

Specifying the SSH Version

By default, the SSH server supports both SSHv1 and SSHv2, connecting either SSHv1 clients or SSHv2 clients.

Run the **ip ssh version** command to configure the SSH version supported by the SSH server.

If only SSHv1 or SSHv2 is configured, only the SSH client of the configured version can be connected to the SSH server.

Configuring the SSH Authentication Timeout

By default, the user authentication timeout is 120s.

Run the **ip ssh time-out** command to configure the user authentication timeout of the SSH server. Use the **no** form of the command to restore the default timeout. The SSH server starts the timer after receiving a user connection request. If authentication does not succeed before the timeout is reached, authentication times out and fails.

Configuring the Maximum Number of SSH Authentication Retries

By default, the maximum number of user authentication retries is 3.

Run the **ip ssh authentication-retries** command to configure the maximum number of user authentication retries on the SSH server. Use the **no** form of the command to restore the default number of user authentication retries. If authentication still does not succeed when the maximum number of user authentication retries is reached, user authentication fails.

Specifying the SSH Encryption Mode

By default, the encryption mode supported by the SSH server is Compatible, that is, supporting cipher block chaining (CBC), counter (CTR) and other encryption modes.

Run the **ip ssh cipher-mode** command to configure the encryption mode supported by the SSH server. Use the **no** form of the command to restore the default encryption mode supported by the SSH server.

Specifying the SSH Message Authentication Algorithm

By default, the message authentication algorithms supported by the SSH server are as follows: (1) For the SSHv1, no algorithm is supported; (2) For the SSHv2, four algorithms, including MD5,SHA1,SHA1-96, and MD5-96, are supported.

Run the **ip ssh hmac-algorithm** command to configure the message authentication algorithm supported by the SSH server. Use the **no** form of the command to restore the default message authentication algorithm supported by the SSH server.

Setting A Monitoring Port ID for the SSH Server

The default port ID is 22.

Run the **ip ssh port** command to set a monitoring port ID for the SSH server. Use either the **no ip ssh port** command or the **ip ssh port 22** command to restore the default setting.

Enabling the Public Key Authentication on the SSH Server

Run the **ip ssh peer** command to associate the public key file on the client with the user name. When the client is authenticated upon login, a public key file is specified based on the user name.

11.3.2 SCP Service

The SSH server provides the SCP service to implement secure file transfer between the server and the client.

Working Principle

- SCP is a protocol that supports online file transfer. It runs on Port 22 based on the BSC RCP protocol, whereas RCP provides the encryption and authentication functions based on the SSH protocol. RCP implements file transfer, and SSH implements authentication and encryption.
- Assume that the SCP service is enabled on the server. When you use an SCP client to upload or download files, the SCP client first analyzes the command parameters, sets up a connection with a remote server, and starts another SCP process based on this connection. This process may run in source or sink mode. (The process running in source mode is the data provider. The process running in sink mode is the destination of data.) The process running in source mode reads and sends files to the peer end through the SSH connection. The process running in sink mode receives files through the SSH connection.

Related Configuration

Enabling the SCP Server

By default, the SCP server function is disabled.

Run the **ip scp server enable** command to enable SCP server function on a network device.

11.3.3 SSH Client

The SSH client is used to set up a secure connection with a remote network device on which the SSH server runs.

Working Principle

For details about the working principle of the SSH client, see the "SSH Communication" in "Basic Concepts."

Related Configuration

Specifying the Source Interface of the SSH Client

By default, the source address of SSH packets is searched based on the destination address.

Run the **ip ssh source-interface** *interface-name* command to specify the source interface of the SSH client.

Establishing a Session with the SSH Server

Run the **ssh** command to log in to a remote device that supports the SSH Server

Recovering an Established SSH Session

- Run the **ssh-session** *session-id* command to recover an established SSH session.

Disconnecting a Suspended SSH Session

- Run the **disconnect ssh-session** *session-id* command to disconnect a specified SSH session.

11.4 Configuration

Configuration	Description and Command	
Configuring the SSH Server	It is mandatory to enable the SSH server.	
	enable service ssh-server	Enables the SSH server.
	disconnect ssh[**vty**] *session-id*	Disconnects an established SSH session.
	crypto key generate {**rsa**|**dsa**}	Generates an SSH key.
	ip ssh version {**1**|**2**}	Specifies the SSH version.
	ip ssh time-out *time*	Configures the SSH authentication timeout.
	ip ssh authentication-retries *retry times*	Configures the maximum number of SSH authentication retries.
	Ip ssh cipher-mode{**cbc** | **ctr** | **others** }	Specifies the SSH encryption mode.
	Ip ssh hmac-algorithm{**md5** | **md5-96** | **sha1** | **sha1-96**}	Specifies the SSH message authentication algorithm.
	ip ssh peer *test* **public-key rsa flash** *:rsa.pub*	Associates an RSA public key file with a user.
	ip ssh peer *test* **public-key dsa flash**:*dsa.pub*	Associates a DSA public key file with a user.
	ip ssh port *port*	Sets a monitoring port ID for the SSH server.
Configuring the SCP Service	Mandatory.	
	ip scp server enable	Enables the SCP server.
Configuring the SSH Client	(Optional)It is used to set up a secure connection with a remote network device that supports the SSH server.	
	ip ssh source-interface *interface-name*	Specifies the source interface of the SSH client.

Configuration	Description and Command	
	ssh [**-v** { **1** | **2** }] [**-c** { **3des** | **aes128-cbc** | **aes192-cbc** | **aes256-cbc** }] [**-l** *username*] [**-m** { **hmac-md5-96** | **hmac-md5-128** | **hmac-sha1-96** | **hmac-sha1-160** }] [**-p** *port-num*] { *ip-addr* | *hostname* }[**/source** { **ip** *A.B.C.D*| **ipv6** *X:X:X:X::X* | **interface** *interface-name* }]	Establishes an encrypted session with a remote network device.

11.4.1 Configuring the SSH Server

Configuration Effect

- Enable the SSH server function on a network device so that you can set up a secure connection with a remote network device through the SSH client. All interactive data is encrypted before transmitted, featuring authentication and security.
- You can use diversified SSH user authentications modes, including local line password authentication, AAA authentication, and public key authentication.
- You can generate or delete an SSH key.
- You can specify the SSH version.
- You can configure the SSH authentication timeout.
- You can configure the maximum number of SSH authentication retries.
- You can specify the SSH encryption mode.
- You can specify the SSH message authentication algorithm.
- You can specify ACL filtering of the SSH server.

Notes

- The precondition of configuring a device as the SSH server is that communication is smooth on the network that the device resides, and the administrator can access the device management interface to configure related parameters.
- The **no crypto key generate** command does not exist. You need to run the **crypto key zeroize** command to delete a key.
- The SSH module does not support hot standby. Therefore, for products that supports hot standby on the supervisor modules, if no SSH key file exist on the new active module after failover, you must run the **crypto key generate** command to re-generate a key before using SSH.

Configuration Steps

Enabling the SSH Server

- Mandatory.
- By default, the SSH server is disabled. In global configuration mode, enable the SSH server and generate an SSH key so that the SSH server state changes to ENABLE.

Specifying the SSH Version

- Optional.
- By default, the SSH server supports SSHv1 and SSHv2, connecting either SSHv1 or SSHv2clients. If only SSHv1 or SSHv2 is

configured, only the SSH client of the configured version can be connected to the SSH server.

Configuring the SSH Authentication Timeout

- Optional.
- By default, the SSH authentication timeout is 120s. You can configure the user authentication timeout as required. The value ranges from 1 to 120. The unit is second.

Configuring the Maximum Number of SSH Authentication Retries

- Optional.
- Configure the maximum number of SSH authentication retries to prevent illegal behaviors such as malicious guessing. By default, the maximum number of SSH authentication retries is 3, that is, a user is allowed to enter the user name and password three times for authentication. You can configure the maximum number of retries as required. The value ranges from 0 to 5.

Specifying the SSH Encryption Mode

- Optional.
- Specify the encryption mode supported by the SSH server. By default, the encryption mode supported by the SSH server is Compatible, that is, supporting CBC, CTR and other encryption modes.

Specifying the SSH Message Authentication Algorithm

- Optional.
- Specify the message authentication algorithm supported by the SSH server. By default, the message authentication algorithms supported by the SSH server are as follows: (1) For the SSHv1, no algorithm is supported; (2) For the SSHv2, four algorithms, including MD5, SHA1, SHA1-96, and MD5-96, are supported.

Setting ACL Filtering of the SSH Server

- Optional.
- Set ACL filtering of the SSH server. By default, ACL filtering is not performed for all connections to the SSH server. According to needs, set ACL filtering to perform for all connections to the SSH server.

Enabling the Public Key Authentication for SSH Users

- Optional.
- Only SSHv2 supports authentication based on the public key. This configuration associates a public key file on the client with a user name. When a client is authenticated upon login, a public key file is specified based on the user name.

Verification

- Run the **show ip ssh** command to display the current SSH version, authentication timeout, and maximum number of authentication retries of the SSH server.
- Run the **show crypto key mypubkey** command to display the public information of the public key to verify whether the key has been generated.
- Configure the public key authentication login mode on the SSH client and specify the private key file. Check whether you can successfully log in to the SSH server from the SSH client. If yes, the public key file on the client is successfully associated with the user name, and public key authentication succeeds.

Related Commands

Enabling the SSH Server

Command	**enable service ssh-server**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	To disable the SSH server, run the **no enable service ssh-server** command in global configuration mode. After this command is executed, the SSH server state changes to DISABLE.

Disconnecting an Established SSH Session

Command	**disconnect ssh**[**vty**] *session-id*
Parameter Description	**vty:** Indicates an established virtual teletype terminal (VTY) session. *session-id*: Indicates the ID of the established SSH session. The value ranges from 0 to 35.
Command Mode	Privileged EXEC mode
Usage Guide	Specify an SSH session ID to disconnect the established SSH session. Alternatively, specify a VTY session ID to disconnect a specified SSH session. Only an SSH session can be disconnected.

Generating an SSH Key

Command	**crypto key generate** {**rsa**|**dsa**}
Parameter Description	**rsa:** Generates an RSA key. **dsa:** Generates a DSA key.
Command Mode	Global configuration mode
Usage Guide	The **no crypto key generate** command does not exist. You need to run the **crypto key zeroize** command to delete a key. SSHv1 uses an RSA key, whereas SSHv2 uses an RSA or DSA key. If an RSA key is generated, both SSHv1 and SSHv2 are supported. If only a DSA key is generated, only SSHv2 can use the key.

Specifying the SSH Version

Command	**ip ssh version** {**1**|**2**}
Parameter Description	**1:** Indicates that the SSH server only receives the connection requests sent by SSHv1 clients. **2:** Indicates that the SSH server only receives the connection requests sent by SSHv2 clients.
Command Mode	Global configuration mode
Usage Guide	Run the **no ip ssh version** command to restore the default settings. By default, the SSH server supports both SSHv1 and SSHv2.

Configuring the SSH Authentication Timeout

Command	**ip ssh time-out** *time*
Parameter	*time***:** Indicates the SSH authentication timeout. The value ranges from 1 to 120. The unit is second.

Description	
Command Mode	Global configuration mode
Usage Guide	Run the **no ip ssh time-out** command to restore the default SSH authentication timeout, which is 120s.

Configuring the Maximum Number of SSH Authentication Retries

Command	**ip ssh authentication-retries** *retry times*
Parameter Description	*retry times***:** Indicates the maximum number of user authentication retries. The value ranges from 0 to 5.
Command Mode	Global configuration mode
Usage Guide	Run the **no ip ssh authentication-retries** command to restore the default number of user authentication retries, which is 3.

Specifying the SSH Encryption Mode

Command	**ip ssh cipher-mode**{**cbc** | **ctr** | **others** }
Parameter Description	**cbc:** Sets the encryption mode supported by the SSH server to the CBC mode. Corresponding algorithms include DES-CBC,3DES-CBC,AES-128-CBC,AES-192-CBC,AES-256-CBC, and Blowfish-CBC. **ctr:** Sets the encryption mode supported by the SSH server to the CTR mode. Corresponding algorithms include AES128-CTR, AES192-CTR, and AES256-CTR. **others:** Sets the encryption mode supported by the SSH server to others. The corresponding algorithm is RC4.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the encryption mode supported by the SSH server. On FS devices, the SSHv1 server supports the DES-CBC, 3DES-CBC, and Blowfish-CBC encryption algorithms; the SSHv2 server supports the AES128-CTR, AES192-CTR, AES256-CTR, DES-CBC, 3DES-CBC, AES-128-CBC, AES-192-CBC, AES-256-CBC, Blowfish-CBC, and RC4 encryption algorithms. These algorithms can be grouped into three encryption modes: CBC, CTR, and others. As the cryptography continuously develops, it is approved that encryption algorithms in the CBC and others modes can be decrypted in a limited period of time. Therefore, organizations or companies that have high security requirements can set the encryption mode supported by the SSH server to CTR to increase the security level of the SSH server.

Specifying the SSH Message Authentication Algorithm

Command	**ip ssh hmac-algorithm**{**md5** | **md5-96** | **sha1** | **sha1-96**}
Parameter Description	**md5:** Indicates that the message authentication algorithm supported by the SSH server is MD5. **md5-96:** Indicates that the message authentication algorithm supported by the SSH server is MD5-96. **sha1:** Indicates that the message authentication algorithm supported by the SSH server is SHA1. **sha1-96:** Indicates that the message authentication algorithm supported by the SSH server is SHA1-96.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the message authentication algorithm supported by the SSH server. On FS devices, the SSHv1 server does support any message authentication algorithm; the SSHv2 server supports the MD5**,** SHA1**,** SHA1-96**, and** MD5-96 message authentication algorithms. You can select message authentication

	algorithms supported by the SSH server as required.

Setting A Monitoring Port ID for the SSH Server

Command	**ip ssh port** *port*
Parameter Description	*port*: Indicates the monitoring port ID of the SSH server. The value ranges from 1025 to 65535.
Command Mode	Global configuration mode
Usage Guide	Use either the **no ip ssh port** or the **ip ssh port 22** to restore the monitoring port ID of the SSH server to the default value.

Configuring RSA Public Key Authentication

Command	**ip ssh peer** *test* **public-key rsaflash**:*rsa.pub*
Parameter Description	*test***:** Indicates the user name. **rsa:** Indicates that the public key type is RSA. *rsa.pub*: Indicates the name of a public key file.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the RSA public key file associated with user *test*. Only SSHv2 supports authentication based on the public key. This command associates the public key file on the client with the user name. When the client is authenticated upon login, a public key file is specified based on the user name.

Configuring DSA Public Key Authentication

Command	**ip ssh peer** *test* **public-key dsaflash**:*dsa.pub*
Parameter Description	*test***:** Indicates the user name. **dsa**: Indicates that the public key type is DSA. *dsa.pub***:** Indicates the name of a public key file.
Command Mode	Global configuration mode
Usage Guide	This command is used to configure the DSA key file associated with user **test**. Only SSHv2 supports authentication based on the public key. This command associates the public key file on the client with the user name. When the client is authenticated upon login, a public key file is specified based on the user name.

Configuration Example

The following configuration examples describe only configurations related to SSH.

Generating a Public Key on the SSH Server

Configuration Steps	● Run the **crypto key generate { rsa | dsa }** command to generate a RSA public key for the server.

SSH Server	FS#configure terminal FS(config)# crypto key generate rsa Choose the size of the rsa key modulus in the range of 512 to 2048 and the size of the dsa key modulus in the range of 360 to 2048 for your Signature Keys. Choosing a key modulus greater than 512 may take a few minutes. How many bits in the modulus [512]: ● If the generation of the RSA key is successful, the following information is displayed: % Generating 512 bit RSA1 keys ...[ok] % Generating 512 bit RSA keys ...[ok] ● If the generation of the RSA key fails, the following information is displayed: % Generating 512 bit RSA1 keys ...[fail] % Generating 512 bit RSA keys ...[fail]
Verification	● Run the **show crypto key mypubkey rsa** command to display the public information about the RSA key. If the public information about the RSA key exists, the RSA key has been generated.
SSH Server	FS(config)#show crypto key mypubkey rsa % Key pair was generated at: 1:49:47 UTC Jan 4 2013 Key name: RSA1 private Usage: SSH Purpose Key Key is not exportable. Key Data: AAAAAwEA AQAAAHJM 6izXt1pp rUSOEGZ/ UhFpRRrW nngP4BU7 mG836apf jajSYwcU 8O3LojHL ayJ8G4pG 7j4T4ZSf FKg09kfr 92JpRNHQ gbwaPc5/ 9UnTtX9t qFIKDj1j 0dKBcCfN tr0r/CT+ cs5tlGKV S0lCGifz oB+pYaE= % Key pair was generated at: 1:49:47 UTC Jan 4 2013 Key name: RSA private Usage: SSH Purpose Key Key is not exportable. Key Data: AAAAAwEAAQAAAHJfLwKnzOgO F3RlKhTN /7PmQYoE v0a2VXTX 8ZCa7Sll EghLDLJc w3T5JQXk Rr3iBD5s b1EeOL4b 21ykZt/u UetQ0Q80 slSglfZ9 8o5No3Zz MPM0LnQR

	G4c7/28+ GOHzYkTk 4IiQuTIL HRgtbyEYXCFaaxU=

Specifying the SSH Version

Configuration Steps	● Run the **ip ssh version** { **1 | 2** } command to set the version supported by the SSH server to SSHv2.
SSH Server	FS#configure terminal FS(config)#ip ssh version 2
Verification	● Run the **show ip ssh** command to display the SSH version currently supported by the SSH server.
SSH Server	FS(config)#show ip ssh SSH Enable - version 2.0 Authentication timeout: 120 secs Authentication retries: 3 SSH SCP Server: disabled

Configuring the SSH Authentication Timeout

Configuration Steps	● Run the **ip ssh time-out** *time* command to set the SSH authentication timeout to 100s.
SSH Server	FS#configure terminal FS(config)#ip sshtime-out100
Verification	● Run the **show ip ssh** command to display the configured SSH authentication timeout.
SSH Server	FS(config)#show ip ssh SSH Enable - version 2.0 Authentication timeout: 100 secs Authentication retries: 3 SSH SCP Server: disabled

Configuring the Maximum Number of SSH Authentication Retries

Configuration Steps	● Run the **ip ssh authentication-retries** *retry times* command to set the maximum number of user authentication retries on the SSH server to 2.
SSH Server	FS#configure terminal FS(config)#ip ssh authentication-retries 2

Verification	● Run the **show ip ssh** command to display the configured maximum number of authentication retries.
SSH Server	FS(config)#show ip ssh SSH Enable - version 2.0 Authentication timeout: 100 secs Authentication retries: 3 SSH SCP Server: disabled

Specifying the SSH Encryption Mode

Configuration Steps	● Run the **ip ssh cipher-mode** {**cbc** | **ctr** | **others** }command to set the encryption mode supported by the SSH server to CTR.
SSH Server	FS#configure terminal FS(config)# ip ssh cipher-mode ctr
Verification	● Select the CTR encryption mode on the SSH client, and verify whether you can successfully log in to the SSH server from the SSH client.

Specifying the SSH Message Authentication Algorithm

Configuration Steps	● Run the **ip ssh hmac-algorithm** {**md5** | **md5-96** | **sha1** | **sha1-96** } command to set the message authentication algorithm supported by the SSH server to SHA1.
SSH Server	FS#configure terminal FS(config)# ip ssh hmac-algorithmsha1
Verification	● Select the SHA1 message authentication algorithm on the SSH client, and verify whether you can successfully log in to the SSH server from the SSH client.

Setting A Monitoring Port ID for the SSH Server

Configuration Steps	● Run the **ip ssh port** *port* command to set a monitoring port ID to 10000.
SSH Server	FS# configure terminal FS(config)# ip ssh port 10000
Verification	● Run the show ip ssh command to display information about a monitoring port ID for the SSH server. FS(config)#show ip ssh SSH Enable - version 2.0 SSH Port: 10000 SSH Cipher Mode: cbc,ctr,others SSH HMAC Algorithm: md5-96,md5,sha1-96,sha1

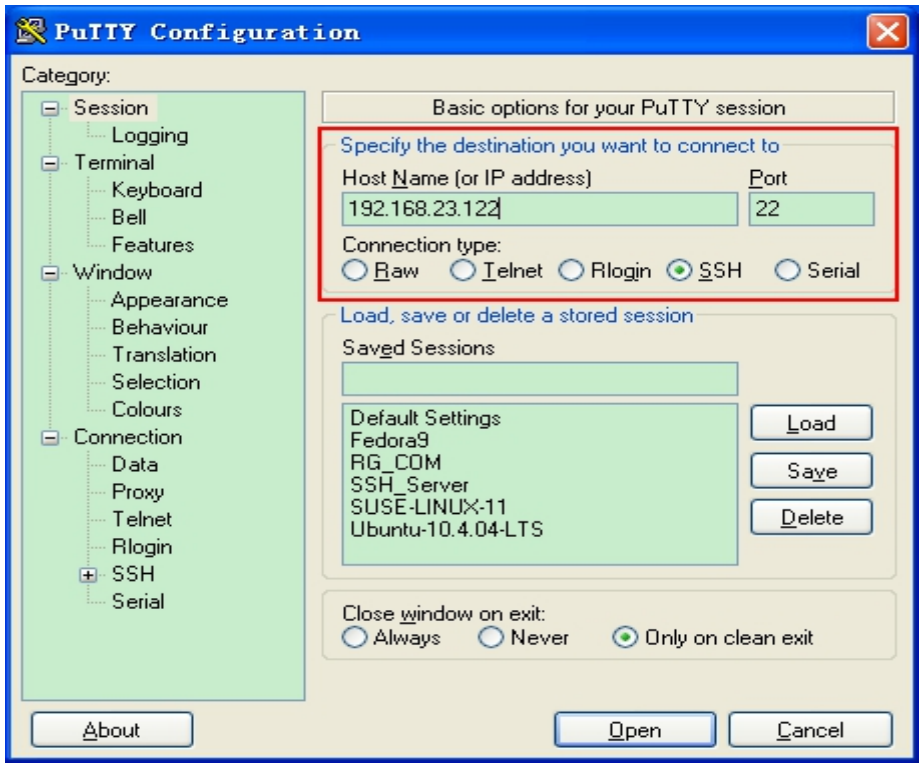
	Authentication timeout: 120 secs Authentication retries: 3 SSH SCP Server: disabled

Configuring the Public Key Authentication

Configuration Steps	● Run the **ip ssh peer** *username* **public-key** { **rsa** | **dsa**}*filename* command to associate a public key file of the client with a user name. When the client is authenticated upon login, a public key file (for example, RSA) is specified based on the user name.
SSH Server	FS#configure terminal FS(config)# ip ssh peer test public-key rsaflash:rsa.pub
Verification	● Configure the public key authentication login mode on the SSH client and specify the private key file. Check whether you can successfully log in to the SSH server from the SSH client. If yes, the public key file on the client is successfully associated with the user name, and public key authentication succeeds.

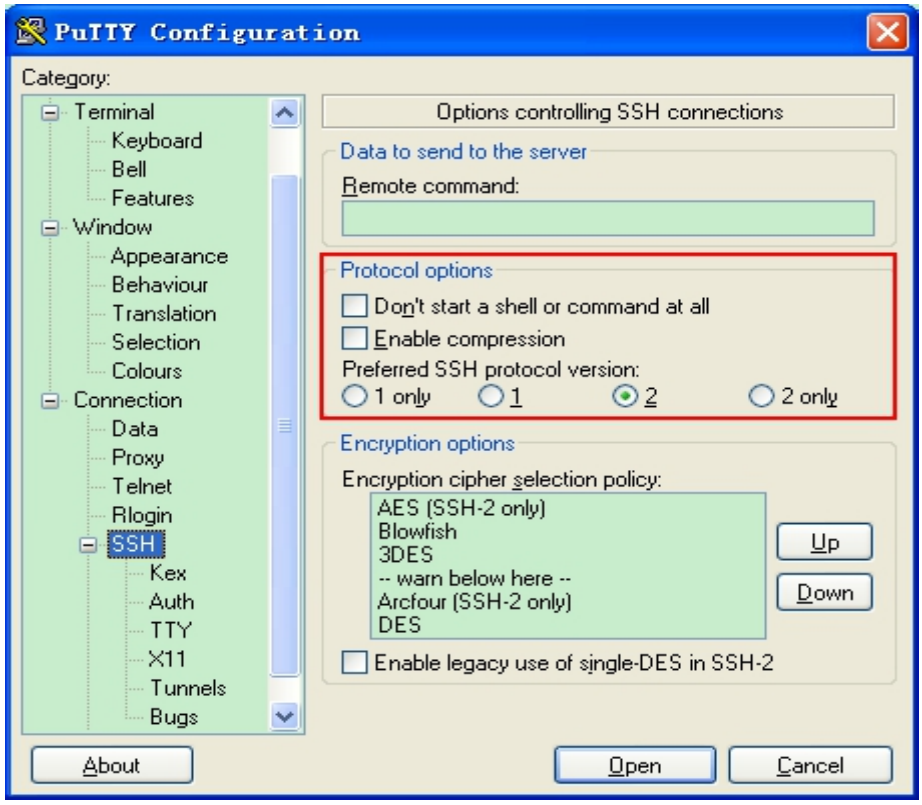
Configuring SSH Device Management

Scenario **Figure**11- 7	 You can use SSH to manage devices on the precondition that the SSH server function is enabled. By default, this function is disabled. The Telnet component that comes with the Windows does not support SSH. Therefore, a third-party client software must be used. Currently, well-compatible client software includes PuTTY, Linux, and SecureCRT. The following takes the PuTTY as an example to introduce the configurations of the SSH client.
Configuration Steps	● Start the PuTTY software. ● On the **Session** option tab of PuTTY, type in the host IP address **192.168.23.122** and SSH port number **22**, and select the connection type **SSH**. ● On the **SSH** option tab of PuTTY, select the preferred SSH protocol version **2**. ● On the **SSH authentication** option tab of PuTTY, select the authentication method **Attempt "keyboard-interactive" auth**. ● Click **Open** to connect to the SSH server. ● Type in the correct user name and password to enter the terminal login interface.
SSH Client	Figure 11-8



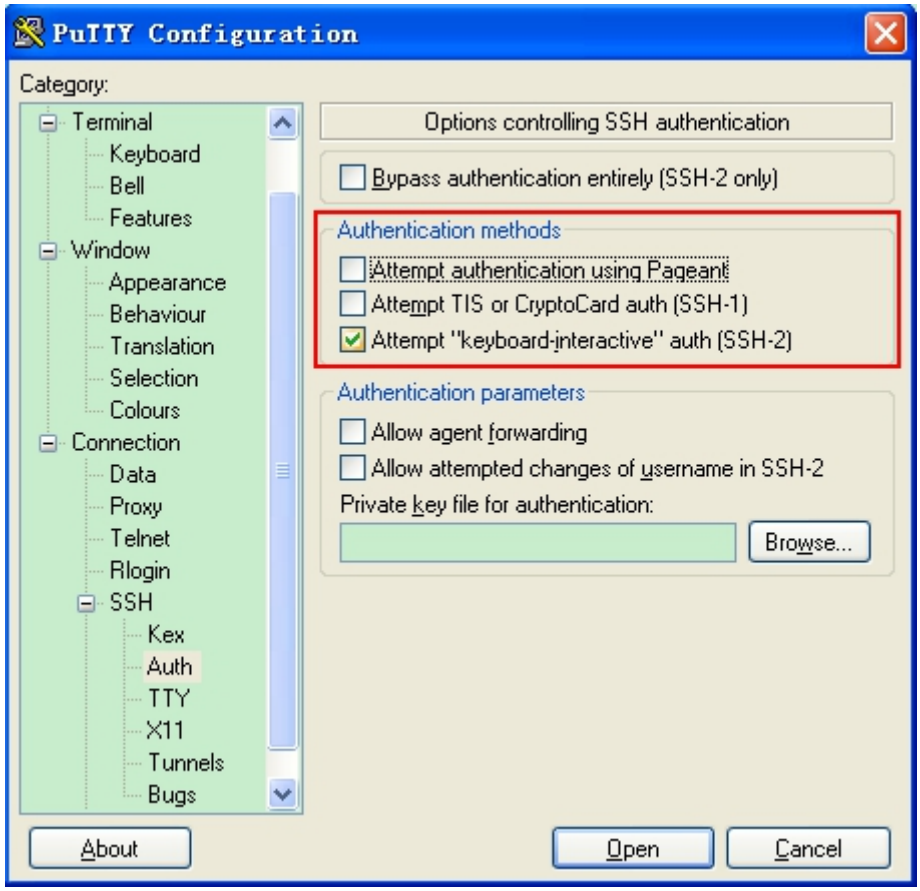
Host Name (or IP address) indicates the IP address of the host to be logged in. In this example, the IP address is **192.168.23.122**. **Port** indicates the port ID 22, that is, the default ID of the port listened by SSH. **Connection type** is **SSH**.

Figure 11- 9



As shown in Figure 11- 9, select **2** as the preferred SSH protocol version in the **Protocol options** pane because SSHv2 is used for login.

Figure 11- 10



As shown in Figure 11- 10, select **Attempt "keyboard-interactive" auth** as the authentication method to support authentication based on the user name and password.

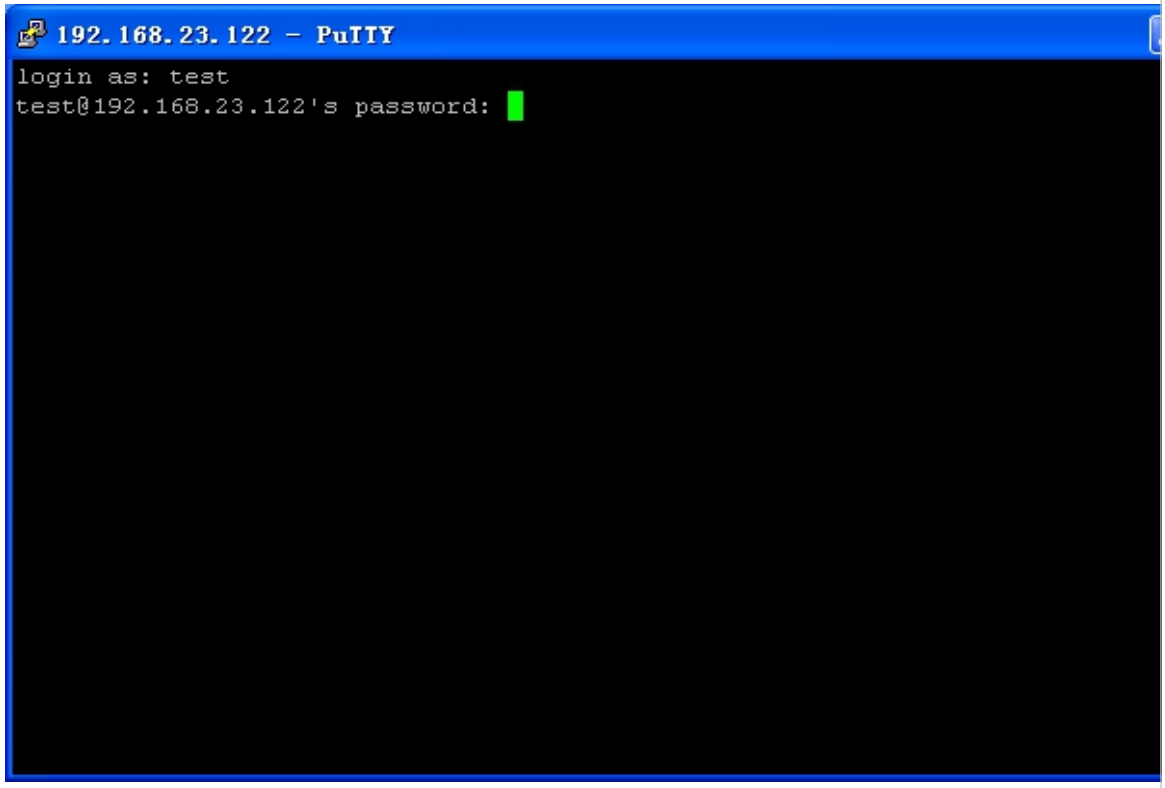
Then, click **Open** to connect to the configured server host, as shown in Figure 11- 10.

Figure 11- 11

The **PuTTY Security Alert** box indicates that you are logging in to the client of the server 192.168.23.122, and asks you whether to receive the key sent from the server.

If you select **Yes**, a login dialog box is displayed, as shown in Figure 11- 11.

Figure 11- 12



Type in the correct user name and password, and you can log in to the SSH terminal interface, as shown in Figure 11- 12.

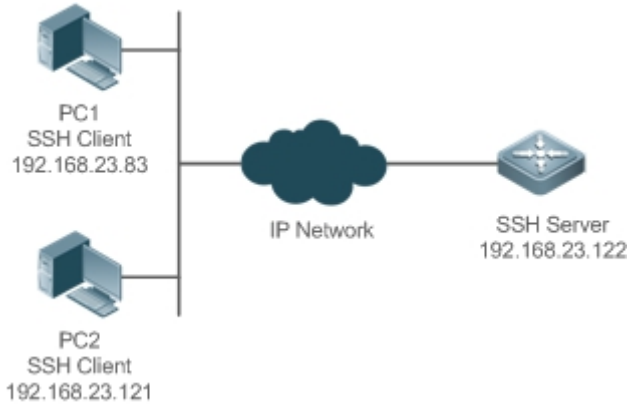
Figure 11- 13

Verification

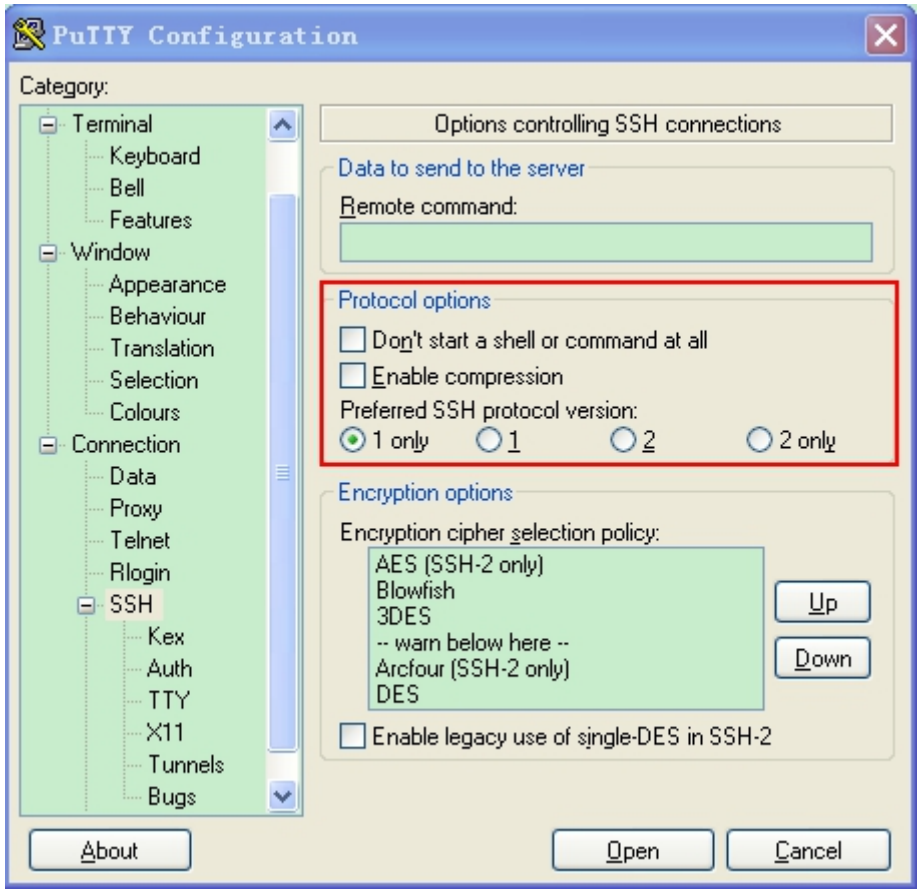
- Run the **show ip ssh** command to display the configurations that are currently effective on the SSH server.
- Run the **show ssh** command to display information about every SSH connection that has been established.

	FS#show ip ssh SSH Enable - version 1.99 Authentication timeout: 120 secs Authentication retries: 3 FS#show ssh Connection Version Encryption Hmac State Username 0 2.0 aes256-cbc hmac-sha1 Session started test

Configuring SSH Local Line Authentication

Scenario **Figure** 11-14	 SSH users can use the local line password for user authentication, as shown in Figure 11-14.To ensure security of data exchange, PC 1 and PC 2 function as the SSH clients, and use the SSH protocol to log in to the network device where the SSH server is enabled. The requirements are as follows: ● SSH users use the local line password authentication mode. ● Five lines, including Line 0 to Line 4, are activated concurrently. The login password is "passzero" for Line 0 and "pass" for the remaining lines. Any user name can be used.
Configuration Steps	Configure the SSH server as follows: ● Enable the SSH server function globally. By default, the SSH server supports two SSH versions: SSHv1 and SSHv2. ● Configure the key. With this key, the SSH server decrypts the encrypted password received from the SSH client, compares the decrypted plain text with the password stored on the server, and returns a message indicating the successful or unsuccessful authentication. SSHv1 uses the RSA key, whereas SSHv2 uses the RSA or DSA key. ● Configure the IP address of the FastEthernet 0/1 interface on the SSH server. The SSH client is connected to the SSH server based on this IP address. The route from the SSH client to the SSH server is reachable. Configure the SSH client as follows: ● Diversified SSH client software is available, including PuTTY, Linux, and SecureCRT. This document takes PuTTY as an example to explain the method for configuring the SSH client. For details about the configuration method, see "Configuration Steps."
SSH Server	Before configuring SSH-related function, ensure that the route from the SSH user to the network segment of the SSH server is reachable. The interface IP address configurations are shown in Figure 11-14. The detailed procedures for configuring IP addresses and routes are omitted. FS(config)# enable service ssh-server

	FS(config)#crypto key generate rsa % You already have RSA keys. % Do you really want to replace them? [yes/no]: Choose the size of the key modulus in the range of 360 to 2048 for your Signature Keys. Choosing a key modulus greater than 512 may take a few minutes. How many bits in the modulus [512]: % Generating 512 bit RSA1 keys ...[ok] % Generating 512 bit RSA keys ...[ok] FS(config)#interface fastEthernet0/1 FS(config-if-fastEthernet0/1)#ip address 192.168.23.122 255.255.255.0 FS(config-if-fastEthernet0/1)#exit FS(config)#line vty 0 FS(config-line)#password passzero FS(config-line)#privilege level 15 FS(config-line)#login FS(config-line)#exit FS(config)#line vty1 4 FS(config-line)#password pass FS(config-line)#privilege level 15 FS(config-line)#login FS(config-line)#exit
SSH Client(PC1/PC2)	Figure 11- 15



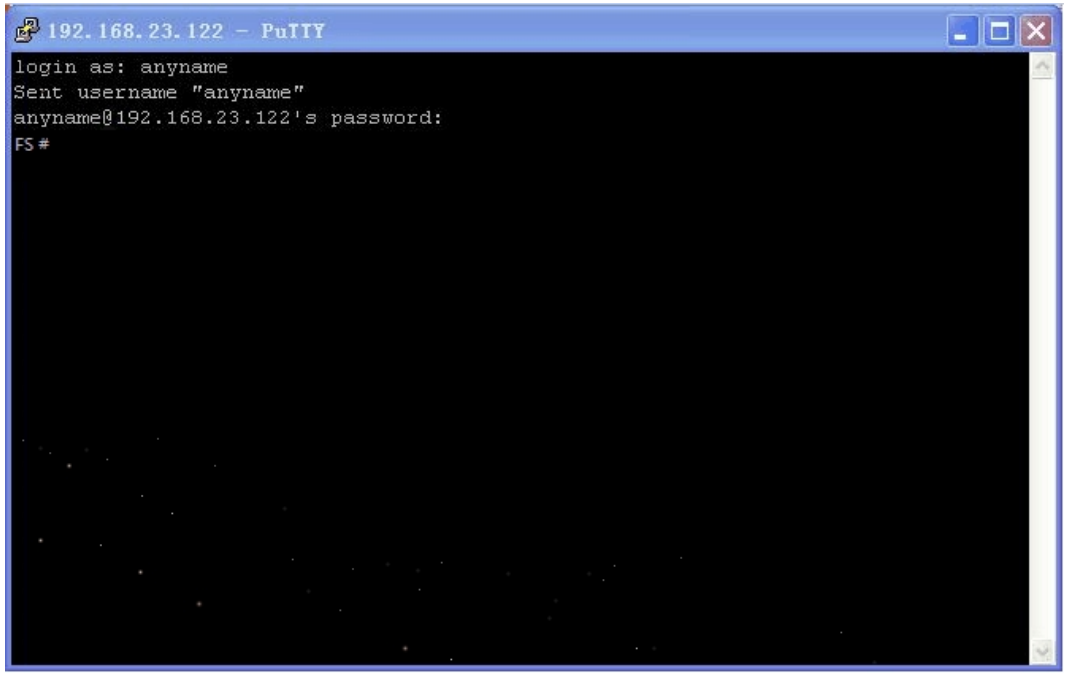
Set the IP address and port ID of the SSH server. As shown in the network topology, the IP address of the server is 192.168.23.122, and the port ID is 22 (For details about the configuration method, see "Configuring SSH Device Management."). Click **Open** to start the SSH server. As the current authentication mode does not require a user name, you can type in any user name, but cannot leave the user name unspecified. (In this example, the user name is "anyname".)

Verification

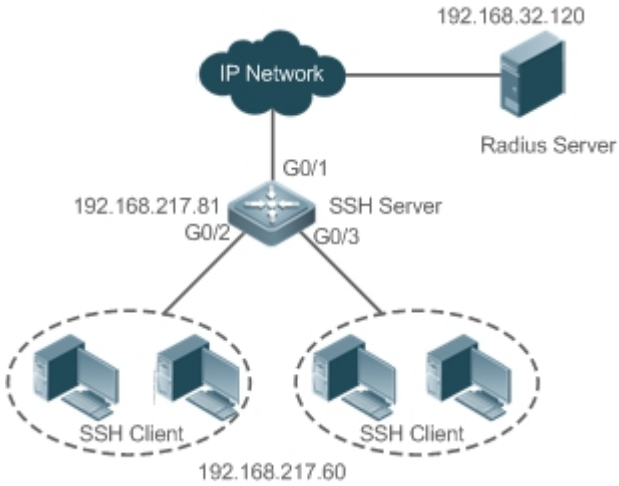
- Run the **show running-config** command to display the current configurations.
- Verify that the SSH client configurations are correct.

SSH Server

```
FS#show running-config
Building configuration...
!
enable secret 5 $1$eyy2$xs28FDw4s2q0tx97
enable service ssh-server
!
interface fastEthernet0/1
ip address 192.168.23.122 255.255.255.0
!
line vty 0
privilege level 15
 login
 password passzero
line vty 1 4
privilege level 15
```

	login password pass ! end
SSH Client	Set up a connection, and enter the correct password. The login password is "passzero" for Line 0 and "pass" for the remaining lines. Then, the SSH server operation interface is displayed, as shown in Figure 11- 16. Figure 11- 16  FS#show users Line User Host(s) Idle Location ---------- ---------- ---------- ---------- ---------- * 0 con 0 --- idle 00:00:00 --- 1 vty 0 --- idle 00:08:02 192.168.23.83 2 vty 1 --- idle 00:00:58 192.168.23.121

Configuring AAA Authentication of SSH Users

Scenario **Figure** 11- 17	 SSH users can use the AAA authentication mode for user authentication, as shown in Figure 11- 17.To ensure security of data exchange, the PC functions as the SSH client, and uses the SSH protocol to log in to the network device where the SSH server is enabled. To better perform security management, the AAA authentication mode is used on the user login interface of the SSH client. Two authentication methods, including Radius server authentication and local authentication, are provided in the AAA authentication method list to ensure reliability. The Radius server authentication method is preferred. If the Radius server does not respond, select the local authentication method.
Configuration Steps	● The route from the SSH client to the SSH server is reachable, and the route from the SSH server to the Radius server is also reachable. ● Configure the SSH server on the network device. The configuration method is already described in the previous example, and therefore omitted here. ● Configure the AAA parameters on the network device. When the AAA authentication mode is used, method lists are created to define the identity authentication and types, and applied to a specified service or interface.
SSH Server	FS(config)# enable service ssh-server FS(config)#crypto key generate rsa % You already have RSA keys. % Do you really want to replace them? [yes/no]: Choose the size of the key modulus in the range of 360 to 2048 for your Signature Keys. Choosing a key modulus greater than 512 may take a few minutes. How many bits in the modulus [512]: % Generating 512 bit RSA1 keys ...[ok] % Generating 512 bit RSA keys ...[ok] FS(config)#crypto key generate dsa Choose the size of the key modulus in the range of 360 to 2048 for your Signature Keys. Choosing a key modulus greater than 512 may take

```
a few minutes.
How many bits in the modulus [512]:
% Generating 512 bit DSA keys ...[ok]
FS(config)#interface gigabitEthernet1/1
FS(config-if-gigabitEthernet1/1)#ip address 192.168.217.81 255.255.255.0
FS(config-if-gigabitEthernet1/1)#exit
FS#configure terminal
FS(config)#aaa new-model
FS(config)#radius-server host 192.168.32.120
FS(config)#radius-server key aaaradius
FS(config)#aaa authentication login methodgroup radius local
FS(config)#line vty 0 4
FS(config-line)#login authentication method
FS(config-line)#exit
FS(config)#username user1 privilege 1 password 111
FS(config)#username user2 privilege 10 password 222
FS(config)#username user3 privilege 15 password 333
FS(config)#enable secret w
```

Verification

- Run the **show running-config** command to display the current configurations.
- This example assumes that the SAM server is used.
- Set up a remote SSH connection on the PC.
- Check the login user.

```
FS#show run
aaa new-model
!
aaa authentication login method group radius local
!
username user1 password 111
username user2 password 222
username user2 privilege 10
username user3 password 333
username user3 privilege 15
no service password-encryption
!
```

```
radius-server host 192.168.32.120
radius-server key aaaradius
enable secret 5 $1$hbgz$ArCsyqty6yyzzp03
enable service ssh-server
!
interface gigabitEthernet1/1
  no ip proxy-arp
ip address 192.168.217.81 255.255.255.0
!
ip route 0.0.0.0 0.0.0.0 192.168.217.1
!
line con 0
line vty 0 4
  login authentication method
!
End
```

On the SSH client, choose **System Management**>**Device Management**, and add the device IP address **192.168.217.81** and the device key **aaaradius**.

Choose **Security Management**>**Device Management Rights**, and set the rights of the login user.

Choose **Security Management**>**Device Administrator**, and add the user name **user** and password **pass**.

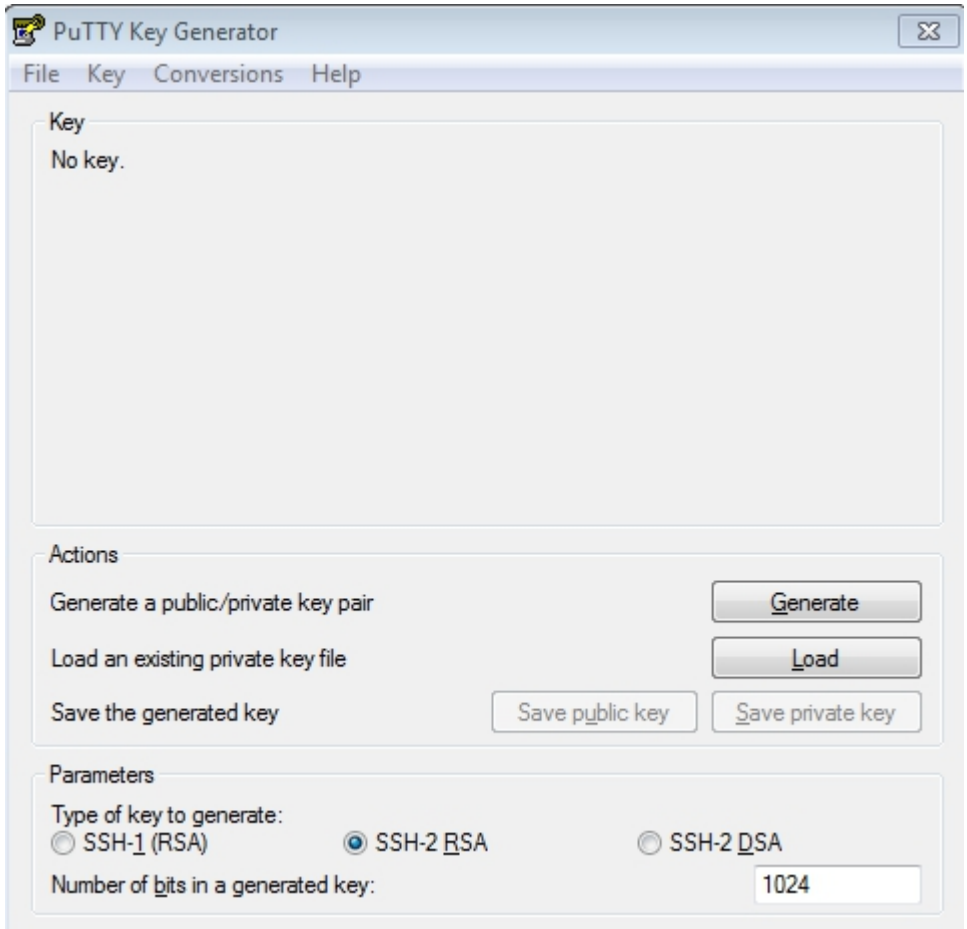
Configure the SSH client and set up a connection to the SSH server. For details, see the previous example.

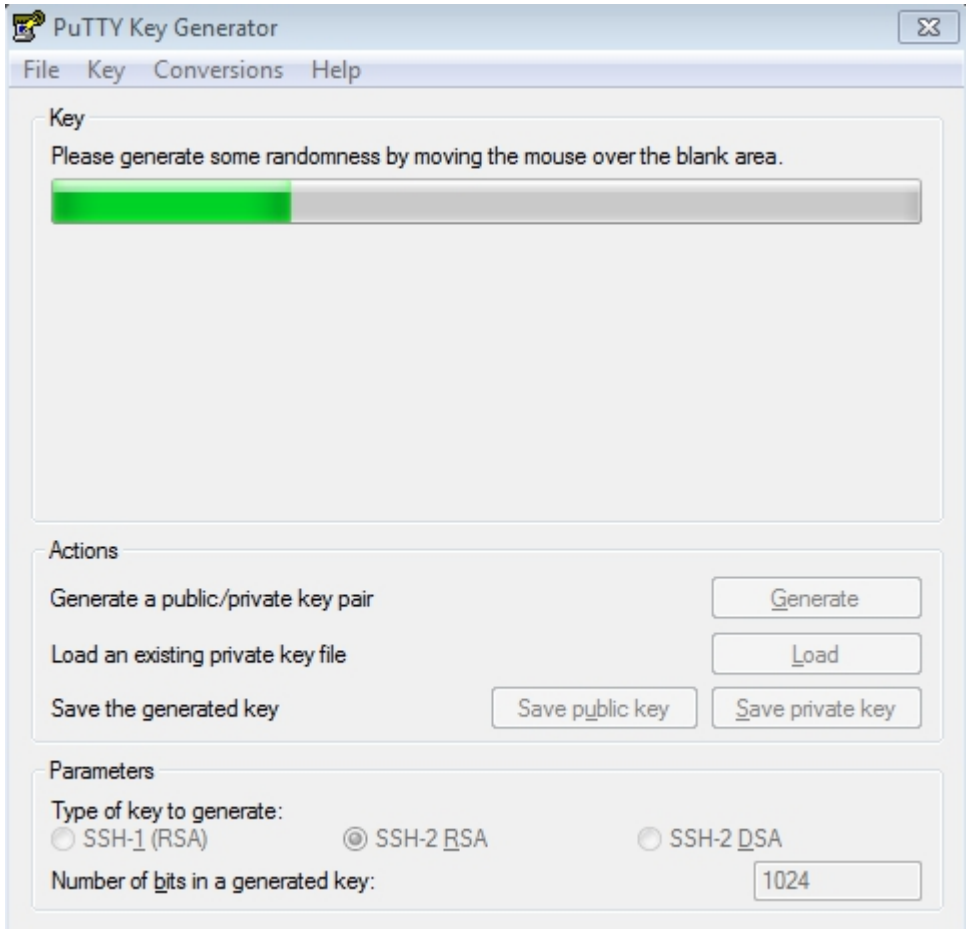
Type in the user name **user** and password **pass**. Verify that you can log in to the SSH server successfully.

```
FS#show users
    Line          User          Host(s)          Idle          Location
   0 con 0                      idle             00:00:31
* 1 vty 0      user        idle                 00:00:33     192.168.217.60
```

↘ Configuring Public Key Authentication of SSH Users

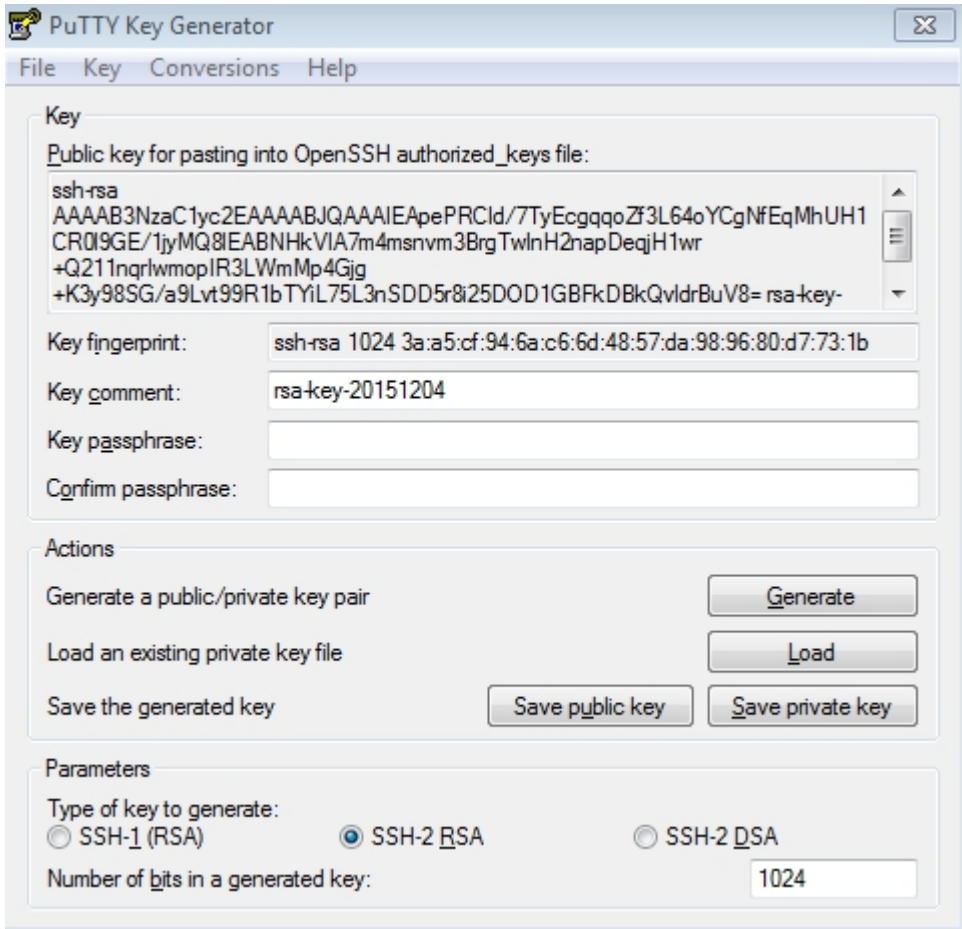
Scenario **Figure** 11- 18	 SSH users can use the public key for user authentication, and the public key algorithm is RSA or DSA, as shown in Figure 11- 18.SSH is configured on the client so that a secure connection is set up between the SSH client and the SSH server.
Configuration Steps	● To implement public key authentication on the client, generate a key pair (for example, RSA key) on the client, place the public key on the SSH server, and select the public key authentication mode.

	After the key pair is generated on the client, you must save and upload the public key file to the server and complete the server-related settings before you can continue to configure the client and connect the client with the server. ● After the key is generated on the client, copy the public key file from the client to the flash of the SSH server, and associate the file with an SSH user name. A user can be associated with one RSA public key and one DSA public key.
SSH Client	Run the **puttygen.exe** software on the client. Select **SSH-2 RSA** in the **Parameters** pane, and click **Generate** to generate a key, as shown in Figure 11- 19. Figure 11- 19  When a key is being generated, you need to constantly move the mouse over a blank area outside the green progress bar; otherwise, the progress bar does not move and key generation stops, as shown in Figure 11- 20. Figure 11- 20



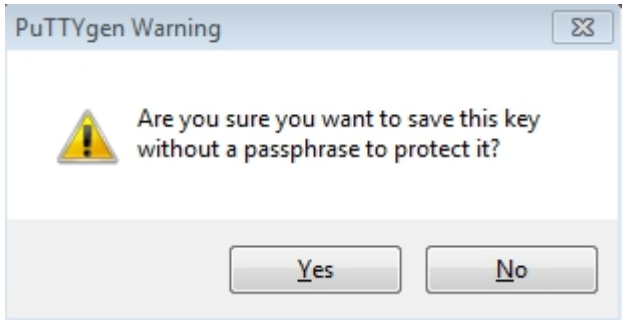
To ensure security of the RSA public key authentication, the length of the generated RSA key pair must be equal to or larger than 768 bits. In this example, the length is set to 1024 bits.

Figure 11- 21



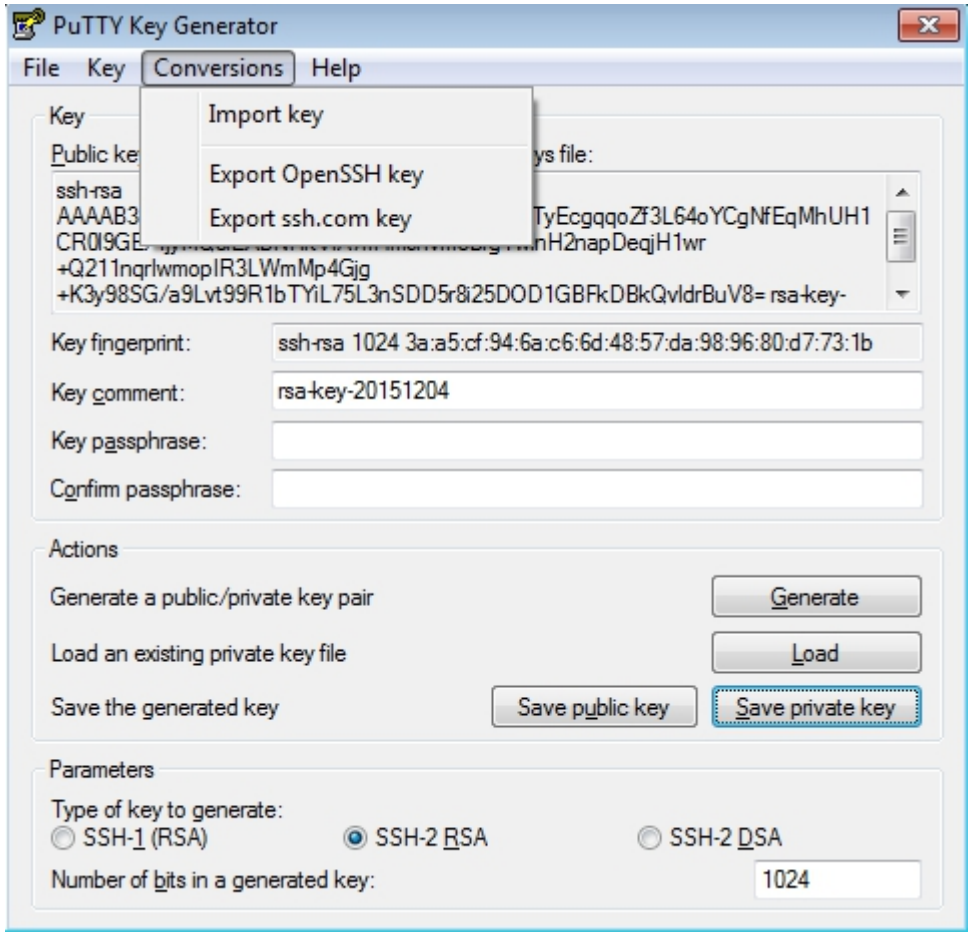
After the key pair is generated, click **Save public key**, type in the public key name **test_key.pub**, select the storage path, and click **Save**. Then click **Save private key**. The following prompt box is displayed. Select **Yes**, type in the public key name **test_private**, and click **Save**.

Figure 11- 22

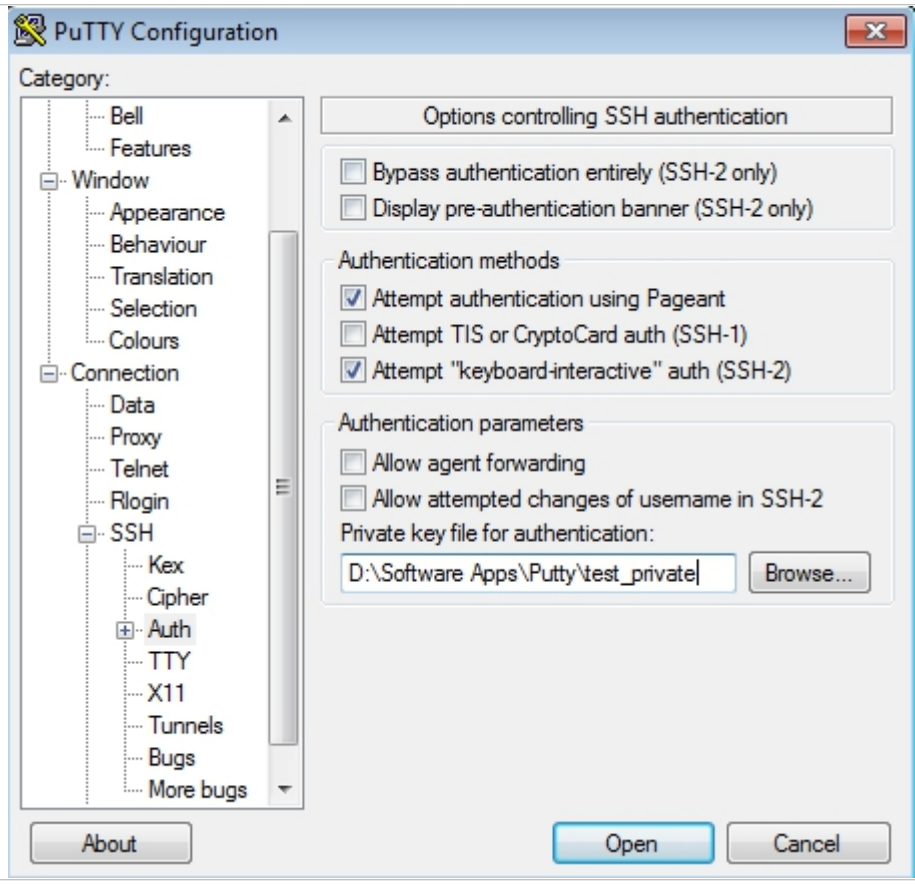


You must select the OpenSSH key file; otherwise, the key file cannot be used. The **puttygen.exe** software can be used to generate a key file in OpenSSH format, but this file cannot be directly used by the PuTTY client. You must use **puttygen.exe** to convert the private key to the PuTTY format. Format conversion is not required for the public key file stored on the server, and the format of this file is still OpenSSH, as shown in Figure 11- 23.

Figure 11- 23



SSH Server	FS#configure terminal FS(config)# ip ssh peer test public-key rsaflash:test_key.pub
Verification	● After completing the basic configurations of the client and the server, specify the private key file **test_private** on the PuTTY client, and set the host IP address to **192.168.23.122** and port ID to **22** to set up a connection between the client and the server. In this way, the client can use the public key authentication mode to log in to the network device.
	Figure 11- 24



Common Errors

- The **no crypto key generate** command is used to delete a key.

11.4.2 Configuring the SCP Service

Configuration Effect

After the SCP function is enabled on a network device, you can directly download files from the network device and upload local files to the network device. In addition, all interactive data is encrypted, featuring authentication and security.

Notes

- The SSH server must be enabled in advance.

Configuration Steps

Enabling the SCP Server

- Mandatory.
- By default, the SCP server function is disabled. Run the **ip scp server enable** command to enable the SCP server function in global configuration mode.

Verification

Run the **show ip ssh** command to check whether the SCP server function is enabled.

Related Commands

Enabling the SCP Server

Command	**ip scp server enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	This command is used to enable the SCP server. Run the **no ip scp server enable** command to disable the SCP server.

Configuration Example

Enabling the SCP Server

Configuration Steps	● Run the **ip scp server enable** command to enable the SCP server.
	FS#configure terminal FS(config)#ip scp server enable
Verification	● Run the **show ip ssh** command to check whether the SCP server function is enabled.
	FS(config)#show ipssh SSH Enable - version 1.99 Authentication timeout: 120 secs Authentication retries: 3 SSH SCP Server: enabled

Configuring SSH File Transfer

Scenario **Figure** 11- 25	 The SCP service is enabled on the server, and SCP commands are used on the client to transfer data to the server.
Configuration Steps	● Enable the SCP service on the server. The SCP server uses SSH threading. When connecting to a network device for SCP transmission, the client occupies a VTY session (You can finds out that the user type is SSH by running the show user command). ● On the client, use SCP commands to upload files to the server, or download files from the server. Syntax of the SCP command: scp [-1246BCpqrv] [-c cipher] [-F ssh_config] [-iidentity_file] [-l limit] [-o ssh_option] [-P port] [-S program] [[user@]host1:]file1 [...] [[user@]host2:]file2

	Descriptions of some options: -1: Uses SSHv1 (If not specified, SSHv2 is used by default); -2: Uses SSHv2 (by default); -C: Uses compressed transmission. -c: Specifies the encryption algorithm to be used. -r:Transmits the whole directory; -i: Specifies the key file to be used. -l: Limits the transmission speed (unit: Kbit/s). For other parameters, see the filescp.0. Most options are related to terminals. Few options are supported on both terminals and servers. FS's SCP servers do not support d-p-q-r options. When these options are applied, there are prompts.
SSH Server	FS#configure terminal FS(config)# ip scp server enable
Verification	● File transmission example on the Ubuntu 7.10 system: Set the username of a client to **test** and copy the **config.text** file from the network device with the IP address of 192.168.195.188 to the **/root** directory on the local device.
	root@dhcpd:~#scp test@192.168.23.122:/config.text /root/config.text test@192.168.195.188's password: config.text 100% 1506 1.5KB/s 00:00 Read from remote host 192.168.195.188: Connection reset by peer

11.4.3 Configuring the SSH Client

Configuration Effect

On the network device that supports the SSH server, enable the SSH server function, and specify the user authentication method and supported SSH versions. Then, you can use the built-in SSH client function of the device to set up a secure connection with the SSH server, implementing remote device management.

Notes

- The SSH server function must be configured in advance on the device that needs to remotely support the SSH server.
- The SSH client must communicate with the SSH server properly.

Configuration Steps

Specifying the Source Interface of the SSH Client

- (Optional) This configuration must be performed on the SSH-client device.

Establishing a Session with the SSH Server

- (Optional) Use the **ssh** command on the client to set up a connection with a remote server.
- Before using this command, enable the SSH server function and configure the SSH key and authentication mode on the server.

Recovering an Established SSH Session

- (Optional) Run the related command to recover a session after temporary stop if necessary.

Disconnecting a Suspended SSH Session

- (Optional) This configuration must be performed on the SSH client if you need to disconnect a specified SSH session.

Verification

Run the **show ssh-session** command to display information about every established SSH client session.

Related Commands

Specifying the Source Interface of the SSH Client

Command	**ip ssh source-interface** *interface-name*
Parameter Description	*interface-name*: Specifies an interface, the IP address of which will be used as the source address of an SSH client session.
Command Mode	Global configuration mode
Usage Guide	This command is used to specify an interface, the IP address of which will be used as the global source address of an SSH client session. When the **ssh** command is used to connect to an SSH server, this global configuration will be used if a source interface or a source address is not specified for this connection. Run the **no ip ssh source-interface** command to restore the default setting.

Establishing a Session with the SSH Server

Command	**ssh** [**oob**] [**-v** {**1** | **2** }][**-c** {**3des** | **aes128-cbc** | **aes192-cbc** | **aes256-cbc** }] [**-l** *username*][**-m** {**hmac-md5-96** | **hmac-md5-128** | **hmac-sha1-96** | **hmac-sha1-160** }] [**-p** *port-num*]{ *ip-addr*| *hostname*}[**via** *mgmt-name*][**/source** {**ip** *A.B.C.D* **| ipv6** *X:X:X:X::X* **| interface** *interface-name*}]
Parameter Description	N/A
Command Mode	User EXEC mode
Usage Guide	The **ssh** command is used to set up a secure and encrypted connection from the local device (an SSH client) to another device (an SSH server) or any other server that supports SSHv1 or SSHv2. This connection provides a mechanism similar to the Telnet connection except that all data transmitted over this connection is encrypted. Based on authentication and encryption, the SSH client can set up a secure connection on an insecure network. SSHv1 supports only the DES (56-bit key) and 3DES (168-bit key) encryption algorithms. SSHv2 supports the following Advanced Encryption Standards (AES):ASE128-CBC, AES192-CBC, AES256-CBC, AES128-CTR, AES192-CTR, and AES256-CTR. SSHv1 does not support the Hashed Message Authentication Code (HMAC). If you specify an unmatched encryption or authentication algorithm when selecting an SSH version, the

	unmatched algorithm will be ignored when a connection is set up.

Recovering an Established SSH Client Session

Command	**ssh-session** *session-id*
Parameter Description	*session-id*: Indicates the ID of an established SSH client session.
Command Mode	User EXEC mode
Usage Guide	This command is used to restore the use of an established SSH client session. When the **ssh** command is used to initiate an SSH client session, you can press Ctrl+Shift+6+X to temporarily exit the session. To recover this session, run the **ssh-session** command. In addition, if the session is already established, you can run the **show ssh-session** command to display information about the established session.

Disconnecting a Suspended SSH Session

Command	**disconnect ssh-session** *session-id*
Parameter Description	*session-id*: Indicates the ID of a suspended SSH client session.
Command Mode	User EXEC mode
Usage Guide	You can specify an SSH client session ID to disconnect the specified SSH client session.

Configuration Example

Specifying the Source Interface of the SSH Client

Configuration Steps	● Run the **ip ssh source-interface** *interface-name* command to specify an interface, the IP address of which will be used as the global source address of an SSH client session.
	FS#configure terminal FS(config)#ipsshsource-interface gigabitEthernet 0/1
Verification	N/A

Establishing a Session with the SSH Server

Scenario **Figure** 11- 26	 The SSH server function is enabled on the server. The **ssh** command is used on the client to set up a secure connection with the server.
Configuration Steps	● Enable the SSH server function on the server. ● Configure the SSH key on the server. ● Configure the authentication mode of the SSH server, and use the local line authentication mode for Line 0 to Line 4.

	• Configure the IP address of the Gi 0/1 interface of the SSH server. The client will use this address as the source address to connect to the SSH server. • Configure the SSH client, and specify the source address of the SSH client. By default, the SSH server supports two SSH versions: SSHv1 and SSHv2. With this key, the SSH server decrypts the encrypted password received from the SSH client, compares the decrypted plain text with the password stored on the server, and returns a message indicating the successful or unsuccessful authentication. SSHv1 uses an RSA key, whereas SSHv2 uses an RSA or DSA key. The authentication mode used by the SSH server is local line authentication. The local user name is admin, and the password is 123456. The SSH client is connected to the SSH server based on this IP address. The routes from the SSH clients to the SSH server are reachable. Configure the IP address of the Gi 0/1 interface of the SSH server. The client will use this address as the source address to connect to the SSH server.
SSH Server	FS#configure terminal FS(config)#enable service ssh-server FS(config)#crypto key generate rsa % You already have RSA keys. % Do you really want to replace them? [yes/no]: Choose the size of the key modulus in the range of 360 to 2048 for your Signature Keys. Choosing a key modulus greater than 512 may take a few minutes. How many bits in the modulus [512]: % Generating 512 bit RSA1 keys ...[ok] % Generating 512 bit RSA keys ...[ok] FS(config)#line vty 0 4 FS(config-line)#login local FS(config-line)#exit FS(config)#username admin password 123456 FS(config)#username admin privilege 15 FS(config-line)#exit FS(config)#interface gigabitEthernet0/1 FS(config-if-gigabitEthernet0/1)#ip address 192.168.23.122 255.255.255.0 FS(config-if-gigabitEthernet0/1)#exit
SSH Client	FS(config)#interface gigabitEthernet0/1 FS(config-if-gigabitEthernet0/1)#ip address 192.168.23.83 255.255.255.0

	FS(config-if-gigabitEthernet0/1)#exit FS(config)#ipsshsource-interface gigabitEthernet 0/1
Verification	● Run the **show running-config| include username** and **show ip ssh** commands to verify whether the SSH server configurations are correct. ● On the SSH client, set up a connection with a remote SSH server. After the connection is set up, type in the correct password **123456**. The SSH server operation interface is displayed. Check the login user on the Console of the SSH client.
	FS(config)#sh running-config | include username username admin password admin username admin privilege 15 FS(config)#sh running-config | begin line line con 0 line vty 0 4 login local ! ! end
	● Verify whether the SSH client configurations are correct.
	FS#ssh -l admin 192.168.23.122 %Trying 192.168.23.122, 22,...open admin@192.168.23.122's password: FS# FS#sh users Line User Host(s) Idle Location 0 con 0 idle 00:00:00 * 1 vty 0 admin idle 00:00:36 192.168.217.20

11.5 Monitoring

Displaying

Description	Command
Displays the effective SSH server configurations.	**show ipssh**
Displays the established SSH connection.	**show ssh**
Displays the public information of the SSH public key.	**show crypto key mypubkey**
Displays the established SSH client session.	**show ssh-session**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs SSH sessions.	**debug ssh**
Debugs SSH client sessions.	**debug ssh client**

12 Configuring URPF

12.1 Overview

Unicast Reverse Path Forwarding (URPF) is a function that protects the network against source address spoofing.

URPF obtains the source address and inbound interface of a received packet, and searches a forwarding entry in the forwarding table based on the source address. If the entry does not exist, the packet is dropped. If the outbound interface of the forwarding entry does not match the inbound interface of the packet, the packet is also dropped. Otherwise, the packet is forwarded.

URPF is implemented in two modes:

- Strict mode: It is often deployed on a point-to-point (P2P) interface, and inbound and outbound data streams must go through the network of the P2P interface.
- Loose mode: It is applicable to the asymmetric routes or multihomed network that have the problem of asymmetric traffic.

Protocols and Standards

- RFC 2827: Network Ingress Filtering: DDOS Attacks which employ IP Source Address Spoofing
- RFC 3704: Ingress Filtering for Multi-homed Networks

12.2 Applications

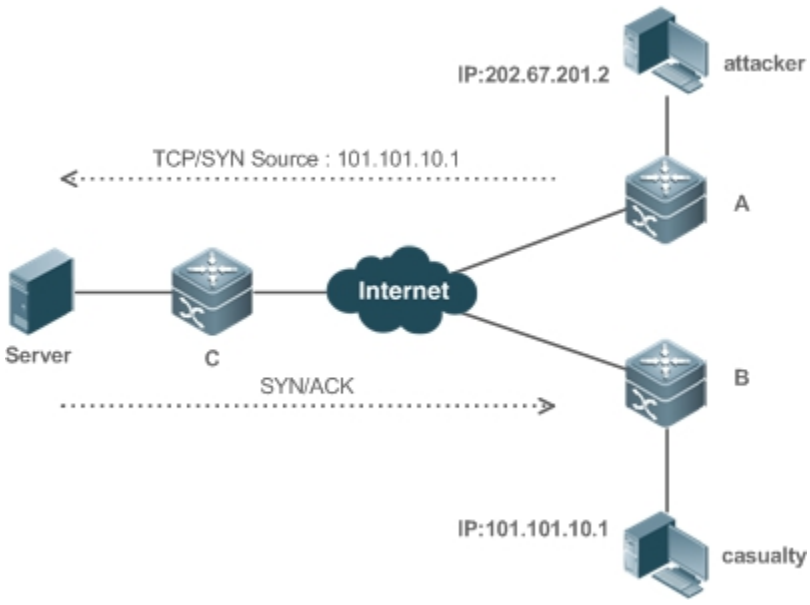
Application	Description
Strict Mode	Block the packets with spoofed sourced addresses at the access layer or aggregation layer to prevent sending these packets from PCs to the core network.
Loose Mode	On a multihomed network, the user network is connected to multiple Internet service providers (ISPs), and the inbound and outbound traffic is not symmetric. Deploy the URPF loose mode on the outbound interface connected to ISPs to prevent invalid packets from attacking the user network.

12.2.1 Strict Mode

Scenario

An attacker initiates an attack by sending packets with the spoofed source address 11.0.0.1. As a result, the server sends a lot of SYN or ACK packets to the hosts that do not initiate the attack, and the host with the real source address 11.0.0.1 is also affected. Even worse, if the network administrator determines that this address initiates an attack to the network, and therefore blocks all data streams coming from this source address, the denial of service (DoS) of this source address occurs.

Figure 12- 1



Remarks	The attacker sends spoofing packets using a spoofed address of the casualty.

Deployment

- Deploy the URPF strict mode on device A to protect the device against source address spoofing.

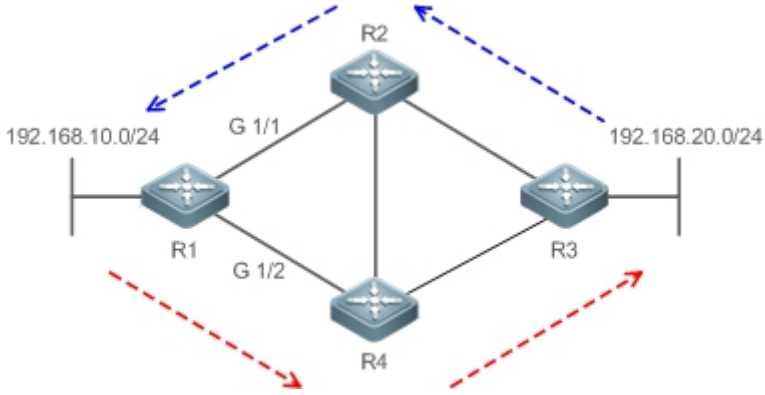
12.2.2 Loose Mode

Scenario

The asymmetric route is a common network application used to control the network traffic or to meet the routing policy requirements.

As shown in Figure 12- 2, if the URPF strict mode is enabled on the G1/1 interface of R 1, R1 receives a packet from the network segment 192.168.20.0/24 on the G1/1 interface, but the interface obtained through the URPF check is G1/2. Therefore, this packet fails in the URPF check and is dropped.

Figure 12- 2



Deployment

- Reversely search a route based on the source IP address of a received packet. The purpose is to find a route, and it is not required that the outbound interface of the next hop on the route must be the inbound interface of the received packet.

- The URPF loose mode can resolve the asymmetric traffic problem of the asymmetric route and prevents access of invalid data streams.

12.3 Features

Basic Concepts

URPF Strict Mode

Obtain the source address and inbound interface of a received packet, and search a forwarding entry in the forwarding table based on the source address. If the entry does not exist, the packet is dropped. If the outbound interface of the forwarding entry does not match the inbound interface of the packet, the packet is also dropped. The strict mode requires that the inbound interface of a received packet must be the outbound interface of the route entry to the source address of the packet.

URPF Loose Mode

Reversely search a route based on the source IP address of a received packet. The purpose is to find a route, and it is not required that the outbound interface of the next hop on the route must be the inbound interface of the received packet. However, the route cannot be a route of a host on the local network.

URPF Packet Loss Rate

The URPF packet loss rate is equal to the number of packets dropped due to the URPF check per second. The unit is packets/second, that is, pps.

Calculation Interval of the URPF Packet Loss Rate

It is the interval from the previous time the packet loss rate is calculated to the current time the packet loss rate is calculated.

Sampling Interval of the URPF Packet Loss Rate

It the interval at which the number of lost packets is collected for calculating the packet loss rate. This interval must be equal to or longer than the calculation interval of the packet loss rate.

Threshold of the URPF Packet Loss Rate

It refers to the maximum packet loss rate that is acceptable. When the packet loss rate exceeds the threshold, alarms can be sent to users through syslogs or trap messages. You can adjust the threshold of the packet loss rate based on the actual conditions of the network.

Alarm Interval of the URPF Packet Loss Rate

It is the interval at which alarms are sent to users. You can adjust the alarm based on the actual conditions of the network to prevent frequently output of logs or trap messages.

Calculation of the URPS Packet Loss Rate

Between the period of time from enabling of URPF to the time that the sampling interval arrives, the packet loss rate is equal to the number of lost packets measured within the sampling interval divided by the URPF enabling duration. After that, the packet loss rate is calculated as follows: Current packet loss rate = (Current number of lost packets measured at the calculation interval – Number of lost packets measured before the sampling interval)/Sampling interval

Overview

Feature	Description
Enabling URPF	Enable URPF to perform a URPF check,thus protecting the device against source address spoofing.
Notifying the URPF Packet Loss Rate	To facilitate monitoring of information about lost packets after URPF is enabled, FS devices support the use of syslogs and trap messages to proactively notify users of the packet loss information detected in the URPF check.

12.3.1 Enabling URPF

Enable URPF to perform a URPF check on IPv4 or IPv6 packets, thus protecting the device against source address spoofing.

Working Principle

URPF can be applied to IP packets based on configurations, but the following packets are not checked by URPF:

1. After URPF is enabled, the source address of a packet is checked only if the destination address of the packet is a unicast address, and is not checked if the packet is a multicast packet or an IPv4 broadcast packet.

2. If the source IP address of a DHCP/BOOTP packet is 0.0.0.0 and the destination IP address is 255.255.255.255, the packet is not checked by URPF.

3. A loopback packet sent by the local device to itself is not checked by URPF.

URPF Configured in Interface Configuration Mode

URPF, including IPv4 URPF and IPv6 URPF, is performed on packets received on the configured interface.

By default, the default route is not used for the URPF check. You can configure data to use the default route for the URPF check if necessary.

A switch supports configuration of URPF on a routed port of L3 aggregate port (AP). Some switches also support configuration of URPF on a switch virtual interface (SVI). (For details about the switch products, contact FS technical support engineers.) The following constraints exists:

- After URPF is enabled on interfaces, a URPF check is performed on all packets received on physical ports corresponding to these interfaces, which increase the scope of packets checked by URPF. If a packet received on a tunnel port is also received on the preceding physical ports, the packet is also checked by URPF. In such a scenario, be cautious in enabling URPF.
- After URPF is enabled, the route forwarding capacity of the device will be reduced by half.
- After the URPF strict mode is enabled, if a packet received on an interface matches an equal-cost route during the URPF check, the packet will be processed according to the URPF loose mode.

Related Configuration

Enabling URPF for a Specified Interface

By default, URPF is disabled for a specified interface.

Run the **ip verify unicast source reachable-via** {**rx** | **any** }[**allow-default**][*acl-name*] command to enable or disable the IPv4 or IPv6 URPF function for a specified interface.

By default, the default route is not used for the URPF check. You can use the **allow-default** keyword to use the default route for the URPF check if necessary.

12.3.2 Notifying the URPF Packet Loss Rate

To facilitate monitoring of information about lost packets after URPF is enabled, FS devices support the use of syslogs and trap messages to proactively notify users of the packet loss information detected in the URPF check.

Working Principle

Between the period of time from enabling of URPF to the time that the sampling interval arrives, the packet loss rate is equal to the number of lost packets measured within the sampling interval divided by the URPF enabling duration. After that, the packet loss rate is calculated as follows: Current packet loss rate = (Current number of lost packets measured at the calculation interval – Number of lost packets measured before the sampling interval)/Sampling interval

After the function of monitoring the URPF packet loss information is enabled, the device can proactively send syslogs or trap messages to notify users of the packet loss information detected in the URPF check so that users can monitor the network status conveniently.

Related Configuration

Configuring the Calculation Interval of the URPF Packet Loss Rate

By default, the calculation interval of the URPF packet loss rate is 30s. If the calculation interval is found too short, run the **ip verify urpf drop-rate compute interval** *seconds* command to modify the calculation interval.

The calculation interval of the URPF packet loss rate ranges from 30 to 300.

Configuring the Alarm Interval of the URPF Packet Loss Rate

By default, the alarm interval of the URPF packet loss rate is 300s. If the alarm interval is found inappropriate, run the **ip verify urpf drop-rate notify hold-down** *seconds* command to modify the alarm interval of the URPF packet loss rate.

The unit of the alarm interval is second. The value ranges from 30 to 300.

Configuring the Function of Monitoring the URPF Packet Loss Information

By default, the function of monitoringthe URPF packet loss information is disabled.

Run the **ip verify urpf drop-rate notify** command to enable or disable the function of monitoringthe URPF packet loss information.

Configuring the Threshold of the URPF Packet Loss Rate

By default, the threshold of the URPF packet loss rate is 1000 pps. If the threshold is fond inappropriate, run the **ip verify urpf notification threshold** *rate-value* command to modify the threshold of the URPF packet loss rate.

The unit of the threshold is pps. The value ranges from 0 to 4,294,967,295.

12.4 Configuration

Configuration Item	Description and Command	
Enabling URPF	(Mandatory) It is used to enable URPF.	
	ip unicast source reachable-via { rx | any } [allow-default] (Interface configuration mode)	Enables URPF for a specified interface.
Configuring the Function of Monitoring the URPF Packet Loss Information	(Optional) It is used to enable the function of monitoring the URPF packet loss information.	
	ip verify urpf drop-rate compute interval *seconds*	Configures the calculation interval of the URPF packet loss rate.
	ip verify urpf drop-rate notify	Configures the function of monitoring URPF packet loss information.
	ip verify urpf drop-rate notify hold-down *seconds*	Configures the alarm interval of the URPF packet loss rate.
	Ip verify urpf notification threshold *rate-value*	Configures the threshold of the URPF packet loss rate.

12.4.1 Enabling URPF

Configuration Effect

- Enable URPF to perform a URPF check on IP packets, thus protecting the device against source address spoofing.
- URPF can be enabled in interface configuration mode
- URPF enabled in interface configuration mode supports both the strict and loose modes.

Notes

- URPF is implemented with the help of the existing unicast routes on the network. Therefore, unicast routes must be configured on the network.
- URPF cannot be enabled on a range of interfaces.

Configuration Steps

Enabling IPv4 URPF for a Specified Interface

- Mandatory.
- Switches supports configuration of IPv4 URPF on a routed port or L3 AP port, other products supports configuration of IPv4 URPF on a routed port.

Verification

Enable URPF and check the source address as follows:

- If the strict mode is used, check whether a packet is forwarded only when the forwarding table contains the source address of the received IPv4 packet and the outbound interface of the searched forwarding entry matches the inbound interface of the packet; otherwise, the packet is dropped.

- If the loose mode is used, check whether a packet is forwarded when a forwarding entry can be found in the forwarding table for the source address of the received IPv4 packet; otherwise, the packet is dropped.

Related Commands

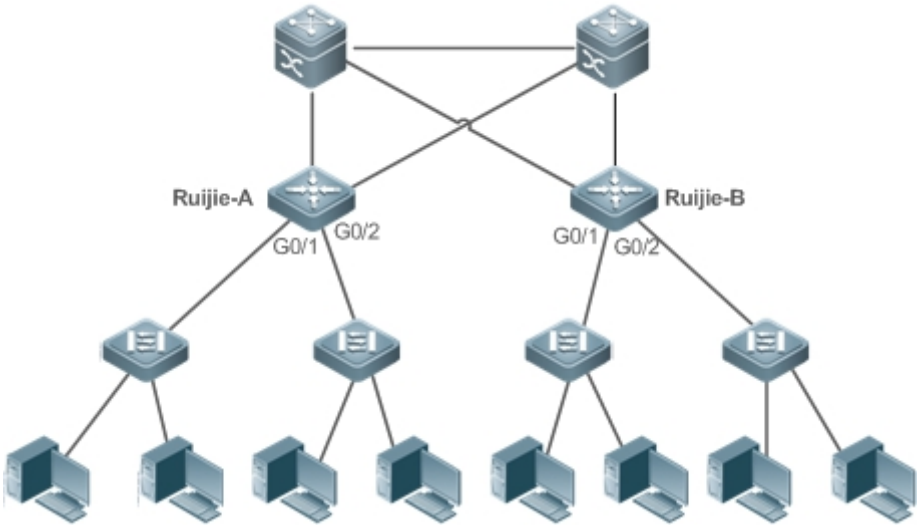
Enabling IPv4 URPF for a Specified Interface

Command	**ip verify unicast source reachable-via { rx | any } [allow-default]**
Parameter Description	**rx**: Indicates that the URPF check is implemented in strict mode. The strict mode requires that the outbound interface of the forwarding entry found in the forwarding table based on the source address of a received IP packet must match the inbound interface of the packet. **any**: Indicates that the URPF check is implemented in loose mode. The loose mode only requires that a forwarding entry can be found in the forwarding table based on the source address of a received IP packet. **allow-default**: (Optional) Indicates that the default route can be used for the URPF check.
Command Mode	Interface configuration mode
Usage Guide	Based on the source address of a received IP packet, URPF checks whether any route to the source address exists in the forwarding table and accordingly determines whether the packet is valid. If no forwarding entry is matched, the packet is determined as invalid. You can enable URPF in interface configuration mode to perform a URPF check on packets received on the interface. By default, the default route is not used for the URPF check. You can use the **allow-default** keyword to use the default route for the URPF check if necessary. By default, packets that fail in the URPF check will be dropped. A switch will enable URPF check on IPv4 Packets. A switch supports configuration of URPF on a routed port or L3 AP port. In addition, the following constraints exists: 1. After URPF is enabled on interfaces, a URPF check is performed on all packets received on physical ports corresponding to these interfaces, which increase the scope of packets checked by URPF. If a packet received on a tunnel port is also received on the preceding physical ports, the packet is also checked by URPF. In such a scenario, be cautious in enabling URPF. 2. After URPF is enabled, the route forwarding capacity of the device will be reduced by half. 3. After the URPF strict mode is enabled, if a packet received on an interface matches an equal-cost route during the URPF check, the packet will be processed according to the URPF loose mode.

Configuration Example

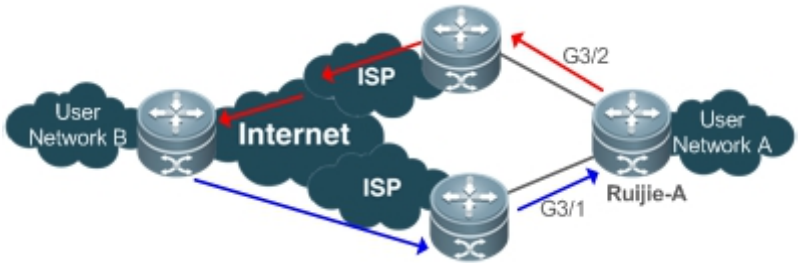
Configuring the Strict Mode

	Block the packets with spoofed sourced addresses at the access layer or aggregation layer to prevent sending these packets from PCs to the core network. To meet the preceding requirement, enable URPF in strict mode on the interface between the aggregation device and the access device.

Scenario **Figure 12- 3**	
Verification	As shown in Figure 12-3, enable URPF in strict mode on the aggregation devices, including FS A and FS B. The configurations are as follows:
FS-A	FS-A# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS-A (config)# interface gigabitEthernet0/1 FS-A (config-if-GigabitEthernet 0/1)#ip address 195.52.1.1 255.255.255.0 FS-A (config-if-GigabitEthernet 0/1)#ip verify unicast source reachable-via rx FS-A (config-if-GigabitEthernet 0/1)# ip verify urpf drop-rate notify FS-A (config-if-GigabitEthernet 0/1)#exit FS-A (config)# interface gigabitEthernet0/2 FS-A (config-if-GigabitEthernet 0/2)#ip address 195.52.2.1 255.255.255.0 FS-A (config-if-GigabitEthernet 0/2)#ip verify unicast source reachable-via rx FS-A (config-if-GigabitEthernet 0/2)# ip verify urpf drop-rate notify FS-A (config-if-GigabitEthernet 0/2)#exit
FS-B	FS-B# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS-B (config)# interface gigabitEthernet0/1 FS-B (config-if-GigabitEthernet 0/1)#ip address 195.52.3.1 255.255.255.0 FS-B (config-if-GigabitEthernet 0/1)#ip verify unicast source reachable-via rx FS-B (config-if-GigabitEthernet 0/1)# ip verify urpf drop-rate notify FS-B (config-if-GigabitEthernet 0/1)#exit FS-B (config)# interface gigabitEthernet0/2 FS-B (config-if-GigabitEthernet 0/2)#ip address 195.52.4.1 255.255.255.0

	FS-B (config-if-GigabitEthernet 0/2)#ip verify unicast source reachable-via rx FS-B (config-if-GigabitEthernet 0/2)# ip verify urpf drop-rate notify FS-B (config-if-GigabitEthernet 0/2)#exit
Verification	If source address spoofing exists on the network, run the show ip urpf command to display the number of spoofing packets dropped by URPF.
A	FS-A#show ip urpf interface gigabitEthernet 0/1 IP verify source reachable-via RX IP verify URPF drop-rate notify enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 124 Number of drop-rate notification counts in this interface is 0 FS-A#show ip urpf interface gigabitEthernet 0/2 IP verify source reachable-via RX IP verify URPF drop-rate notify enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 133 Number of drop-rate notification counts in this interface is 0
B	FS-B#show ip urpf interface gigabitEthernet 0/1 IP verify source reachable-via RX IP verify URPF drop-rate notify enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 124 Number of drop-rate notification counts in this interface is 0 FS-B#show ip urpf interface gigabitEthernet 0/2 IP verify source reachable-via RX IP verify URPF drop-rate notify enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 250 Number of drop-rate notification counts in this interface is 0

Configuring the Loose Mode

	On the egress device FS A of user network A, to prevent invalid packets from attacking the user network, enable URPF in loose mode on the outbound interfaces G3/1 and G3/2 that connect to two ISPs.
Scenario **Figure 12-4**	
FS-A	FS-A# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS-A (config)# interface gigabitEthernet3/1 FS-A (config-if-GigabitEthernet 3/1)# ip address 195.52.1.2 255.255.255.252 FS-A (config-if-GigabitEthernet 3/1)# ip verify unicast source reachable-via any FS-A (config-if-GigabitEthernet 3/1)# ip verify urpf drop-rate notify FS-A (config-if-GigabitEthernet 3/1)# exit FS-A (config)# interface gigabitEthernet3/2 FS-A (config-if-GigabitEthernet 3/2)# ip address 152.95.1.2 255.255.255.252 FS-A (config-if-GigabitEthernet 3/2)# ip verify unicast source reachable-via any FS-A (config-if-GigabitEthernet 3/2)# ip verify urpf drop-rate notify FS-A (config-if-GigabitEthernet 3/2)# end
Verification	If source address spoofing exists on the network, run the **show ip urpf** command to display the number of spoofing packets dropped by URPF.
A	FS #show ip urpf IP verify URPF drop-rate compute interval is 300s IP verify URPF drop-rate notify hold-down is 300s Interface gigabitEthernet3/1 IP verify source reachable-via ANY IP verify URPF drop-rate notify enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 4121 Number of drop-rate notification counts in this interface is 2 Interface gigabitEthernet3/2

	IP verify source reachable-via ANY IP verify URPF drop-rate notify enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 352 Number of drop-rate notification counts in this interface is 0

12.4.2 Configuring the Function of Monitoring the URPF Packet Loss Information

Configuration Effect

- After the function of monitoring the URPF packet loss information is enabled, the device can proactively send syslogs or trap messages to notify users of the packet loss information detected in the URPF check so that users can monitor the network status conveniently.

Notes

- URPF must be enabled.

Configuration Steps

Configuring the Calculation Interval of the URPF Packet Loss Rate

- Optional.
- Global configuration mode

Configuring the Alarm Interval of the URPF Packet Loss Rate

- Optional.
- Global configuration mode

Configuring the Function of Monitoring the URPF Packet Loss Information

- Optional.
- Interface configuration mode

Configuring the Threshold of the URPF Packet Loss Rate

- Optional.
- Interface configuration mode

Verification

Simulate a source address spoofing attack, enable URPF, and check as follows:

- Enable the alarm function. After the packet loss rate exceeds the threshold, check whether an alarm can be generated normally.

Related Commands

Configuring the Calculation Interval of the URPF Packet Loss Rate

Command	**ip verify urpf drop-rate compute interval** *seconds*
Parameter Description	**interval *seconds***: Indicates the calculation interval of the URPF packet loss rate. The unit is second. The value ranges from 30 to 300. The default value is 30s.
Command Mode	Global configuration mode
Usage Guide	The calculation interval of the URPF packet loss rate is configured in global configuration mode. The configuration is applied to the global and interface-based calculation of the URPF packet loss rate.

Configuring the Alarm Interval of the URPF Packet Loss Rate

Command	**ip verify urpf drop-rate notify hold-down** *seconds*
Parameter Description	**hold-down *seconds***: Indicates the alarm interval of the URPF packet loss rate. The unit is second. The value ranges from 30 to 300. The default value is 30s.
Command Mode	Global configuration mode
Usage Guide	The alarm interval of the URPF packet loss rate is configured in global configuration mode. The configuration is applied to the global and interface-based alarms of the URPF packet loss rate.

Configuring the Function of Monitoring the IPv4 URPF Packet Loss Information

Command	**ip verify urpf drop-rate notify**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	After the function of monitoring the URPF packet loss information is enabled, the device can proactively send syslogs or trap messages to notify users of the packet loss information detected in the URPF check so that users can monitor the network status conveniently.

Configuring the Threshold of the IPv4 URPF Packet Loss Rate

Command	**ip verify urpf notification threshold** *rate-value*
Parameter Description	**threshold *rate-value***: Indicates the threshold of the URPF packet loss rate. The unit is pps. The value ranges from 0 to 4,294,967,295. The default value is 1,000 pps.
Command Mode	Interface configuration mode
Usage Guide	If the threshold is 0, a notification is sent for every packet that is dropped because it fails in the URPF check. You can adjust the threshold based on the actual situation of the network.

Configuration Example

Setting the Calculation Interval of the URPF Packet Loss Rate to 120s

Configuration Steps	Set the calculation interval of the URPF packet loss rate to 120s in global configuration mode.

	FS#configure terminal FS(config)# ip verify urpf drop-rate compute interval 120 FS(config)# end
Verification	Run the **show ip urpf** command to check whether the configuration takes effect.
	FS# show ip urpf IP verify URPF drop-rate compute interval is 120s

Setting the Alarm Interval of the URPF Packet Loss Rate to 120s

Configuration Steps	Set the alarm interval of the URPF packet loss rate to 120s in global configuration mode.
	FS#configure terminal FS(config)# ip verify urpf drop-rate notify hold-down 120 FS(config)# end
Verification	Run the **show ip urpf** command to check whether the configuration takes effect.
	FS# show ip urpfIP verify URPF drop-rate notify hold-down is 120s

Enabling the Function of Monitoring the IPv4 URPF Packet Loss Information on the Interface GigabitEthernet 0/1

Configuration	Enable the function of monitoring the IPv4 URPF packet loss information on the interface GigabitEthernet 0/1.
	FS#configure terminal FS(config)# interface gigabitEthernet0/1 FS(config-if-GigabitEthernet 0/1)# ip verify unicast source reachable-via rx FS(config-if-GigabitEthernet 0/1)# ip verify urpf drop-rate notify
Verification	Run the **show ip urpf** command to check whether the function of monitoring the IPv4 URPF packet loss information is enabled on the interface GigabitEthernet 0/1.
	FS# show ip urpf interface gigabitEthernet 0/1 IP verify source reachable-via RX IP verify URPF drop-rate notify is enabled IP verify URPF notification threshold is 1000pps Number of drop packets in this interface is 0 Number of drop-rate notification counts in this interface is 0

Setting the Threshold of the IPv4 URPF Packet Loss Rate to 2,000 pps on the Interface GigabitEthernet 0/1

Configuration	Set the threshold of the IPv4 URPF packet loss rate to 2,000 pps on the interface GigabitEthernet 0/1.
	FS#configure terminal FS(config)# interface gigabitEthernet0/1 FS(config-if-GigabitEthernet 0/1)# ip verify unicast source reachable-via rx FS(config-if-GigabitEthernet 0/1)#ip verify urpf notification threshold 2000
Verification	Run the **show ip urpf** command to check the threshold of the IPv4 URPF packet loss rate and the threshold of the IPv6 URPF packet loss rate.
	FS# show ip urpf interface gigabitEthernet 0/1 IP verify source reachable-via RX IP verify URPF drop-rate notify is enabled IP verify URPF notification threshold is 2000pps Number of drop packets in this interface is 0 Number of drop-rate notification counts in this interface is 0

12.5 Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears statistics of the number of packets dropped during the IPv4 URPF check.	**clear ip urpf** [**interface** *interface-name*]

Displaying

Description	Command
Displays the IPv4 URPF configuration and statistics.	**show ip urpf** [**interface** *interface-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the URPF events.	**debug urpf event**
Debugs the URPF timers.	**debug urpf timer**

13 Configuring CPP

13.1 Overview

The CPU Protect Policy (CPP) provides policies for protecting the CPU of a switch.

In network environments, various attack packets spread, which may cause high CPU usages of the switches, affect protocol running and even difficulty in switch management. To this end, switch CPUs must be protected, that is, traffic control and priority-based processing must be performed for various incoming packets to ensure the processing capabilities of the switch CPUs.

CPP can effectively prevent malicious attacks in the network and provide a clean environment for legitimate protocol packets.

CPP is enabled by default. It provides protection during the entire operation of switches.

13.2 Applications

Application	Description
Preventing Malicious Attacks	When various malicious attacks such as ARP attacks intrude in a network, CPP divides attack packets into queues of different priorities so that the attack packets will not affect other packets.
Preventing CPU Processing Bottlenecks	Even when no attacks exist, it would become a bottleneck for CPU to handle excessive normal traffic. CPP can limit the rate of packets being sent to the CPU to ensure normal operation of switches.

13.2.1 Preventing Malicious Attacks

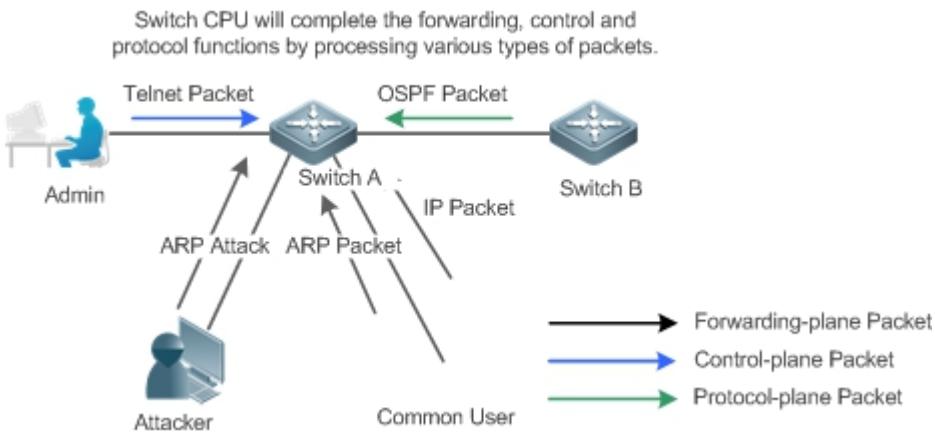
Scenario

Network switches at all levels may be attacked by malicious packets, typically ARP attacks.

As shown in Figure 13-1, switch CPUs process three types of packets: forwarding-plane, control-plane and protocol-plane. Forwarding-plane packets are used for routing, including ARP packets and IP route disconnection packets. Control-plane packets are used to manage services on switches, including Telnet packets and HTTP packets. Protocol-plane packets serve for running protocols, including BPDU packets and OSPF packets.

When an attacker initiates attacks by using ARP packets, the ARP packets will be sent to the CPU for processing. Since the CPU has limited processing capabilities, the ARP packets may force out other packets (which may be discarded) and consume many CPU resources (for processing ARP attack packets). Consequently, the CPU fails to work normally. In the scenario as shown in Figure 13-1, possible consequences include: common users fail to access the network; administrators fail to manage switches; the OSPF link between switch A and the neighbor B is disconnected and route learning fails.

Figure 13-1 Networking Topology of Switch Services and Attacks



Deployment

- By default, CPP classifies ARP packets, Telnet packets, IP route disconnection packets, and OSFP packets into queues of different priorities. In this way, ARP packets will not affect other packets.
- By default, CPP limits the rates of ARP packets and the rates of the priority queue where the ARP packets reside to ensure that the attack packets do not occupy too many CPU resources.
- Packets in the same priority queue with ARP packets may be affected by ARP attack packets. You can divide the packets and the ARP packets into different priority queues by means of configuration.
- When ARP attack packets exist, CPP cannot prevent normal ARP packets from being affected. CPP can only differentiate the packet type but cannot distinguish attack packets from normal packets of the same type. In this case, the Network Foundation Protection Policy (NFPP) function can be used to provide higher-granularity attack prevention.

For description of NFPP configurations, see the *Configuring NFPP*.

13.2.2 Preventing CPU Processing Bottlenecks

Scenario

Even though no attacks exist, many packets may need to be sent to the CPU for processing at an instant.

For example, the accesses to the core device of a campus network are counted in ten thousands. The traffic of normal ARP packets may reach dozens of thousands packets per second (PPS). If all packets are sent to the CPU for processing, the CPU resources cannot support the processing, which may cause protocol flapping and abnormal CPU running.

Deployment

- By default, the CPP function limits the rates of ARP packets and the rates of the priority queue where the APR packets reside to control the rate of ARP packets sent to the CPU and ensure that the CPU resource consumption is within a specified range and that the CPU can normally process other protocols.
- By default, the CPP function also limits the rates of other packets at the user level, such as Web authentication and 802.1X authentication packets.

13.3 Features

Basic Concepts

➘ QOS, DiffServ

Quality of Service (QoS) is a network security mechanism, a technology used to solve the problems of network delay and congestion.

DiffServ refers to the differentiated service model, which is a typical model implemented by QoS for classifying service streams to provide differentiated services.

➘ Bandwidth, Rate

Bandwidth refers to the maximum allowable data rate, which refers to the rate threshold in this document. Packets whose rates exceed the threshold will be discarded.

The rate indicates an actual data rate. When the rate of packets exceeds the bandwidth, packets out of the limit will be discarded. The rate must be equal to or smaller than the bandwidth.

The bandwidth and rate units in this document are packets per second (pps).

➘ L2, L3, L4

The structure of packets is hierarchical based on the TCP/IP model.

L2 refers to layer-2 headers, namely, the Ethernet encapsulation part; L3 refers to layer-3 headers, namely, the IP encapsulation part; L4 refers to layer-4 headers, usually, the TCP/UDP encapsulation part.

➘ Priority Queue, SP

Packets are cached inside a switch and packets in the output direction are cached in queues. Priority queues are mapped to Strict Priorities (SPs). Queues are not equal but have different priorities.

The SP is a kind of QoS scheduling algorithm. When a higher priority queue has packets, the packets in this queue are scheduled first. Scheduling refers to selecting packets from queues for output and refers to selecting and sending the packets to the CPU in this document.

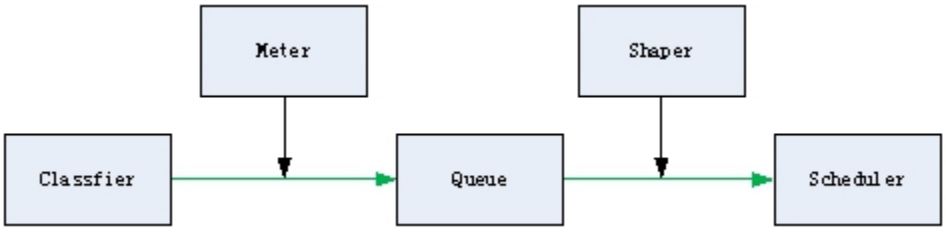
➘ CPU interface

Before sending packets to the CPU, a switch will cache the packets. The process of sending packets to the CPU is similar to the process of packet output. The CPU interface is a virtual interface. When packets are sent to the CPU, the packets will be output from this virtual interface. The priority queue and SP mentioned above are based on the CPU interface.

Overview

CPP protects the CPU by using the standard QoS DiffServ model.

Figure 13- 2 CPP Implementation Model



Feature	Description
Classfier	Classifies packet types and provides assurance for the subsequent implementation of QoS policies.
Meter	Limits rates based on packet types and controls the bandwidth for a specific packet type.
Queue	Queue packets to be sent to the CPU and select different queues based on packet types.
Scheduler	Selects and schedules queues to be sent to the CPU.
Shaper	Performs rate limit and bandwidth control on priority queues and the CPU interface.

13.3.1 Classifier

Working Principle

The Classifier classifies all packets to be sent to the CPU based on the L2, L3 and L4 information of the packets. Classifying packets is the basis for implementing QoS policies. In subsequent actions, different policies are implemented based on the classification to provide differentiated services. A switch provides fixed classification. The management function classifies packet types based on the protocols supported by the switch, for example, STP BPDU packets and ICMP packets. Packet types cannot be customized.

13.3.2 Meter

Working Principle

The Meter limits the rates of different packets based on the preset rate thresholds. You can set different rate thresholds for different packet types. When the rate of a packet type exceeds the corresponding threshold, the packets out of the limit will be discarded.

By using the Meter, you can control the rate of a packet type sent to the CPU within a threshold to prevent specific attack packets from exerting large impacts on the CPU resources. This is the level-1 protection of the CPP.

Related Configuration

- By default, each packet type corresponds to a rate threshold (bandwidth) and Meter policies are implemented based on the rate threshold.
- In application, you can run the **cpu-protect type** *packet-type* **bandwidth** *bandwidth-value* command to set Meter policies for specified packet types.

13.3.3 Queue

Working Principle

Queues are used to classify packets at level 2. You can select the same queue for different packet types; meanwhile, queues cache packets inside switches and provide services for the Scheduler and Shaper.

CPP queues are SP queues. The SPs of the packets are determined based on the time when they are added to a queue. Packets with a larger queue number have a higher priority.

Related Configuration

- By default, each packet type is mapped to an SP queue.
- In application, you can run the **cpu-protect type** *packet-type* **traffic-class** *traffic-class-num* command to select SP queues for specific packet types.

13.3.4 Scheduler

Working Principle

The Scheduler schedules packets based on SPs of queues. That is, packets in a queue with a higher priority are scheduled first.

Before being scheduled, packets to be sent to the CPU are cached in queues. When being scheduled, the packets are sent to the CPU for processing.

Only the SP scheduling policy is supported and cannot be modified.

13.3.5 Shaper

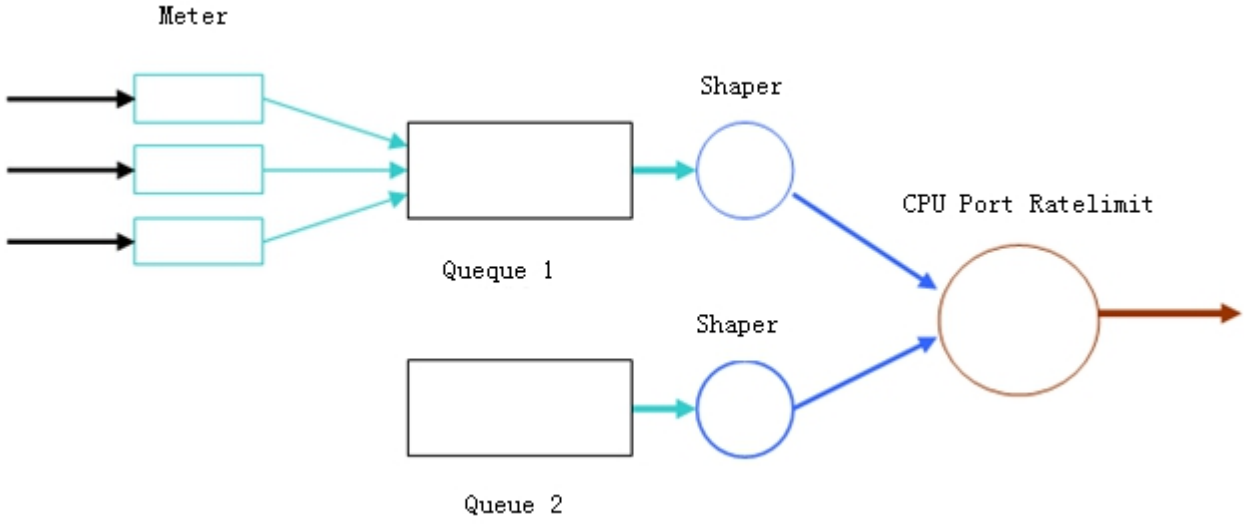
Working Principle

The Shaper is used to shape packets to be sent to the CPU, that is, when the actual rate of packets is greater than the shaping threshold, the packets must stay in the queue and cannot be scheduled. When packet rates fluctuate, the Shaper ensures that the rates of packets sent to the CPU are smooth (no more than the shaping threshold).

When the Shaper is available, packets in a queue with a lower priority may be scheduled before all packets in a queue with a higher priority are scheduled. If the rate of packets in a queue with certain priority exceeds the shaping threshold, scheduling of the packets in this queue may be stopped temporarily. Therefore, the Shaper can prevent packets in queues with lower priorities from starvation (which means that only packets in queues with higher priorities are scheduled and packets in queues with higher priorities are not scheduled).

Since the Shaper limits the scheduling rates of packets, it actually plays the rate limit function. The Shaper provides level-2 rate limit for priority queues and all packets sent to the CPU (CPU interface). The Shaper and Meter functions provide 3-level rate limit together and provide level-3 protection for the CPU.

Figure 13- 3 Level Rate Limit of the CPP



Related Configuration

Configuring the Shaper for priority queues

- By default, each priority queue determines a shaping threshold (bandwidth).
- In application, you can run the **cpu-protect traffic-class** *traffic-class-num* **bandwidth** *bandwidth_value* command to perform Shaper configuration for a specific priority queue.

Configuring the Shaper for the CPU Interface

- By default, the CPU interface determines a shaping threshold (bandwidth).
- Run the **cpu-protect cpu bandwidth** *bandwidth_value* command to perform Shaper configuration for the CPU interface.

13.4 Configuration

Configuration	Description and Command	
Configuring CPP	(Optional and configured by default) It is used to adjust the configuration parameters of CPP.	
	cpu-protect type *packet-type* **bandwidth**	Configures the Meter for a packet type.
	cpu-protect type *packet-type* **traffic-class**	Configures the priority queue for a packet type.
	cpu-protect traffic-class *traffic-class-num* **bandwidth**	Configures the Shaper for a priority queue.
	cpu-protect cpu bandwidth	Configures the Shaper for the CPU interface.

13.4.1 Configuring CPP

Configuration Effect

- By configuring the Meter function, you can set the bandwidth and rate limit for a packet type. Packets out of the limit will be directly discarded.
- By configuring the Queue function, you can select a priority queue for a packet type. Packets in a queue with a higher priority will be scheduled first.
- By configuring the Shaper function, you can set the bandwidth and rate limit for a CPU interface and a priority queue. Packets out of the limit will be directly discarded.

Notes

- Pay special attention when the bandwidth of a packet type is set to a smaller value, which may affect the normal traffic of the same type. To provide per-user CPP, combine the NFPP function.
- When the Meter and Shaper functions are combined, 3-level protection will be provided. Any level protection fights alone may bring negative effects. For example, if you want to increase the Meter of a packet type, you also need to adjust the Shaper of the corresponding priority queue. Otherwise, the packets of this type may affect other types of packets in the same priority queue.

Configuration Steps

Configuring the Meter for a packet type

- You can use or modify the default value but cannot disable it.
- You need to modify the configuration in the following cases: when packets of a type are not attackers but are discarded, you need to increase the Meter of this packet type. If attacks of a packet type cause abnormal CPU running, you need to decrease the Meter of this packet type.
- This configuration is available on all switches in a network environment.

Configuring the priority queue for a packet type

- You can use or modify the default value but cannot disable it.

- You need to modify the configuration in the following cases: When attacks of a packet type cause abnormality of other packets in the same queue, you can put the packet type in an unused queue. If a packet type cannot be discarded but the packet type is in the same queue with other packet types in use, you can put this packet type in a queue with a higher priority.
- This configuration is available on all switches in a network environment.

Configuring the Shaper for a priority queue

- You can use or modify the default value and cannot disable it.
- You need to modify the configuration in the following cases: If the Meter value of a packet type is greater which causes that other packets in the corresponding priority queue do not have sufficient bandwidth, you need to increase the Shaper for this priority queue. If attack packets are put in a priority queue and no other packets are in use, you need to increase the Shaper of this priority queue.
- This configuration is available on all switches in a network environment.

Configuring the Shaper for the CPU interface

- You can use or modify the default value and cannot disable it.
- You are not advised to change the Shaper of the CPU interface.
- This configuration is available on all switches in a network environment.

Verification

- Modify the configurations when the system runs abnormally, and view the system running after the modification to check whether the configurations take effect.
- Check whether the configurations take effect by viewing corresponding configurations and statistic values. For details, see the following commands.

Related Commands

Configuring the Meter for a packet type

Command	**cpu-protect type** *packet-type* **bandwidth** *bandwidth_value*
Parameter Description	*packet-type*: Specifies a packet type. Packet types are defined. *bandwidth_value***:** Sets the bandwidth, in the unit of packets per second (pps).
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the priority queue for a packet type

Command	**cpu-protect type** *packet-type* **traffic-class** *traffic-class-num*
Parameter Description	*packet-type*: Specifies a packet type. Packet types are defined. *traffic-class-num***:** Specifies a priority queue.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Shaper for a priority queue

Command	**cpu-protect traffic-class** *traffic-class-num* **bandwidth** *bandwidth_value*
Parameter Description	*traffic-class-num***:** Specifies a priority queue. *bandwidth_value***:** Sets the bandwidth, in the unit of pps.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Shaper for a CPU interface

Command	**cpu-protect cpu bandwidth** *bandwidth_value*
Parameter Description	*bandwidth_value***:** Sets the bandwidth, in the unit of pps.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Preventing packet attacks and network flapping by using CPP

Scenario	● ARP, IP, OSPF, dot1x, VRRP, Telnet and ICMP streams are available in the system. In the current configurations, ARP and 802.1X are in priority queue 2; IP, ICMP and Telnet streams are in priority queue 4; OSPF streams are in priority queue 3; VRRP streams are in priority queue 6. The Meter for each packet type is 10,000 pps; the shaper for each priority queue is 20,000 pps; the Shaper for the CPU interface is 100,000 pps. ● ARP attacks and IP scanning attacks exist in the system, which causes abnormal running of the system, authentication failure, Ping failure, management failure, and OSPF flapping.
Configuration Steps	● Put ARP attack packets in priority queue 1 and limit the bandwidth for ARP packets or the corresponding priority queue. ● Put OSPF packets in priority queue 5. ● Put IP Ping failure attack packets in priority queue 3 and limit the bandwidth for IP packets or the corresponding priority queue.
	FS# configure terminal FS(config)# cpu-protect type arp traffic-class 1 FS(config)# cpu-protect type arp bandwidth 5000 FS(config)# cpu-protect type ospf traffic-class 5 FS(config)# cpu-protect type v4uc-route traffic-class 3 FS(config)# cpu-protect type traffic-class 3 bandwidth 5000 FS(config)# end
Verification	Run the **show cpu-protect** command to view the configuration and statistics.
	FS#show cpu-protect %cpu port bandwidth: 100000(pps) Traffic-class Bandwidth(pps) Rate(pps) Drop(pps) ---------- ---------- --------- --------- 0 6000 0 0

1	6000	0	0
2	6000	0	0
3	6000	0	0
4	6000	0	0
5	6000	0	0
6	6000	0	0
7	6000	0	0

Packet Type	Traffic-class	Bandwidth(pps)	Rate(pps)	Drop(pps)	Total	Total Drop
bpdu	6	128	0	0	0	0
arp	1	3000	0	0	0	0
tpp	6	128	0	0	0	0
dot1x	2	1500	0	0	0	0
gvrp	5	128	0	0	0	0
rldp	5	128	0	0	0	0
lacp	5	256	0	0	0	0
rerp	5	128	0	0	0	0
reup	5	128	0	0	0	0
lldp	5	768	0	0	0	0
cdp	5	768	0	0	0	0
dhcps	2	1500	0	0	0	0
dhcps6	2	1500	0	0	0	0
dhcp6-client	2	1500	0	0	0	0
dhcp6-server	2	1500	0	0	0	0
dhcp-relay-c	2	1500	0	0	0	0
dhcp-relay-s	2	1500	0	0	0	0
option82	2	1500	0	0	0	0
tunnel-bpdu	2	128	0	0	0	0
tunnel-gvrp	2	128	0	0	0	0
unknown-v6mc	1	128	0	0	0	0
xgv6-ipmc	1	128	0	0	0	0
stargv6-ipmc	1	128	0	0	0	0
unknown-v4mc	1	128	0	0	0	0
xgv-ipmc	2	128	0	0	0	0
stargv-ipmc	2	128	0	0	0	0
udp-helper	1	128	0	0	0	0
dvmrp	4	128	0	0	0	0
igmp	2	1000	0	0	0	0
icmp	3	1600	0	0	0	0
ospf	4	2000	0	0	0	0
ospf3	4	2000	0	0	0	0
pim	4	1000	0	0	0	0
pimv6	4	1000	0	0	0	0
rip	4	128	0	0	0	0

ripng	4	128	0	0	0	0
vrrp	6	256	0	0	0	0
vrrpv6	6	256	0	0	0	0
ttl0	0	128	0	0	0	0
ttl1	0	2000	0	0	0	0
hop-limit	0	800	0	0	0	0
local-ipv4	3	4000	0	0	0	0
local-ipv6	3	4000	0	0	0	0
v4uc-route	1	800	0	0	0	0
v6uc-route	1	800	0	0	0	0
rt-host	4	3000	0	0	0	0
mld	2	1000	0	0	0	0
nd-snp-ns-na	1	3000	0	0	0	0
nd-snp-rs	1	1000	0	0	0	0
nd-snp-ra-redirect	1	1000	0	0	0	0
erps	5	128	0	0	0	0
mpls-ttl0	4	128	0	0	0	0
mpls-ttl1	4	128	0	0	0	0
mpls-ctrl	4	128	0	0	0	0
isis	4	2000	0	0	0	0
bgp	4	2000	0	0	0	0
cfm	5	512	0	0	0	0
web-auth	2	2000	0	0	0	0
fcoe-fip	4	1000	0	0	0	0
fcoe-local	4	1000	0	0	0	0
bfd	6	5120	0	0	0	0
micro-bfd	6	5120	0	0	0	0
micro-bfd-v6	6	5120	0	0	0	0
dldp	6	3200	0	0	0	0
other	0	4096	0	0	0	0
trill	4	1000	0	0	0	0
efm	5	1000	0	0	0	0
ipv6-all	0	2000	0	0	0	0
ip-option	0	800	0	0	0	0
mgmt	-	4000	4	0	4639	0
dns	2	200	0	0	0	0
sdn	0	5000	0	0	0	0
sdn_of_fetch	0	5000	0	0	0	0
sdn_of_copy	0	5000	0	0	0	0
sdn_of_trap	0	5000	0	0	0	0
vxlan-non-uc	1	512	0	0	0	0
local-telnet	3	1000	0	0	0	0
local-snmp	3	1000	0	0	0	0
local-ssh	3	1000	0	0	0	0

13.5 Monitoring

Clearing

Description	Command
Clears the CPP statistics.	**clear cpu-protect counters** [**device** *device_num*]
Clears the CPP statistics on the master device.	**clear cpu-protect counters mboard**

Displaying

Description	Command
Displays the configuration and statistics of a packet type.	**show cpu-protect type** *packet-type* [**device** *device_num*]
Displays the configuration and statistics of a priority queue.	**show cpu-protect traffic-class** *traffic-class-num* [**device** *device_num*]
Displays the configuration on a CPU interface.	**show cpu-protect cpu**
Displays all configurations and statistics on the master device.	**show cpu-protect** {**mboard** | **summary** }
Displays all configurations and statistics of CPP.	**show cpu-protect** [**device** *device_num*]

Debugging

N/A

The preceding monitoring commands are available on both chassis and cassette devices in either the standalone mode or the Stacking mode.

If the **device** value is not specified, the **clear** command is used to clear the statistics of all nodes in the system and the **show** command is used to display the configurations on the master device.

In the Stacking mode, the parameter **device** indicates a chassis or cassette device. If the **device** value is not specified, it indicates the master chassis or the master device.

14 Configuring DHCP Snooping

14.1 Overview

DHCP Snooping: DHCP Snooping snoops DHCP interactive packets between clients and servers to record and monitor users' IP addresses and filter out illegal DHCP packets, including client request packets and server response packets. The legal user database generated from DHCP Snooping records may serve security applications like IP Source Guard.

Protocols and Standards

- RFC 2131: Dynamic Host Configuration Protocol
- RFC 2132: DHCP Options and BOOTP Vendor Extensions

14.2 Applications

Application	Description
Guarding against DHCP service spoofing	In a network with multiple DHCP servers, DHCP clients are allowed to obtain network configurations only from legal DHCP servers.
Guarding against DHCP packet flooding	Malicious network users may frequently send DHCP request packets.
Guarding against forged DHCP packets	Malicious network users may send forged DHCP request packets, for example, DHCP-RELEASE packets.
Guarding against IP/MAC spoofing	Malicious network users may send forged IP packets, for example, tampered source address fields of packets.
Preventing Lease of IP Addresses	Network users may lease IP addresses rather than obtaining them from a DHCP server.
Detecting ARP attack	Malicious users forge ARP response packets to intercept packets during normal users' communication.

14.2.1 Guarding Against DHCP Service Spoofing

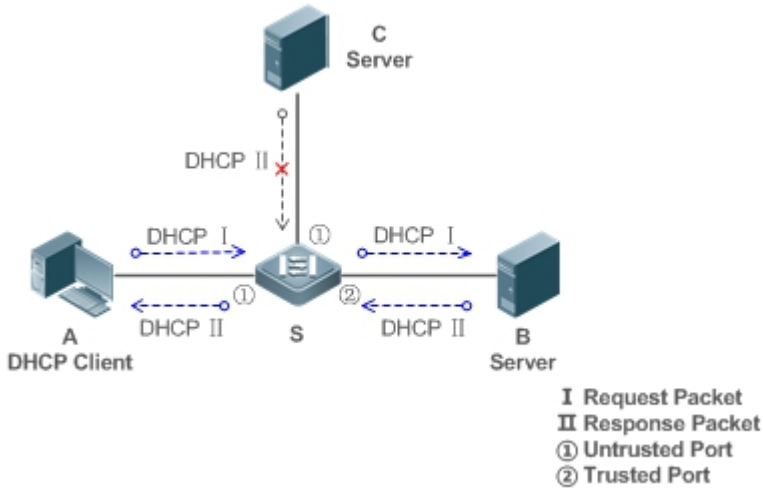
Scenario

Multiple DHCP servers may exist in a network. It is essential to ensure that user PCs obtain network configurations only from the DHCP servers within a controlled area.

Take the following figure as an example. The DHCP client can only communicate with trusted DHCP servers.

- Request packets from the DHCP client can be transmitted only to trusted DHCP servers.
- Only the response packets from trusted DHCP servers can be transmitted to the client.

Figure 14- 1



Remarks:	S is an access device. A is a user PC. B is a DHCP server within the controlled area. C is a DHCP server out of the controlled area.

Deployment

- Enable DHCP Snooping on S to realize DHCP packet monitoring.
- Set the port on S connecting to B as trusted to transfer response packets.
- Set the rest of ports on S as untrusted to filter response packets.

14.2.2 Guarding Against DHCP Packet Flooding

Scenario

Potential malicious DHCP clients in a network may send high-rate DHCP packets. As a result, legitimate users cannot obtain IP addresses, and access devices are highly loaded or even break down. It is necessary to take actions to ensure network stability.

With the DHCP Snooping rate limit function for DHCP packets, a DHCP client can only send DHCP request packets at a rate below the limit.

- The request packets from a DHCP client are sent at a rate below the limit.
- Packets sent at rates beyond the limit will be discarded.
- Enable DHCP Snooping correlation with ARP, and delete the non-existing entries.

Deployment

- Enable DHCP Snooping on S to realize DHCP monitoring.
- Limit the rates of DHCP packets from the untrusted ports.
- Enable DHCP Snooping correlation with ARP, and detect whether the user is online.

14.2.3 Guarding Against Forged DHCP Packets

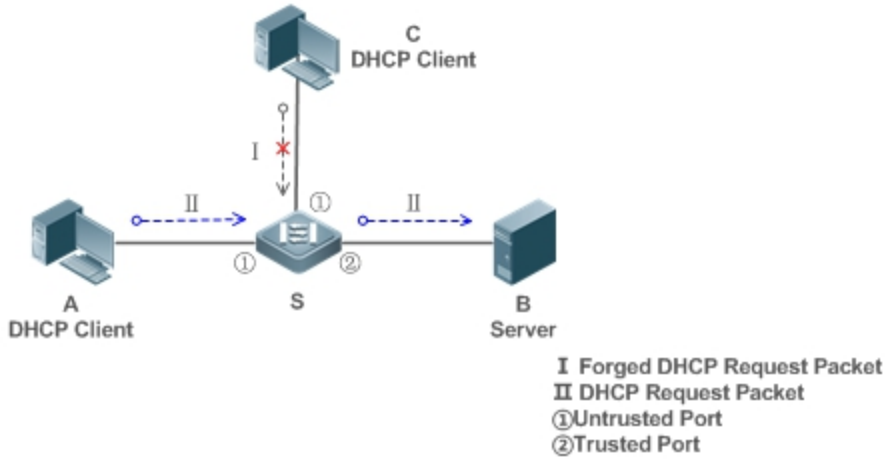
Scenario

Potential malicious clients in a network may forge DHCP request packets, consuming applicable IP addresses from the servers and probably preempting legal users' IP addresses. Therefore, it is necessary to filter out illegal DHCP packets.

For example, as shown in the figure below, the DHCP request packets sent from DHCP clients will be checked.

- The source MAC address fields of the request packets from DHCP clients must match the **chaddr** fields of DHCP packets.
- The Release packets and Decline packets from clients must match the entries in the DHCP Snooping binding database.

Figure 14- 2



Remarks:	S is an access device. A and C are user PCs. B is a DHCP server within the controlled area.

Deployment

- Enable DHCP Snooping on S to realize DHCP monitoring.
- Set the port on S connecting to B as trusted to transfer response packets.
- Set the rest of ports on S as untrusted to filter response packets.
- Enable DHCP Snooping Source MAC Verification on untrusted ports of S to filter out illegal packets.

12.2.4 Guarding Against IP/MAC Spoofing

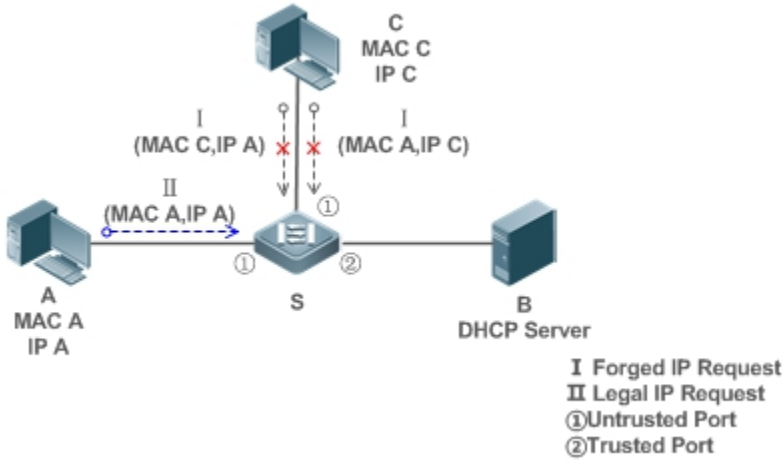
Scenario

Check IP packets from untrusted ports to filter out forged IP packets based on IP or IP-MAC fields.

For example, in the following figure, the IP packets sent by DHCP clients are validated.

- The source IP address fields of IP packets must match the IP addresses assigned by DHCP.
- The source MAC address fields of layer-2 packets must match the **chaddr** fields in DHCP request packets from clients.

Figure 14- 3



Remarks:	S is an access device. A and C are user PCs. B is a DHCP server within the controlled area.

Deployment

- Enable DHCP Snooping on S to realize DHCP monitoring.
- Set all downlink ports on the S as DHCP Snooping untrusted.
- Enable IP Source Guard on S to filter IP packets.
- Enable IP Source Guard in IP-MAC based mode to check the source MAC and IP address fields of IP packets.

12.2.5 Preventing Lease of IP Addresses

Scenario

Validate the source addresses of IP packets from untrusted ports compared with DHCP-assigned addresses.

If the source addresses, connected ports, and layer-2 source MAC addresses of ports in IP packets do not match the assignments of the DHCP server, such packets will be discarded.

The networking topology scenario is the same as that shown in the previous figure.

Deployment

- The same as that in the section "Guarding Against IP/MAC Spoofing".

12.2.6 Detecting ARP Attacks

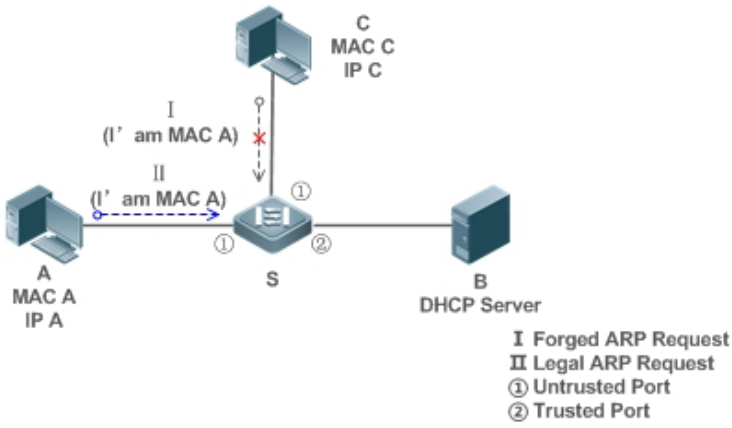
Scenario

Check the ARP packets from untrusted ports and filter out the ARP packets unmatched with the assignments of the DHCP server.

For example, in the following figure, the ARP packets sent from DHCP clients will be checked.

- The ports receiving ARP packets, the layer-2 MAC addresses, and the source MAC addresses of ARP packets senders shall be consistent with the DHCP Snooping histories.

Figure 14- 4



Remarks:	S is an access device. A and C are user PCs. B is a DHCP server within the controlled area.

Deployment

- Enable DHCP Snooping on S to realize DHCP monitoring.
- Set all downlink ports on the S as untrusted.
- Enable IP Source Guard and ARP Check on all the untrusted ports on S to realize ARP packet filtering.

All the above security control functions are only effective to DHCP Snooping untrusted ports.

14.3 Features

Basic Concepts

DHCP Request Packets

Request packets are sent from a DHCP client to a DHCP server, including DHCP-DISCOVER packets, DHCP-REQUEST packets, DHCP-DECLINE packets, DHCP-RELEASE packets and DHCP-INFORM packets.

DHCP Response Packets

Response packets are sent from a DHCP server to a DHCP client, including DHCP-OFFER packets, DHCP-ACK packets and DHCP-NAK packets.

DHCP Snooping Trusted Ports

IP address request interaction is complete via broadcast. Therefore, illegal DHCP services will influence normal clients' acquisition of IP addresses and lead to service spoofing and stealing. To prevent illegal DHCP services, DHCP Snooping ports are divided into two types: trusted ports and untrusted ports. The access devices only transmit DHCP response packets received on trusted ports, while such packets from untrusted ports are discarded. In this way, we may configure the ports connected to a legal DHCP Server as trusted and the other ports as untrusted to shield illegal DHCP Servers.

On switches, all switching ports or layer-2 aggregate ports are defaulted as untrusted, while trusted ports can be specified.

DHCP Snooping Packet Suppression

To shield all the DHCP packets on a specific client, we can enable DHCP Snooping packet suppression on its untrusted ports.

VLAN-based DHCP Snooping

DHCP Snooping can work on a VLAN basis. By default, when DHCP Snooping is enabled, it is effective to all the VLANs of the current client. Specify VLANs help control the effective range of DHCP Snooping flexibly.

DHCP Snooping Binding Database

In a DHCP network, clients may set static IP addresses randomly. This increases not only the difficulty of network maintenance but also the possibility that legal clients with IP addresses assigned by the DHCP server may fail to use the network normally due to address conflict. Through snooping packets between clients and servers, DHCP Snooping summarizes the user entries including IP addresses, MAC address, VLAN ID (VID), ports and lease time to build the DHCP Snooping binding database. Combined with ARP detection and ARP check, DHCP Snooping controls the reliable assignment of IP addresses for legal clients.

DHCP Snooping Rate Limit

DHCP Snooping rate limit function can be configured through the rate limit command of Network Foundation Protection Policy (NFPP). For NFPP configuration, see the *Configuring NFPP*.

DHCP Option82

DHCP Option82, an option for DHCP packets, is also called DHCP Relay Agent Information Option. As the option number is 82, it is known as Option82. Option82 is developed to enhance the security of DHCP servers and improve the strategies of IP address assignment. The option is often configured for the DHCP relay services of a network access device like DHCP Relay and DHCP Snooping. This option is transparent to DHCP clients, and DHCP relay components realize the addition and deduction of the option.

Illegal DHCP Packets

Through DHCP Snooping, validation is performed on the DHCP packets passing through a client. Illegal DHCP packets are discarded, user information is recorded into the DHCP Snooping binding database for further applications (for example, ARP detection). The following types of packets are considered illegal DHCP packets.

- The DHCP response packets received on untrusted ports, including DHCP-ACK, DHCP-NACK and DHCP-OFFER packets
- The DHCP request packets carrying gateway information **giaddr**, which are received on untrusted ports
- When MAC verification is enabled, packets with source MAC addresses different with the value of the **chaddr** field in DHCP packets
- DHCP-RELEASE packets with the entry in the DHCP Snooping binding database Snooping while with untrusted ports inconsistent with settings in this binding database
- DHCP packets in wrong formats, or incomplete

Overview

Feature	Description
Filtering DHCP packets	Perform legality check on DHCP packets and discard illegal packets (see the previous section for the introduction of illegal packets). Transfer requests packets received on trusted ports only.
Building the DHCP Snooping binding database	Snoop the interaction between DHCP clients and the server, and generate the DHCP Snooping binding database to provide basis for other filtering modules.

14.3.1 Filtering DHCP Packets

Perform validation on DHCP packets from untrusted ports. Filter out the illegal packets as introduced in the previous section "Basic Concepts".

Working Principle

During snooping, check the receiving ports and the packet fields of packets to realize packet filtering, and modify the destination ports of packets to realize control of transmit range of the packets.

Checking Ports

In receipt of DHCP packets, a client first judges whether the packet receiving ports are DHCP Snooping trusted ports. If yes, legality check and binding entry addition are skipped, and packets are transferred directly. For not, both the check and addition are needed.

Checking Packet Encapsulation and Length

A client checks whether packets are UDP packets and whether the destination port is 67 or 68. Check whether the packet length match the length field defined in protocols.

Checking Packet Fields and Types

According to the types of illegal packet introduced in the section "Basic Concepts", check the fields **giaddr** and **chaddr** in packets and then check whether the restrictive conditions for the type of the packet are met.

Related Configuration

Enabling Global DHCP Snooping

By default, DHCP Snooping is disabled.

It can be enabled on a device using the **ip dhcp snooping** command.

Global DHCP Snooping must be enabled before VLAN-based DHCP Snooping is applied.

Configuring VLAN-based DHCP Snooping

By default, when global DHCP Snooping is effective, DHCP Snooping is effective to all VLANs.

Use the [**no**] **ip dhcp snooping vlan** command to enable DHCP Snooping on specified VLANs or delete VLANs from the specified VLANs. The value range of the command parameter is the actual range of VLAN numbers.

Configuring DHCP Snooping Source MAC Verification

By default, the layer-2 MAC addresses of packets and the **chaddr** fields of DHCP packets are not verified.

When the **ip dhcp snooping verify mac-address** command is used, the source MAC addresses and the **chaddr** fields of the DHCP request packets sent from untrusted ports are verified. The DHCP request packets with different MAC addresses will be discarded.

14.3.2 Building the Binding Database

DHCP Snooping detects the interactive packets between DHCP clients and the DHCP server, and generate entries of the DHCP Snooping binding database according to the information of legal DHCP packets. All these legal entries are provided to other security modules of a client as the basis of filtering packets from network.

Working Principle

During snooping, the binding database is updated timely based on the types of DHCP packets.

Generating Binding Entries

When a DHCP-ACK packet on a trusted port is snooped, the client's IP address, MAC address, and lease time field are extracted together with the port ID (a wired interface index) and VLAN ID. Then, a binding entry of it is generated.

Deleting Binding Entries

When the recorded lease time of a binding entry is due, it will be deleted if a legal DHCP-RELEASE/DHCP-DECLINE packet sent by the client or a DHCP-NCK packet received on a trusted port is snooped, or the **clear** command is used.

Related Configuration

No configuration is needed except enabling DHCP Snooping.

14.4 Configuration

Configuration	Description and Command	
Configuring basic functions of DHCP Snooping	(Mandatory) It is used to enable DHCP Snooping.	
	ip dhcp snooping	Enables DHCP Snooping.
	ip dhcp snooping suppression	Enables DHCP Snooping packet suppression.
	ip dhcp snooping vlan	Enables VLAN-based DHCP Snooping.
	ip dhcp snooping verify mac-address	Configures DHCP Snooping source MAC verification.
	ip dhcp snooping database write-delay	Writes the DHCP Snooping binding database to Flash periodically.
	ip dhcp snooping database write-to-flash	Writes the DHCP Snooping binding database to Flash manually.
	renew ip dhcp snooping database	Imports Flash storage to the DHCP Snooping Binding database.
	ip dhcp snooping trust	Configures DHCP Snooping trusted ports.
	ip dhcp snooping bootp	Enables BOOTP support.
	ip dhcp snooping check-giaddr	Enables DHCP Snooping to support the function of processing Relay requests.
	ip dhcp snooping loose-forward	Enables loose forwarding.
Configuring Option82	(Optional)It is used to optimize the address assignment by DHCP servers.	
	ip dhcp snooping Information option	Adds Option82 functions to DHCP request packets.
	ip dhcp snooping information option format remote-id	Configures the sub-potion **remote-id** of Option82 as a user-defined character string.
	ip dhcp snooping vlan information option format-type circuit-id string	Configures the sub-option **circuit-id** of Option82 as a user-defined character string.
	ip dhcp snooping information option strategy	Congifures the strategy of Option82.

14.4.1 Configuring Basic Features

Configuration Effect

- Enable DHCP Snooping.
- Generate the DHCP Snooping binding database.
- Control the transmit range of DHCP packets.
- Filter out illegal DHCP packets.

Notes

- The ports on clients connecting a trusted DHCP server must be configured as trusted.
- DHCP Snooping is effective on the wired switching ports, layer-2 aggregate ports, and layer-2 encapsulation sub-interfaces.
- DHCP Snooping and DHCP Relay are mutually exclusive in VRF scenarios.

Configuration Steps

Enabling Global DHCP Snooping

- Mandatory.
- Unless otherwise noted, the feature should be configured on access devices.

Enabling or Disabling VLAN-based DHCP Snooping

- DHCP Snooping can be disabled if not necessary for some VLANs.
- Unless otherwise noted, the feature should be configured on access devices.

Configuring DHCP Snooping Trusted Ports

- Mandatory.
- Configure the ports connecting a trusted DHCP server as trusted.

Enabling DHCP Snooping Source MAC Validation

- This configuration is required if the **chaddr** fields of DHCP request packets match the layer-2 source MAC addresses of data packets.
- Unless otherwise noted, the feature should be enabled on all the untrusted ports of access devices.

Writing the DHCP Snooping Binding Database to Flash Periodically

- Enable this feature to timely save the DHCP Snooping binding database information in case that client reboot.
- Unless otherwise noted, the feature should be configured on access devices.

Enabling BOOTP Support

- Optional
- Unless otherwise noted, the feature should be configured on access devices.

Enabling DHCP Snooping to Process Relay Requests

- Optional.
- Unless otherwise noted, the feature should be enabled on access devices.

Enabling Loose Forwarding

- Optional.
- Unless otherwise noted, the feature is disabled.

Verification

Configure a client to obtain network configurations through the DHCP protocol.

- Check whether the DHCP Snooping Binding database is generated with entries on the client.

Related Commands

Enabling or Disabling DHCP Snooping

Command	[**no**] **ip dhcp snooping**
Parameter Description	**N/A**
Command Mode	Global configuration mode
Usage Guide	After global DHCP Snooping is enabled, you can check DHCP Snooping using the **show ip dhcp snooping** command.

Configuring VLAN-based DHCP Snooping

Command	[**no**] **ip dhcp snooping vlan** { *vlan-rng* | {*vlan-min* [*vlan-max*] } }
Parameter Description	*vlan-rng*: Indicates the range of VLANs *vlan-min*: The minimum VLAN ID *vlan-max*: The maximum VLAN ID
Command Mode	Global configuration mode
Usage Guide	Use this command to enable or disable DHCP Snooping on specified VLANs. This feature is available only after global DHCP Snooping is enabled.

Configuring DHCP Snooping Packet Suppression

Command	[**no**] **ip dhcp snooping suppression**
Parameter Description	**N/A**
Command Mode	Interface configuration mode
Usage Guide	Use this command to reject all DHCP request packets at the port, that is, to forbid all users under the port to apply for addresses via DHCP.

Configuring DHCP Snooping Source MAC Verification

Command	[**no**] **ip dhcp snooping verify mac-address**
Parameter	**N/A**

Description	
Command Mode	Global configuration mode
Usage Guide	Through the source MAC address verification, the MAC addresses in link headers and the CLIENT MAC fields in the request packets sent by a DHCP CLIENT are checked for consistence. When the source MAC address verification fails, packets will be discarded.

Writing DHCP Snooping Database to Flash Periodically

Command	[**no**] **ip dhcp snooping database write-delay** [*time*]
Parameter Description	*time*: Indicates the interval between two times of writing the DHCP Snooping database to the Flash.
Command Mode	Global configuration mode
Usage Guide	Use this command to write the DHCP Snooping database to FLASH document. This can avoid binding information loss which requires re-obtaining IP addresses to resume communication after the device restarts.

Writing the DHCP Snooping Database to Flash Manually

Command	**ip dhcp snooping database write-to-flash**
Parameter Description	**N/A**
Command Mode	Global configuration mode
Usage Guide	Use this command to write the dynamic user information in the DHCP Snooping database in FLASH documents in real time. If a device is upgraded from a non-QinQ version to a QinQ version (or vice versa), binding entries cannot be restored from FLASH documents because of version differences between FLASH documents.

Importing Backep File Storage to the DHCP Snooping Binding Database

Command	**renew ip dhcp snooping database**
Parameter Description	**N/A**
Command Mode	Privileged configuration mode
Usage Guide	Use this command to import the information from backup file to the DHCP Snooping binding database.

Configuring DHCP Snooping Trusted Ports

Command	[**no**] **ip dhcp snooping trust**
Parameter Description	**N/A**
Command Mode	Interface configuration mode

Usage Guide	Use this command to configure a port connected to a legal DHCP server as a trusted port. The DHCP response packets received by trusted ports are transferred, while those received by untrusted ports are discarded.

Enabling or Disabling BOOTP Support

Command	[**no**] **ip dhcp snooping bootp**
Parameter Description	**N/A**
Command Mode	Global configuration mode
Usage Guide	Use this command to support the BOOPT protocol.

Enabling DHCP Snooping to Process Relay Requests

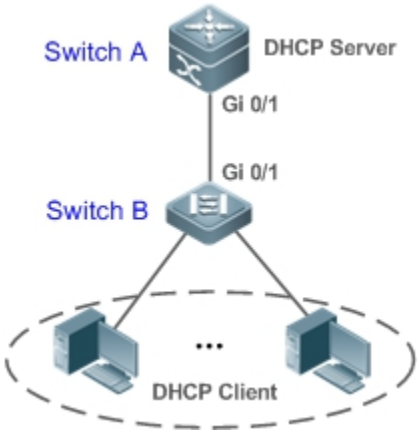
Command	[**no**] **ip dhcp snooping check-giaddr**
Parameter Description	**N/A**
Command Mode	Global configuration mode
Usage Guide	After the feature is enabled, services using DHCP Snooping binding entries generated based on Relay requests, such as IP Source Guard/802.1x authentication, cannot be deployed. Otherwise, users fail to access the Internet. After the feature is enabled, the **ip dhcp snooping verify mac-address** command cannot be used. Otherwise, DHCP Relay requests will be discarded and as a result, users fail to obtain addresses.

Enabling DHCP Snooping Loose Forwarding

Command	**ip dhcp snooping loose-forward**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After this feature is enabled, when the capacity of DHCP Snooping binding entries is reached, DHCP packets of new users are forwarded and obtain addresses, but DHCP Snooping does not record binding entries of new users.

Configuration Example

DHCP Client Obtaining IP addresses Dynamically from a Legal DHCP Server

Scenario **Figure 14- 5**	
Configuration Steps	● Enable DHCP Snooping on an access device (Switch B in this case). ● Configure the uplink port (port Gi 0/1 in this case) as a trusted port.
B	B#configure terminal Enter configuration commands, one per line. End with CNTL/Z. B(config)#ip dhcp snooping B(config)#interface gigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)#ip dhcp snooping trust B(config-if-GigabitEthernet 0/1)#end
Verification	Check the configuration on Switch B. ● Check whether DHCP Snooping is enabled, and whether the configured DHCP Snooping trusted port is uplink. ● Check the DHCP Snooping configuration on Switch B, and especially whether the trusted port is correct.
B	B#show running-config ! ip dhcp snooping ! interface GigabitEthernet 0/1 B#show ip dhcp snooping Switch DHCP Snooping status : ENABLE DHCP Snooping Verification of hwaddr status : DISABLE DHCP Snooping database write-delay time : 0 seconds DHCP Snooping option 82 status : DISABLE DHCP Snooping Support BOOTP bind status : DISABLE Interface Trusted Rate limit (pps) ---------- ------- ---------- GigabitEthernet 0/1 YES unlimited B#show ip dhcp snooping binding Total number of bindings: 1 MacAddress IpAddress Lease(sec) Type VLAN Interface ---------- ---------- ---------- ---------- ----- ---------- 0013.2049.9014 172.16.1.2 86207 DHCP-Snooping 1 GigabitEthernet 0/11

Common Errors

- The uplink port is not configured as a DHCP trusted port.
- Another access security option is already configured for the uplink port, so that a DHCP trusted port cannot be configured.

14.4.2 Configuring Option82

Configuration Effect

- Enable a DHCP server to obtain more information and assign addresses better.
- The Option82 function is client-oblivious.

Notes

- The Opion82 functions for DHCP Snooping and DHCP Relay are mutually exclusive.

Configuration Steps

- To realize optimization of address allocation, implement the configuration.
- Unless otherwise noted, enable this function on access devices with DHCP Snooping enabled.

Verification

Check whether the DHCP Snooping configuration options are configured successfully.

Related Commands

Adding Option82 to DHCP Request Packets

Command	[**no**] **ip dhcp snooping information option** [**standard-format | dot1x-format**]
Parameter Description	**standard-format**: Indicates a standard format of the Option82 options **dot1x-format**: Indicates a dot1x format of the Option82 options
Command Mode	Global configuration mode
Usage Guide	Use this command to add Option82 to DHCP request packets so that a DHCP server assigns addresses according to such information. When dot1x-format is used, if DHCP Relay is configured on the local device and the uplink port connected to the DHCP Server is an SVI port, DHCP Snooping needs to be disabled in the VLAN to which the SVI port belongs.

Configuring Sub-option remote-id of Option82 as User-defined Character String

Command	[**no**] **ip dhcp snooping information option format remote-id** { **string** *ASCII-string* | **hostname** }
Parameter Description	**string** *ASCII-string*: Indicates the content of the extensible format, the Option82 option **remote-id**, is a user-defined character string **hostname**: Indicates the content of the extensible format, the Option82 option **remote-id**, is a host name.
Configuration mode	Global configuration mode
Usage Guide	Use this command to configure the sub-option **remote-id** of the Option82 as user-defined content, which is added to DHCP request packets. A DHCP server assigns addresses according to Option82 information.

Configuring Sub-Option circuit -id of Option82 as User-defined Character String

Command	[**no**] **ip dhcp snooping vlan** *vlan-id* **information option format-type circuit-id string** *ascii-string*
Parameter Description	*vlan-id*: Indicates the VLAN where a DHCP request packet is *ascii-string*: Indicates the user-defined string
Configuration mode	Interface configuration mode
Usage Guide	Use this command to configure the sub-option **circuit-id** of the Option82 as user-defined content, which is added to DHCP request packets. A DHCP server assigns addresses according to Option82 information.

Configuring the Strategy of Option82

Command	**ip dhcp snooping information option strategy {keep | drop | replace}**
Parameter Description	**keep:** Indicates reception of request packets with Option82. Option82 is kept and the packets are forwarded. **drop:** Indicates reception of request packets with Option82. The packets are dropped. **replace: I**ndicates reception of request packets with Option82. Option82 of the packets are replaced with Option82 configured latest. The packets are forwarded.
Configuration mode	Global configuration mode
Usage Guide	This command only works for request packets with Option82. If the strategy is keep or drop, there is no need to configure sub-option **circuit-id** of the Option82. If the strategy is replace, sub-option **circuit-id** of the Option82 needs configuring. In terms of request packets without Option82, configured sub-option **circuit-id** of the Option82 is added.

Configuration Example

Configuring Option82 to DHCP Request Packets

Configuration Steps	● Configuring basic functions of DHCP Snooping. ● Configuring Option82.
B	FS# configure terminal FS(config)# ip dhcp snooping information option FS(config)# end
Verification	Check the DHCP Snooping configuration.
B	B#show ip dhcp snooping Switch DHCP Snooping status : ENABLE DHCP Snooping Verification of hwaddr status : DISABLE DHCP Snooping database write-delay time : 0 seconds DHCP Snooping option 82 status : ENABLE DHCP Snooping Support bootp bind status : DISABLE Interface Trusted Rate limit (pps) ---------- ------- ---------- GigabitEthernet 0/1 YES unlimited

Common Errors

- N/A

14.5 Monitoring

Clearing

Running the clear commands may lose vital information and thus interrupt services.

Description	Command
Clears dynamic user inforamtion of DHCP Snooping database.	**clear ip dhcp snooping binding** [*ip*] [*mac*] [**vlan** *vlan-id*] [**interface** *interface-id*]

Displaying

Description	Command
Displays DHCP Snooping configuration.	**show ip dhcp snooping**
Displays the DHCP Snooping binding database.	**show ip dhcp snooping binding**

Debugging

System resources are occupied when debugging information is output. Disable the debugging switch immediately after use.

Description	Command
Debugs DHCP Snooping events.	**debug snooping ipv4 event**
Disables debugging DHCP Snooping events.	**no debug snooping ipv4 event**
Debugs DHCP Snooping packets.	**debug snooping ipv4 packet**
Disables debugging DHCP Snooping packets.	**no debug snooping ipv4 packet**
Enables debugging MAC-based DHCP Snooping.	**debug snooping ipv4 mac-address** *H.H.H*
Disables debugging MAC-based DHCP Snooping.	**no debug snooping ipv4 mac-address** *H.H.H*
Enables debugging all DHCP Snooping	**debug snooping ipv4 all**
Disables debugging all DHCP Snooping	**no debug snooping ipv4 all**

15 Configuring DHCPv6 Snooping

15.1 Overview

DHCPv6 Snooping: Dynamic Host Configuration Protocol version 6 (DHCPv6) snooping enables recording and monitoring of IPv6 address usage by snooping DHCPv6 packets exchanged between the client and the server, and filters illegal DHCPv6 packets, including request packets from the client and response packets from the server. The user data entries generated by DHCPv6 snooping recording can serve security applications such as IPv6 Source Guard.

Protocols and Standards

- RFC3315 Dynamic Host Configuration Protocol For IPv6
- RFC5007 DHCPv6 Leasequery
- RFC5460 DHCPv6 Bulk Leasequery

15.2 Applications

Application	Description
Prevention of DHCPv6 Spoofing	There is more than one DHCPv6 server on the network, and DHCPv6 clients can obtain network configuration parameters only from legal DHCPv6 servers.
Prevention of Forged DHCPv6 Packet Attacks	Malicious users on the network frequently send DHCPv6 request packets.
Prevention of Forged DHCPv6 Packet Attacks	Malicious users on the network send forged DHCPv6 request packets such as DHCPv6 release packets.
Prevention of IPv6/MAC Spoofing	Malicious users on the network send forged IPv6 request packets that temper the source address fields.
Prevention of Unauthorized IPv6 Configuration	Users do not obtain IPv6 addresses from the DHCPv6 server as required and configure IPv6 addresses without authorization.

15.2.1 Prevention of DHCPv6 Spoofing

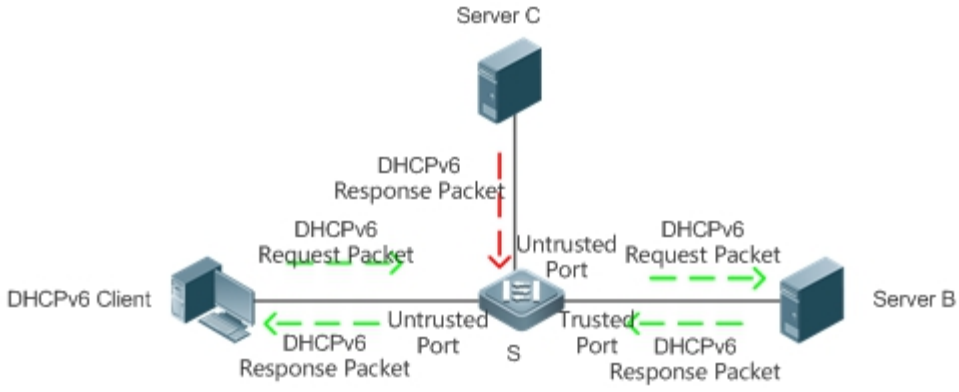
Scenario

There may exist more than one DHCPv6 server on the network, and it is necessary to ensure that user PCs obtain network configuration parameters only from the controlled DHCPv6 servers.

As shown in the following figure, the DHCPv6 client only communicates with trusted DHCPv6 servers.

- The request packets from the DHCPv6 client are transmitted only to a trusted DHCPv6 server.
- Only the response packets from the trusted DHCPv6 server can be transmitted to the client.

Figure 15- 1



Remarks	S is an access device. A is a user PC. B is a controlled DHCPv6 server. C is an uncontrolled DHCPv6 server.

Deployment

- Enable DHCPv6 snooping on the access device S for DHCPv6 packet monitoring.
- Set the port connecting the access device S to the DHCPv6 server B as a DHCPv6 trusted port to forward response packets.
- Set the other ports of the access device S as DHCPv6 untrusted ports to filter response packets.

15.2.2 Prevention of Forged DHCPv6 Packet Attacks

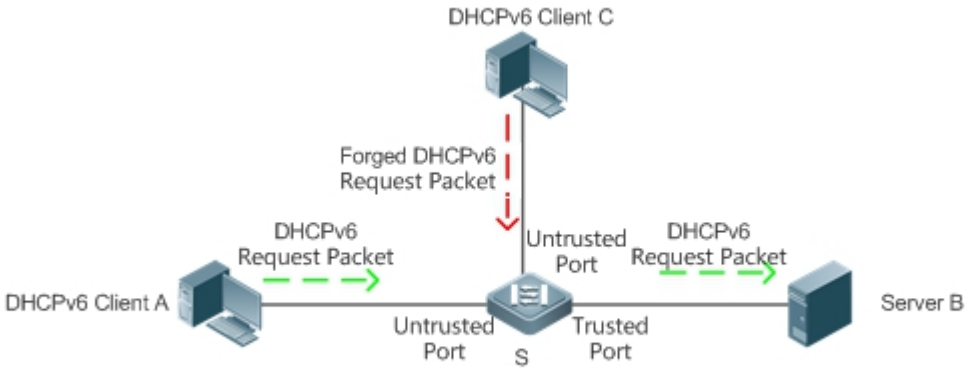
Scenario

There may exist malicious users on the network who forge DHCPv6 request packets. The packets not only consume available IPv6 addresses of the server but may also snatch IPv6 addresses from legal users. Therefore, such packets on the network must be filtered.

As shown in the following figure, the DHCPv6 request packets sent by the DHCPv6 client will be checked.

- Release packets and decline packets from the client must match those recorded in the internal snooping database.

Figure 15- 2



Remarks	S is an access device. A and C are user PCs. B is a controlled DHCPv6 server.

Deployment

- Enable DHCPv6 snooping on the access device S for DHCPv6 monitoring.
- Set the port connecting the access device S to the DHCPv6 server as a DHCPv6 trusted port to forward response packets.
- Set the other ports of the access device S as DHCPv6 untrusted ports to filter DHCPv6 packets.

15.2.3 Prevention of IPv6/MAC Spoofing

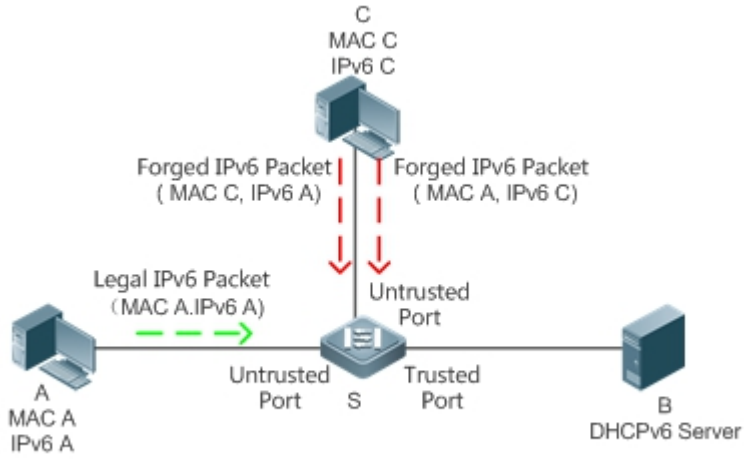
Scenario

When checking IPv6 packets from the untrusted port, you may check IP address fields only or IP+MAC fields to filter forged IPv6 packets.

As shown in the following figure, IPv6 packets sent from the DHCPv6 client will be checked.

- The source address fields of IPv6 packets must match IPv6 addresses assigned by the DHCPv6 client.
- The source Media Access Control (MAC) addresses of Layer-2 packets must match the client MAC addresses in DHCPv6 request packets of the client.

Figure 15- 3



Remarks	S is an access device. A and C are user PCs. B is a controlled DHCPv6 server.

Deployment

- Enable DHCPv6 snooping on the access device S for DHCPv6 monitoring.
- Set all downstream ports on the access device S as DHCPv6 untrusted ports.
- Enable IPv6 Source Guard on the access device S to filter IPv6 packets.
- On the access device S, set the match mode of IPv6 Source Guard as IPv6+MAC to check both MAC fields and IPv6 fields of IPv6 packets.

15.2.4 Prevention of Unauthorized IPv6 Configuration

Scenario

When checking IPv6 packets from untrusted ports, you need to check whether source IPv6 addresses of the packets are consistent with the IPv6 addresses assigned by the DHCPv6.

If the source IPv6 addresses, connection ports, or Layer-2 MAC addresses of IPv6 packets fail to match the assignment records of the DHCPv6 server snooped by the device, the packets should be discarded.

The operating process of the device in the scenario is the same as that in the preceding figure.

Deployment

- See section 15.2.3 "Prevention of IPv6/MAC Spoofing".

15.3 Features

Basic Concepts

DHCPv6 Request Packet

A DHCPv6 request packet is the packet sent from the DHCPv6 client to the DHCPv6 server. It includes DHCPv6 solicit packet, DHCPv6 request packet, DHCPv6 confirm packet, DHCPv6 rebind packet, DHCPv6 release packet, DHCPv6 decline packet, DHCPv6 renew packet, DHCPv6 inform-req packet, and DHCPv6 leasequery packet.

DHCPv6 Response Packet

A DHCPv6 response packet is the packet sent from the DHCPv6 server to the DHCPv6 client. It includes DHCPv6 advertise packet, DHCPv6 reply packet, DHCPv6 reconfigure packet, DHCPv6 relay-reply packet, DHCPv6 leasequery-reply packet, DHCPv6 leasequery-done packet, and DHCPv6 leasequery-data packet.

DHCPv6 Snooping Trusted Port

As the interactive packets used by DHCPv6 to obtain IPv6 addresses or prefixes are multicast packets, there may exist illegal DHCPv6 services affecting IPv6 acquisition, and user information may even be stolen by such illegal services. To prevent such issues, DHCPv6 snooping classifies ports into trusted and untrusted ports, and the devices forwards only the DHCPv6 response packets received by the trusted port and discards all DHCPv6 response packets from the untrusted port. By setting the ports connected to a legal DHCPv6 server as trusted ports and the others as untrusted ports, illegal DHCPv6 servers will be shielded.

On a switch, all switch ports or Layer-2 aggregate ports (APs) are untrusted ports by default, which can be configured as trusted ports.

Filtering DHCPv6 Snooping Request Packets

When DHCPv6 packets are disabled for an individual user, any DHCPv6 packets sent from the user's device shall be shielded. DHCPv6 request packet filtering can be configured on an untrusted port to filter all DHCPv6 request packets received by the port.

VLAN-based DHCPv6 Snooping

DHCPv6 snooping takes effect in the unit of VLAN. If DHCPv6 snooping is enabled by default, the function is enabled on all VLANs of the device. The VLAN on which DHCPv6 snooping takes effect can be flexibly controlled through configuration.

DHCPv6 Snooping User Database

On a DHCPv6 network, a frequently encountered problem is that users may arbitrarily set static IPv6 addresses. Such addresses are difficult to maintain and may conflict with legal user addresses, making the users unable to access the Internet. By snooping the packets exchanged between the client and the server, DHCPv6 snooping forms IPv6 information obtained by users, user MAC, VID, PORT, and lease time into a user record, thus making a DHCPv6 snooping user database to control legal use of IPv6 addresses.

DHCPv6 Option 18 and Option 37

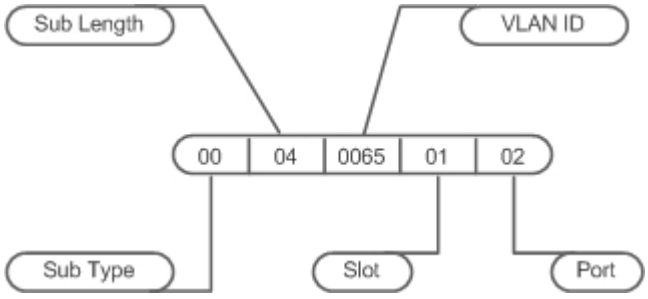
When managing user IP addresses, some network administrators expect to determine the IP addresses to be assigned according to the user locations; that is, they expect to assign IP addresses to users according to the information on the connected network devices, thereby adding user-related device information to DHCP request packets through DHCPv6 option while performing DHCPv6 snooping. The option number for RFC3315 is 18; the option number for RFC4649, the option number used is 37. After the content of Option 18 and Option 37 is parsed on the DHCPv6 server, the server can obtain information of more users according to the content uploaded by Option 18 and option 37 so as to assign IP addresses more accurately.

- Option 18: Interface ID

The default content of Interface ID include the number of the VLAN to which the port receiving request packets from the DHCPv6 client belongs, and the port index (the values of the port index are the slot number and port number); the extension content is a customized character string. Default and extension fillings take effect only for wired interfaces, including switch ports, Layer-2 APs, or Layer-2 encapsulation sub-interfaces.

The Interface ID filling format can be classified into standard and extension formats, only one of which can be used on the same network. When the standard filling format is used, only default content can be filled in for sub-options of Interface ID, as shown in the following figure:

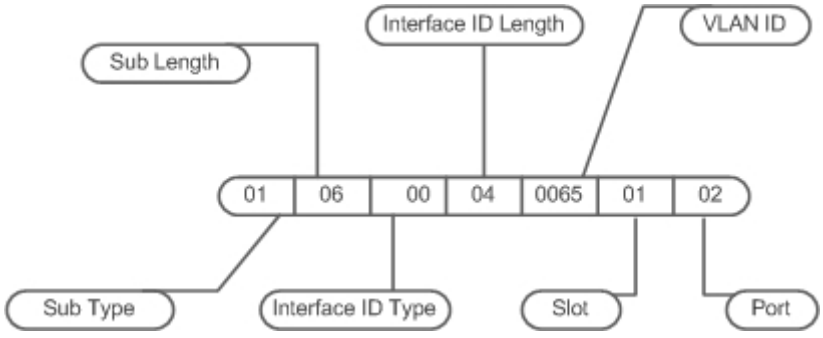
Figure 15- 4



To use customized content, the extension filling format can be used. The content filled in by extension can be default or extension content. To distinguish between the content, add a content type field and a content length field of one byte respectively following the sub-option length. For default content, set the content type as 0; for extension content, set the content type as 1.

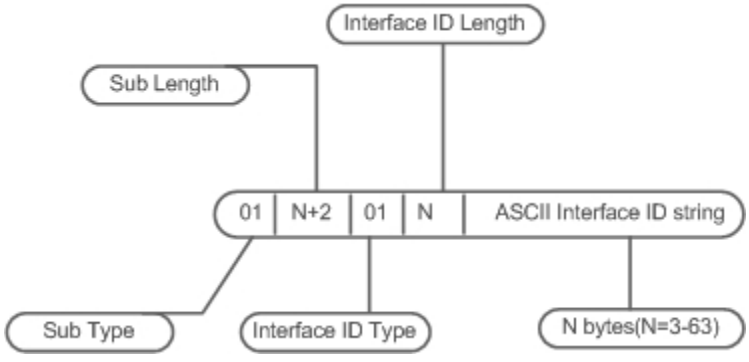
The format of default content is as follows:

Figure 15- 5



The format of extension content is as follows:

Figure 15- 6

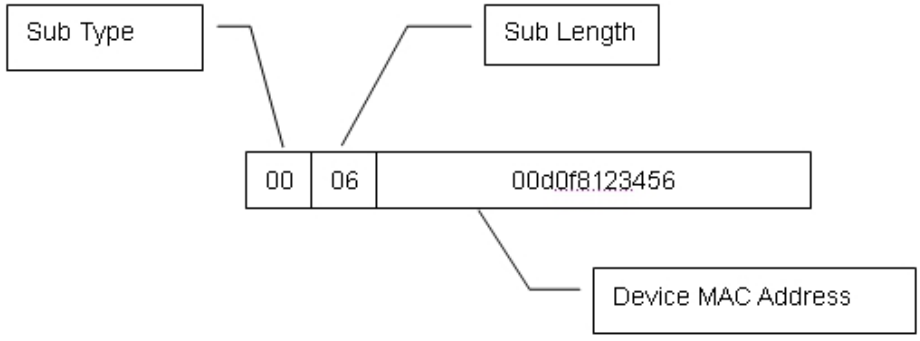


- Option 37: Remote ID

The default content of Remote ID is the bridge MAC address of the DHCPv6 relay that receives request packets from the DHCPv6 client, and the extension content is a customized character string.

The Remote ID filling format can be classified into standard and extension formats, only one of which can be used on the same network. When the standard filling format is used, only default content are filled in for sub-options of Remote ID, as shown in the following figure:

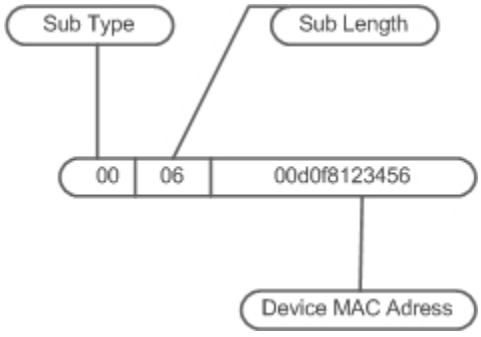
Figure 15- 7



To use customized content, the extension filling format can be used. The content filled in by extension can be default or extension content. To distinguish between the content, add a content type field and a content length field of one byte respectively following the sub-option length. For default content, set the content type as 0; for extension content, set the content type as 1.

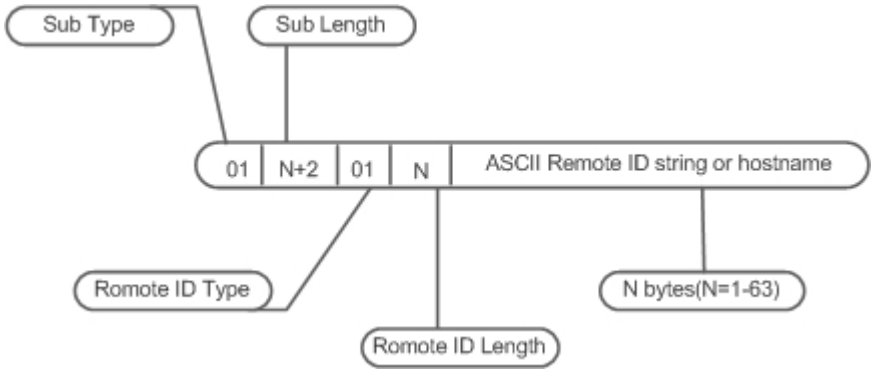
The format of default content is as follows:

Figure 15- 8



The format of extension content is as follows:

Figure 15- 9



- Note

Option 18: The values of port index for Interface ID are the slot number and port number. The port can be a wired switch port, Layer-2 AP,

or Layer-2 encapsulation sub-interface. The port number refers to the sequence number of the port in the slot. The port number of a Layer-2 AP is an AP number. For example, the port number of Fa0/10 is 10, and the port number of AP 11 is 11;

Slot numbers are the sequence numbers of all slots on a device (one device in stack mode). The slot number of an AP is the last one. The sequence numbers of slots start from 0. Run the **show slots** command to display the numbers. For example:

Example 1:

```
FS#show slots (only Dev and slot displayed)
  Dev Slot
  --- ----
  1    0      ------> The slot number is 0.
  1    1      ------> The slot number is 1.
  1    2      ------> The slot number is 2.
```

In this case, the slot number of an AP is 3.

Example 2:

```
FS#show slots (only Dev and slot displayed)
  Dev Slot
  --- ----
  1    0      ------> The slot number is 0.
  1    1      ------> The slot number is 1.
  1    2      ------> The slot number is 2.
  2    0      ------> The slot number is 3.
  2    1      ------> The slot number is 4.
  2    2      ------> The slot number is 5.
```

In this case, the slot number of an AP is 6.

Illegal DHCPv6 Packet

DHCPv6 snooping checks the validity of DHCPv6 packets passing through the device, discards illegal DHCPv6 packets, records user information, and generates a DHCPv6 snooping binding database for query of other functions. The following packets are considered as illegal DHCPv6 packets.

- DHCPv6 response packets received by untrusted ports. For details, see the section DHCPv6 Response Packet.
- Relayed DHCPv6 packets received by untrusted ports, namely DHCPv6 relay-forw packets and DHCPv6 relay-reply packets.
- DHCPv6 relay-reply packets received by trusted ports. The egress for these packets is an untrusted ports according to the entry.
- DHCPv6 release packets; no corresponding users are found in the DHCPv6 snooping user database according to the Layer-2 source MAC and VID of these packets.
- DHCPv6 release packets. The IPv6 addresses or prefixes of these packets do not exist in the DHCPv6 snooping user database.
- DHCPv6 release packets. The IPv6 addresses or prefixes of these packets all exist in the DHCPv6 snooping user database but the untrusted ports of DHCPv6 release packets are inconsistent with those untrusted ports in the DHCPv6 snooping user database.

- DHCPv6 packets in incorrect formats or incomplete packets.

Overview

Features	Description
Filtering Illegal DHCPv6 Packets	Checks the validity of exchanged DHCPv6 packets, and discards illegal packets (see the preceding section for instructions for illegal packets). Forwards only legal response packets to trusted ports.
Establishing a User Database	Snoops interaction between the client and the server, and generates the DHCPv6 snooping user database to provide a basis for other security filtering modules.

15.3.1 Filtering Illegal DHCPv6 Packets

This function is to check the validity of DHCPv6 packets from untrusted ports, filter the packets according to the types of illegal packets described in Basic Concepts above, and control the transmission scope of packets to prevent malicious users from spoofing.

Working Principle

During snooping, the receipt ports of packets and packet fields are checked to filter the packets; the destination ports of packets are modified to control the transmission scope of packets.

Checking Ports

When receiving DHCPv6 packets, the device first determines whether the port receiving packets is a DHCPv6 trusted port. If the port is a trusted port, the packets will be forwarded without validity check, binding, or prefix record generation. If the port is an untrusted port, validity check is required.

Checking whether Packet Encapsulation and Length are Complete

Check whether the packets are User Datagram Protocol (UDP) packets and the destination port is 546 or 547. Check whether the actual length of a packet matches the length field described in the protocol.

Checking Whether DHCPv6 Packet Field and Packet Type are Correct

Check whether the packets are relayed according to the types of illegal packets described in the preceding section Basic Concepts, and then check whether the restrictions specific to a type of packets are met according to the actual type of packets.

Related Configuration

Enabling Global DHCPv6 Snooping

By default, DHCPv6 snooping is disabled.

Run the [**no**] **ipv6 dhcp snooping** command to enable or disable DHCPv6 snooping.

To enable or disable DHCPv6 snooping on different VLANs, global DHCPv6 snooping must be enabled first.

Setting DHCPv6 Snooping on a VLAN

By default, when global DHCPv6 snooping is enabled, DHCPv6 snooping takes effect on all VLANs.

Run the [**no**] **ipv6 dhcp snooping vlan** command to enable or disable DHCPv6 snooping on a VLAN. The range of command parameter values is the actual range of VLAN numbers.

15.3.2 Establishing a User Database

The packets exchanged between the DHCPv6 client and the DHCPv6 server are snooped, and DHCPv6 snooping binding entries and prefix entries are generated according to the information on legal DHCPv6 packets. All the entries are provided for other security configuration modules as an information list of legal users and a basis for network packet filtering.

Working Principle

During snooping, binding database and prefix database are continuously updated according to the types of DHCPv6 packets.

Generating Binding or Prefix Records

When DHCPv6 reply packets are snooped on a trusted port, client IPv6 addresses or prefixes, client MAC addresses, and lease time fields of the packets are extracted, and a binding or prefix record is generated according to the client port ID recorded by the device (wired interface index), and the client VLAN.

Deleting Binding or Prefix Records

When the recorded lease time is over, or the legal DHCPv6 release/DHCPv6 decline packets sent from the client are snooped, or users run the clear command to delete binding or prefix records, the corresponding binding or prefix records are deleted.

Related Configuration

Enable DHCPv6 snooping without extra configuration.

15.4 Configuration

Configuration	Description and Command	
Configuring Basic DHCPv6 Snooping Functions	(Mandatory) It is used to establish DHCPv6 snooping.	
	ipv6 dhcp snooping	Enables DHCPv6 snooping.
	ipv6 dhcp snooping binding-delay	Delays assignment of the DHCPv6 snooping binding entries to the hardware filtering entries.
	ipv6 dhcp snooping filter-dhcp-pkt	Enables DHCPv6 request packet filtering.
	ipv6 dhcp snooping vlan	Enables and disables DHCPv6 snooping for specified VLANs.
	ipv6 dhcp snooping database write-delay	Enables the function for regularly saving DHCPv6 snooping binding and prefix records.
	ipv6 dhcp snooping database write-to-flash	Manually saves DHCPv6 snooping binding and prefix records.
	renew ipv6 dhcp snooping database	Manually imports the user records saved in flash to the DHCPv6 snooping user database.
	ipv6 dhcp snooping trust	Configures DHCPv6 snooping trusted ports.
	ipv6 dhcp snooping link-detection	Clears dynamical biding entries on a port when the port is configured into Link Down state.
Configuring Option 18 and	(Optional) It is used to optimize assignment of DHCPv6 server addresses.	

Configuration	Description and Command	
Option 37	**ipv6 dhcp snooping Information option** [**standard-format**]	Adds Option 18 or Option 37 to DHCPv6 request packets. **standard-format**: Fills in content in a standard format if such keyword exists; otherwise, fills in content in an extension format.
	ipv6 dhcp snooping information option format remote-id [**string** *ASCII-string* | **hostname**]	Configures Remote ID in an extension format. **string**: Indicates that the content filled in is a customized character string. **hostname**: Indicates that the content filled in is hostname.
	ipv6 dhcp snooping vlan *vlan-id* **information option format-type** *interface-id* **string** *ASCII-string*	Configures the customized character string of Interface ID in an extension format.
	ipv6 dhcp snooping vlan *vlan-id* **information option change-vlan-to vlan** *vlan-id*	Configures VLAN mapping for Interface ID in an extension format, which is exclusive from the [**no**] **ipv6 dhcp snooping vlan** *vlan-id* **information option format-type** *interface-id* **string** *ASCII-string* command.

15.4.1 Configuring Basic DHCPv6 Snooping Functions

Configuration Effect

- Enable DHCPv6 snooping.
- Generate DHCPv6 snooping binding and prefix databases.
- Control the transmission scope of DHCPv6 packets.
- Filter illegal DHCPv6 packets.

Notes

- The port connecting the device to a trusted DHCPv6 server must be set as a trusted port.
- The port on which DHCPv6 snooping takes effect can be a wired switch port, Layer-2 AP or Layer-2 encapsulation sub-interface.
- The Link Down entry clearing function applies only to wired ports.

Configuration Steps

Enabling Global DHCPv6 Snooping

- Mandatory.
- If not specified, configure this function on an access device.

Delaying Assignment of DHCPv6 Snooping Binding Entries to Hardware Filtering Entries

- Configure the function if assignment needs to be delayed. Assignment is not delayed by default.
- If not specified, configure this function on an access device.

Enabling DHCPv6 Request Packet Filtering

- Enable the function if users' DHCPv6 requests need to be restricted on a port.
- If not specified, disable the function on the access device.

Enabling and Disabling VLAN-based DHCPv6 Snooping

- Disable DHCPv6 snooping if the function is not needed on a VLAN.
- If not specified, configure this function on an access device.

Enabling Regular Saving of DHCPv6 Snooping Binding Records

- This function should be enabled if DHCPv6 snooping binding records need to be maintained after the device is restarted.
- If not specified, enable the function on the access device.

Configuring DHCPv6 Trusted Ports

- Mandatory.
- Set the port connecting the device to a trusted DHCPv6 device as a DHCPv6 trusted port.

Enabling and Disabling Clearing of Dynamically Bound Entries When the Port is Configured into Link Down State

- On a stable network, enable the function to release spaces occupied by hardware entries and timely clear the entries on the Link Down port.
- If not specified, disable the function on the access device.

Verification

Enable the device to use DHCPv6 to obtain network configuration parameters.

- Check whether user records are generated in the DHCPv6 snooping binding database.

Related Commands

Enabling and Disabling DHCPv6 Snooping

Command	[**no**] **ipv6 dhcp snooping**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	After global DHCPv6 snooping is enabled, run the **show ipv6 dhcp snooping** command to check whether DHCPv6 snooping is enabled.

Delaying Assignment of the DHCPv6 Snooping Binding Entries to the Hardware Filtering Entries

Command	[**no**] **ipv6 dhcp snooping binding-delay**
Parameter Description	***seconds***: Indicates the time for delaying assignment of binding entries to hardware filtering entries, in the unit of seconds. The value is 0 by default.
Command Mode	Global configuration mode
Usage Guide	By default, dynamically bound entries are added to hardware filtering entries in real time. After the function is configured,

	the dynamically generated binding entries are bound to hardware filtering entries only when no IPv6 address conflicts are detected within a specified time period.

Configuring a VLAN on Which DHCPv6 Snooping Takes Effect

Command	[**no**] **ipv6 dhcp snooping vlan** { *vlan-rng* | {*vlan-min* [*vlan-max*] } }
Parameter Description	*vlan-rng*: Indicates the VLAN scope in which DHCPv6 snooping takes effect. *vlan-min*: Indicates the lower VLAN limit where DHCPv6 snooping takes effect. *vlan-max*: Indicates the upper VLAN limit where DHCPv6 snooping takes effect.
Command Mode	Global configuration mode
Usage Guide	DHCPv6 snooping is enabled or disabled on a specified VLAN by configuring the command. This function takes effect only if global DHCPv6 snooping is enabled.

Filtering DHCPv6 Request Packets on a Port

Command	[**no**] **ipv6 dhcp snooping filter-dhcp-pkt**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	All DHCPv6 request packets can be prohibited on the port by configuring the command; that is, all users are prohibited from applying for addresses on the port.

Regularly Writing DHCPv6 Snooping Database Information into Flash

Command	[**no**] **ipv6 dhcp snooping database write-delay** [*time*]
Parameter Description	*time*: Indicates the interval for regularly writing the DHCPv6 snooping database into flash.
Command Mode	Global configuration mode
Usage Guide	The DHCPv6 snooping database can be written into a flash file by configuring the command. The function prevents user information loss after the device restarts. If user information is lost, users have to re-obtain IP addresses for normal communication.

Manually Writing DHCPv6 Snooping Database Information into Flash

Command	**ipv6 dhcp snooping database write-to-flash**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Dynamic user information in the DHCPv6 snooping database can be written into a flash file in real time by running the command.

Manually Importing Information in Flash to the DHCPv6 Snooping Binding Database

Command	**renew ipv6 dhcp snooping database**

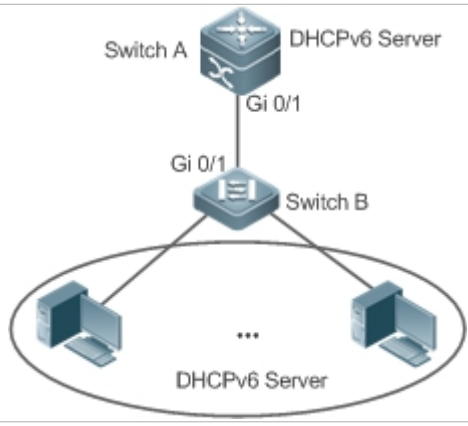
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	Flash file information can be written into the DHCPv6 snooping database in real time by running the command.

Configuring a Port as a Trusted Port

Command	[**no**] **ipv6 dhcp snooping trust**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	The port connecting to a legal DHCPv6 server is configured as a trusted port by configuring the command. The DHCPv6 response packets received by a trusted port are forwarded, while the DHCPv6 response packets received by an untrusted port are discarded.

Configuration Example

Dynamically obtaining IPv6 addresses through the legal DHCPv6 server on a DHCPv6 client

Scenario **Figure 15- 10**	
Configuration Steps	● Enable DHCPv6 snooping on the access device (Switch B). ● Set the uplink port (Gi 0/1) as a trusted port.
B	B#configure terminal Enter configuration commands, one per line. End with CNTL/Z. B(config)#ipv6 dhcp snooping B(config)#interface gigabitEthernet 0/1 B(config-if-GigabitEthernet 0/1)#ipv6 dhcp snooping trust B(config-if-GigabitEthernet 0/1)#end
Verification	Confirm configuration of Switch B. ● Confirm whether DHCPv6 snooping is enabled and whether the DHCPv6 snooping trusted port configured is the uplink port.

	● On Switch B, check the configuration of DHCP snooping, especial whether the trusted port is correct.
B	FS#show ipv6 dhcp snooping DHCPv6 snooping status : ENABLE DHCPv6 snooping database write-delay time : 0 seconds DHCPv6 snooping binding-delay time : 0 seconds DHCPv6 snooping option18/37 status : DISABLE DHCPv6 snooping link detection : DISABLE Interface Trusted Filter DHCPv6 ---------- ------- ---------- GigabitEthernet 0/1 YES DISABLE FS#show ipv6 dhcp snooping binding Total number of bindings: 1 NO. MacAddress IPv6 Address Lease(sec) VLAN Interface ----- ---------- ---------- ----- ---------- 1 00d0.f801.0101 2001::10 42368 2 GigabitEthernet 0/1

Common Errors

- The uplink port is not set as a DHCPv6 trusted port.
- Other access security options are configured on the uplink port, resulting in failure of DHCPv6 trusted port configuration.

15.4.2 Configuring Option 18 and Option 37

Configuration Effect

- The DHCPv6 server can obtain more information during address assignment, thus improving address assignment.
- The option is transparent to the DHCPv6 client, and such function is perception-free to the client.

Configuration Steps

- Run the configuration if the optimization is needed.
- If not specified, enable the function on the device where DHCPv6 snooping is enabled.

Verification

Check the configuration of DHCPv6 snooping to ensure that such function is enabled.

Related Commands

Adding Option18 and Option 37 to DHCPv6 Request Packets

Command	[**no**] **ipv6 dhcp snooping information option** [**standard-format**]

Parameter Description	**standard-format**: Fills in content in a standard format if such keyword exists; otherwise, fills in content in an extension format.
Command Mode	Global configuration mode
Usage Guide	Information on Option 18 and Option 37 is added to DHCPv6 request packets by configuring the command, and the DHCPv6 server assigns addresses according to information on Option 18 and Option 37.

Setting Option 37 (Remote ID) as a Customized Character String

Command	[**no**] **ipv6 dhcp snooping information option format remote-id** { **string** *ASCII-string* | **hostname** }
Parameter Description	**string** *ASCII-string*: Indicates that the content of Remote ID in an extension format is a customized character string. **hostname**: Indicates that the content of Remote ID in an extension format is hostname.
Command Mode	Global configuration mode
Usage Guide	Remote ID is configured in an extension format by configuring the command. Remote ID is customized, and the DHCPv6 server assigns addresses according to information on Option 37.

Setting Option 18 (Interface ID) as a Customized Character String

Command	[**no**] **ipv6 dhcp snooping vlan** *vlan-id* **information option format-type** *interface-id* **string** *ASCII-string*
Parameter Description	*vlan-id*: Indicates the VLAN to which DHCPv6 request packets belong. *ASCII-string*: Indicates the user-customized content to be filled in for Interface-ID.
Command Mode	Interface configuration mode
Usage Guide	Customized character strings of Interface ID are configured in an extension format by configuring the command, and the DHCPv6 server assigns addresses according to information on Option 18.

Setting Option 18 (Interface ID) as a Modified VLAN

Command	[**no**] **ipv6 dhcp snooping vlan** *vlan-id* **information option change-vlan-to vlan** *vlan-id*
Parameter Description	*vlan-id* (the first one): Indicates the VLAN to which DHCPv6 request packets belong. *vlan-id* (the second one): Indicates the VLAN after modification.
Command Mode	Interface configuration mode
Usage Guide	Interface ID is configured as VLAN mapping in an extension format by configuring the command, and the DHCPv6 server assigns addresses according to information on Option 18.

Configuration Example

The following example shows how to add Option 18 and Option 37 to DHCPv6 request packets.

Configuration Steps	• Configure basic DHCPv6 snooping functions.(Omitted) • Enable the function for adding Option 18 and Option 37.
B	FS# configure terminal FS(config)# ipv6 dhcp snooping information option FS(config)# end

Verification	Display the DHCPv6 snooping configuration.
B	FS #show ipv6 dhcp snooping DHCPv6 snooping status : ENABLE DHCPv6 snooping database write-delay time : 0 seconds DHCPv6 snooping binding-delay time : 0 seconds DHCPv6 snooping option 18/37 status : ENABLE DHCPv6 snooping link detection : DISABLE Interface Trusted Filter DHCPv6 ---------- ------- ---------- FastEthernet0/10 YES DISABLE

15.5 Monitoring and Maintenance

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command
Clears dynamic user information in the DHCPv6 snooping database.	**clear ipv6 dhcp snooping binding** [**vlan** *vlan-id* | *mac* | *ipv6* | **interface** *interface-id*]
Clears all entries in the DHCPv6 snooping prefix database.	**clear ipv6 dhcp snooping prefix**
Clears statistics about DHCPv6 snooping handling DHCPv6 packets.	**clear ipv6 dhcp snooping statistics**

Displaying

Description	Command
Displays DHCPv6 snooping configuration.	**show ipv6 dhcp snooping**
Displays the VLANs on which DHCPv6 snooping fails to take effect.	**show ipv6 dhcp snooping vlan**
Displays all dynamically bound entries in the DHCPv6 snooping binding database.	**show ipv6 dhcp snooping binding**
Displays all entries in the DHCPv6 snooping prefix database.	**show ipv6 dhcp snooping prefix**
Displays the counters of DHCPv6 snooping handling packets.	**show ipv6 dhcp snooping statistics**
Displays all statically bound entries added manually and all dynamically bound entries in the DHCPv6 snooping binding database.	**show ipv6 source binding**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command

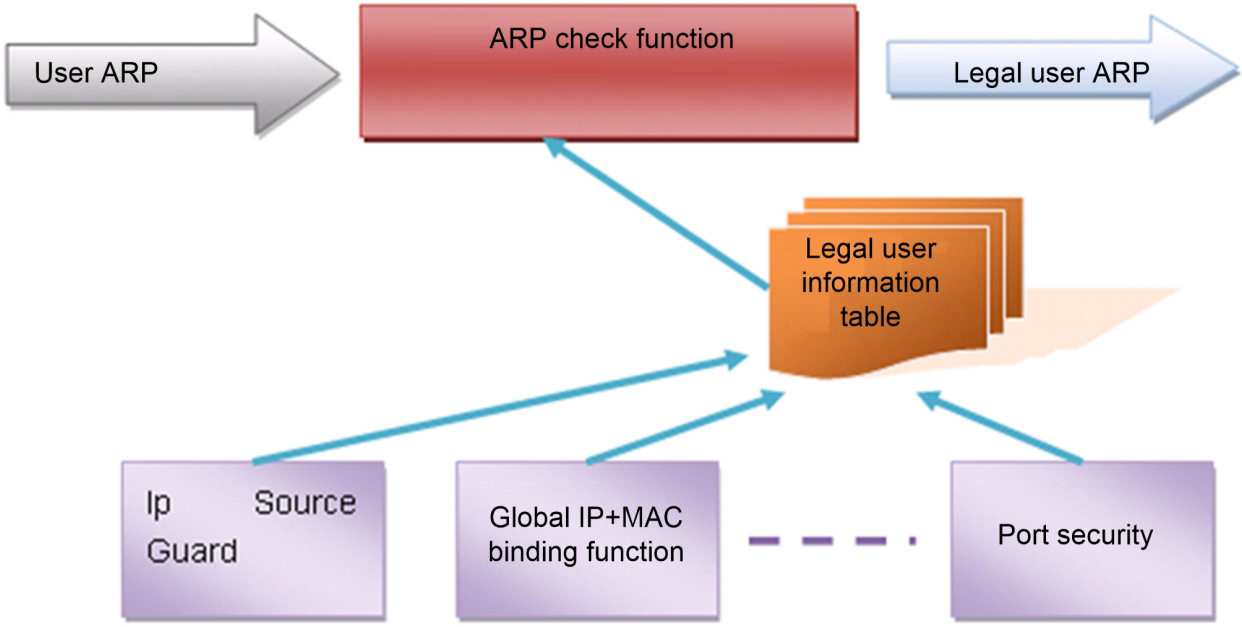
Debugs DHCPv6 snooping events.	**debug snooping ipv6 event**
Disables debugging of DHCPv6 snooping events.	**no debug snooping ipv6 event**
Debugs DHCPv6 snooping packets.	**debug snooping ipv6 packet**
Disables debugging of DHCPv6 snooping packets.	**no debug snooping ipv6 packet**

16 Configuring ARP Check

16.1 Overview

The Address Resolution Protocol (ARP) packet check filters all ARP packets under ports (including wired layer-2 switching ports, layer-2 aggregate ports (APs), and layer-2 encapsulation sub-interfaces) and discards illegal ARP packets, so as to effectively prevent ARP deception via networks and to promote network stability. On devices supporting ARP check, illegal ARP packets in networks will be ignored according to the legal user information (IP-based or IP-MAC based) generated by security application modules such as IP Source Guard, global IP+MAC binding, 802.1X authentication, GSN binding, Web authentication and port security.

Figure 16- 1



The above figure shows that security modules generate legal user information (IP-based or IP-MAC based). ARP Check uses the information to detect whether the Sender IP fields or the <Sender IP, Sender MAC>fields in all ARP packets at ports matches those in the list of legal user information. If not, all unlisted ARP packets will be discarded.

Protocols and Standards

- RFC826: An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses

16.2 Applications

Application	Description
Filtering ARP packets in Networks	Illegal users in networks launch attacks using forged ARP packets.

16.2.1 Filtering ARP Packets in Networks

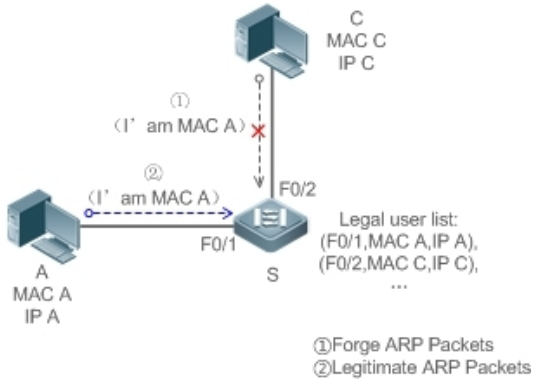
Scenario

Check ARP packets from distrusted ports and filter out ARP packets with addresses not matching the results assigned by the DHCP server.

For example, in the following figure, the ARP packets sent by DHCP clients are checked.

- The ports receiving ARP packets, the source MAC addresses of ARP packets, and the source IP addresses of ARP packets shall be consistent with the snooped DHCP-assigned records.

Figure 16- 2



Remarks:	S is an access device. A and C are user PCs.

Deployment

- Enable DHCP Snooping on S to realize DHCP monitoring.
- Set all the downlink ports on S as DHCP distrusted ports.
- Enable IP Source Guard and ARP Check on all distrusted ports on S to realize ARP packet filtration.

16.3 Features

Basic Concepts

Compatible Security Modules

Presently, the ARP Check supports the following security modules.

- IP-based: IP-based mode: port security, and static configuration of IP Source Guard.
- IP-MAC based: IP-MAC based mode: port security, global IP+MAC binding, 802.1X authorization, IP Source Guard, GSN binding, and Web authentication.

Two Modes of APR Check

The ARP Check has two modes: Enabled and Disabled. The default is Enabled.

1. Enabled Mode

Through ARP Check, ARP packets are detected based on the IP/IP-MAC based binding information provided by the following modules.

- Global IP-MAC binding
- 802.1X authorization
- IP Source Guard
- GSN binding

- Port security
- Web authentication
- Port security IP+MAC binding or IP binding

When only ARP Check is enabled on a port but the above-mentioned modules are not enabled, legal user information cannot be generated, and thereby all ARP packets from this port will be discarded.

When the ARP Check and VRRP functions are enabled on an interface, if the physical IP address and virtual IP address of the interface can be used as the gateway address, the physical IP address and VRRP IP address need to be permitted to pass. Otherwise, ARP packets sent to the gateway will be filtered out.

2. Disabled Mode

ARP packets on a port are not checked.

Overview

Feature	Description
Filtering ARP Packets	Check the source IP and source MAC addresses of ARP packets to filter out illegal ARP packets.

16.3.1 Filtering ARP Packets

Enable ARP Check on specified ports to realize filtration of illegal ARP packets.

Working Principle

A device matches the source IP and source MAC addresses of the ARP packets received at its ports with the legal user information of the device. With successful matching, packets will be transferred, or otherwise they will be discarded.

Related Configuration

Enabling ARP Check on Ports

By default, the ARP Check is disabled on ports.

Use the **arp-check** command to enable ARP Check.

Unless otherwise noted, this function is usually configured on the ports of access devices.

16.4 Configuration

Configuration	Description and Command	
Configuring ARP Check	(Mandatory) It is used to enable APR Check.	
	arp-check	Enables ARP Check.

16.4.1 Configuring ARP Check

Configuration Effect

- Illegal ARP packets are filtered out.

Notes

- When ARP Check is enabled, the number of policies or users of related security applications may decrease.
- ARP Check cannot be configured on mirrored destination ports.
- ARP Check cannot be configured on the trusted ports of DHCP Snooping.
- ARP Check cannot be configured on global IP+MAC exclude ports.
- ARP Check can be enabled only on wired switching ports, layer-2 APs, layer-2 encapsulation sub-interfaces. Enable ARP check for the wired in interface configuration mode

Configuration Steps

Enabling ARP Check

- (Mandatory) The function is disabled by default. To use the ARP Check function, an administrator needs to run a command to enable it.

Verification

- Use the **show run** command to display the system configuration.
- Use the **show interface** { *interface-type interface-number* } **arp-check list** command to display filtering entries.

Related Commands

Enabling ARP Check

Command	**arp-check**
Parameter Description	N/A
Command	Interface configuration mode
Usage Guide	Generate ARP filtration information according to the legal user information of security application modules to filter out illegal ARP packets in networks.

Configuration Example

The following configuration example introduces only ARP Check related configurations.

Enabling ARP Check on ports

Configuration Steps	• Enable ARP Check. Restricted ARP packets must conform to entries of IP Source Guard, port security, or global IP+MAC binding.

```
FS# configure terminal
FS(config)#address-bind 192.168.1.3 00D0.F800.0003
FS(config)#address-bind install
FS(config)#ip source binding 00D0.F800.0002 vlan 1 192.168.1.4 interface gigabitEthernet 0/1
FS(config)# interface GigabitEthernet 0/1
FS(config-if-GigabitEthernet 0/1)#arp-check
FS(config-if-GigabitEthernet 0/1)#ip verify source port-security
FS(config-if-GigabitEthernet 0/1)#switchport port-security
FS(config-if-GigabitEthernet 0/1)#switchport port-security binding 00D0.F800.0001 vlan 1 192.168.1.1
FS(config-if-GigabitEthernet 0/1)#exit
FS(config)#interface gigabitEthernet 0/4
FS(config-if-GigabitEthernet 0/4)#switchport port-security
FS(config-if-GigabitEthernet 0/4)#switchport port-security binding 192.168.1.5
FS(config-if-GigabitEthernet 0/4)#arp-check
FS(config-if-GigabitEthernet 0/4)#exit
FS(config)#interface gigabitEthernet 0/5
FS(config-if-GigabitEthernet 0/5)#arp-check
FS(config-if-GigabitEthernet 0/5)#end
```

Verification

Use the **show interfaces arp-check list** command to display the effective ARP Check list for interfaces.

```
FS# show interface arp-check list

INTERFACE                    SENDER MAC          SENDER IP           POLICY SOURCE
------------------------ -------------------- -------------------- --------------------
GigabitEthernet 0/1          00d0.f800.0003      192.168.1.3         address-bind
GigabitEthernet 0/1          00d0.f800.0001      192.168.1.1         port-security
GigabitEthernet 0/1          00d0.f800.0002      192.168.1.4         DHCP snooping
GigabitEthernet 0/4          00d0.f800.0003      192.168.1.3         address-bind
GigabitEthernet 0/4                              192.168.1.5         port-security
GigabitEthernet 0/5          00d0.f800.0003      192.168.1.3         address-bind
```

Common Errors

- If ARP packets at a port need to be checked but APR-Check is disabled, then APR-Check will not be effective.

16.5 Monitoring

Displaying

Description	Command
Displays the effective ARP Check list based on ports.	**show interface** [*interface-type interface-number*] **arp-checklist**

17 Configuring Dynamic ARP Inspection

17.1 Overview

Dynamic Address Resolution Protocol (ARP) inspection (DAI) checks the validity of received ARP packets. Invalid ARP packets will be discarded.

DAI ensures that only valid ARP packets can be forwarded by devices. DAI mainly performs the following steps:

- Intercepts all ARP request packets and ARP reply packets on untrusted ports in the virtual local area networks (VLANs) where the DAI function is enabled.
- Checks the validity of intercepted ARP packets according to user records stored in a security database.
- Discards the ARP packets that do not pass the validity check.
- Sends the ARP packets that pass the validity check to the destination.
- The DAI validity criteria are the same as those of ARP Check. For details, see the *Configuring ARP Check*.

DAI and ARP Check have same functions. The only difference is that DAI takes effect by VLAN whereas ARP Check takes effect by port.

Protocols and Standards

- RFC826: An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses

17.2 Applications

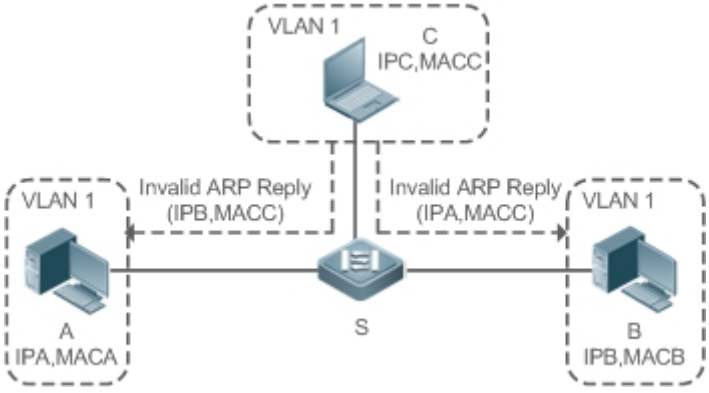
Application	Description
ARP Spoofing Prevention	Prevent ARP spoofing that is mounted by taking advantage of ARP defects.

17.2.1 ARP Spoofing Prevention

Scenario

Due to inherent defects, ARP does not check the validity of received ARP packets. Attackers can take advantage of the defects to mount ARP spoofing. A typical example is man-in-the-middle (MITM) attack. See Figure 17- 1.

Figure 17- 1



Remarks	Device S is a FS access switch enabled with DAI. User A and User B are connected to Device S, and they are in the same subnet.

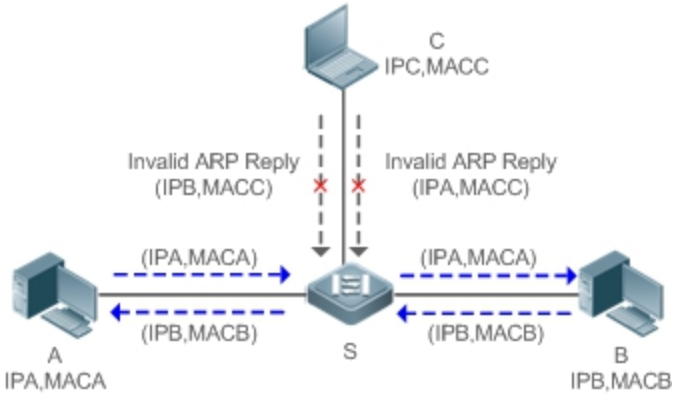
	User C is a malicious user connected to Device S. IP A and MAC A are the IP address and MAC address of User A. IP B and MAC B are the IP address and MAC address of User B. IP C and MAC C are the IP address and MAC address of User C.

When User A needs to initiate network layer communication with User B, User A broadcasts an ARP request in the subnet to query the MAC address of User B. Upon receiving the ARP request packet, User B updates its ARP cache with IP A and MAC A, and sends an ARP reply. Upon receiving the ARP reply packet, User A updates its ARP cache with IP B and MAC B.

In this model, User C can make the ARP entry mapping between User A and User B incorrect by continuously broadcasting ARP reply packets to the network. The reply packets contain IP A, IP B, and MAC C, After receiving these reply packets, User A stores the ARP entry (IP B, MAC C), and User B stores the ARP entry (IP A, MAC C). As a result, the communication between User A and User B is directed to User C, without the knowledge of User A and User B. Here User C acts as the man in the middle by modifying received packets and forwarding them to User A or User B.

If Device S is enabled with DAI, it will filter out forged ARP packets to prevent ARP spoofing as long as the IP addresses of User A and User B meet the validity criteria described in section 17.1 Overview. Figure 17- 2 shows the working process of DAI.

Figure 17- 2



Remarks	Device S is a FS access switch enabled with DAI. User A and User B are connected to Device S, and they are in the same subnet. User C is a malicious user connected to Device S. IP A and MAC A are the IP address and MAC address of User A. IP B and MAC B are the IP address and MAC address of User B. IP C and MAC C are the IP address and MAC address of User C.

The ARP packets of User A and User B are forwarded normally by Device S. The forged ARP packets of User C are discarded because the packets do not match the records in the security database of Device S.

Deployment

- Enable DHCP Snooping on Device S.
- Enable DAI and IP Source Guard on Device S.

17.3 Features

Basic Concepts

Trust Status of Ports and Network Security

ARP packet check is performed according to the trust status of ports. DAI considers packets received from trusted ports as valid without checking their validity, but it checks the validity of packets received from untrusted ports.

For a typical network configuration, you should configure Layer-2 ports connected to network devices as trusted ports, and configure Layer-2 ports connected to hosts as untrusted ports.

Network communication may be affected if a Layer-2 port connected to a network device is configured as an untrusted port.

Overview

Feature	Description
Invalid ARP Packet Filter	Checks the source IP addresses and MAC addresses of ARP packets to filter out invalid packets.
DAI Trusted Port	Permits the ARP packets received from specific ports to pass through without checking their validity.

17.3.1 Invalid ARP Packet Filter

Enable DAI in a specific VLAN to filter out invalid ARP packets. The DAI validity criteria are the same as those of ARP Check.

Working Principle

Upon receiving an ARP packet, the device matches the IP address and MAC address of the packet with the valid user records in its security database. If the packet matches a record, it will be forwarded normally. If it does not match any record, it will be discarded.

DAI and ARP Check use the same set of valid user records. For details, see the packet validity check description in the *Configuring ARP Check*.

Related Configuration

Enabling DAI in a VLAN

By default, DAI is disabled in VLANs.

Run the **ip arp inspection vlan** *vlan-id* command to enable DAI in a specific VLAN.

After DAI is enabled in a VLAN, DAI may not take effect on all ports in the VLAN. A DHCP Snooping trusted port does not perform DAI check.

Disabling DAI in a VLAN

By default, DAI is disabled in VLANs.

After DAI is enabled in a VLAN, you can run the **no ip arp inspection vlan** *vlan-id* command to disable DAI.

Disabling DAI in a VLAN does not mean disabling packet validity check on all ports in the VLAN. The ports with ARP Check effective still check the validity of received ARP packets.

17.3.2 DAI Trusted Port

Configure specific device ports as DAI trusted ports.

Working Principle

The validity of ARP packets received from trusted ports is not checked. The ARP packets received from untrusted ports are checked against the user records in a security database.

Related Configuration

Configuring DAI Trusted Ports

By default, all ports are untrusted ports.

Run the **ip arp inspection trust** command to set ports to trusted state.

A port already enabled with access security control cannot be set to DAI trusted state. To set the port to DAI trusted state, first disable access security control.

In normal cases, uplink ports (ports connected to network devices) can be configured as DAI trusted ports.

17.4 Configuration

Configuration	Description and Command	
Configuring DAI	(Optional) It is used to enable ARP packet validity check.	
	ip arp inspection vlan	Enables DAI.
	ip arp inspection trust	Configures DAI trusted ports.

17.4.1 Configuring DAI

Configuration Effect

- Check the validity of incoming ARP packets in a specific VLAN.

Notes

- DAI cannot be enabled on DHCP Snooping trusted ports.

Configuration Steps

Enabling ARP Packet Validity Check in a Specific VLAN

- Optional.
- Perform this configuration when you need to enable ARP packet validity check on all ports in a VLAN.
- Perform this configuration on FS access devices unless otherwise specified.

Configuring DAI Trusted Ports

- Optional.
- It is recommended to configure uplink ports as DAI trusted ports after DAI is enabled. Otherwise, the uplink ports enabled with other security features and set to trusted state accordingly may filter out valid ARP packets due to the absence of DAI user entries.
- Perform this configuration on FS access devices unless otherwise specified.

Configuring the ARP Packet Reception Rate

- For details, see the rate limit command description in the *Configuring the NFPP*.

Verification

- Construct invalid ARP packets by using a packet transfer tool and check whether the packets are filtered out on DAI-enabled

devices.

- Run the **show** command to check the device configuration.

Related Commands

Enabling DAI

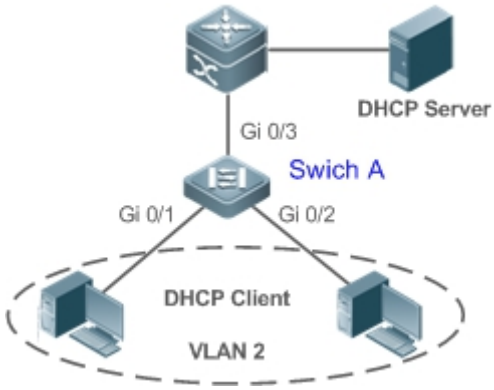
Command	**ip arp inspection vlan** { *vlan-id* | *word* }
Parameter Description	*vlan-id*: Indicates a VLAN ID. *word*: Indicates the VLAN range string, such as 1, 3–5, 7, and 9–11.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring DAI Trusted Ports

Command	**ip arp inspection trust**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Use this command to configure a DAI trusted port so that the ARP packets received by the port can pass through without validity check.

Configuration Example

Allowing Users' PCs to Use only Addresses Allocated by a DHCP Server to Prevent ARP Spoofing

Scenario **Figure 17- 3**	
Configuration Steps	Enable DHCP Snooping on the access switch (Switch A) and configure its uplink port (GigabitEthernet 0/3) connected to the valid DHCP server as a trusted port. Enable IP Source Guard on Switch A. Enable DAI.
Switch A	A#configure terminal Enter configuration commands, one per line. End with CNTL/Z.

	A(config)#vlan 2 A(config-vlan)#exit A(config)#interface range gigabitEthernet 0/1-2 A(config-if-range)#switchport access vlan 2 A(config-if-range)#ip verify source A(config-if-range)#exit A(config)#ip dhcp snooping A(config)#ip arp inspection vlan 2 A(config)#interface gigabitEthernet 0/3 A(config-if-GigabitEthernet 0/3)#switchport access vlan 2 A(config-if-GigabitEthernet 0/3)#ip dhcp snooping trust A(config-if-GigabitEthernet 0/3)#ip arp inspection trust
Verification	● Check whether DHCP Snooping, IP Source Guard, and DAI are enabled and whether trusted ports are configured correctly. ● Check whether the uplink port on Switch A is a DHCP Snooping trusted port. ● Check whether DAI is enabled successfully in the VLAN and the uplink ports are DAI trusted ports.
Switch A	A#show running-config A#show ip dhcp snooping A#show ip arp inspection vlan

Common Errors

- A port with security control enabled is configured as a DAI trusted port.

17.5 Monitoring

Displaying

Description	Command
Displays the DAI state of a specific VLAN.	**show ip arp inspection vlan** [*vlan-id* | *word*]
Displays the DAI configuration state of each Layer-2 port.	**show ip arp inspection interface**

18 Configuring IP Source Guard

18.1 Overview

The IP Source Guard function realizes hardware-based IP packet filtering to ensure that only the users having their information in the binding database can access networks normally, preventing users from forging IP packets.

18.2 Applications

Application	Description
Guarding Against IP/MAC Spoofing Attack	In network environments, users set illegal IP addresses and malicious users launch attacks through forging IP packets.

18.2.1 Guarding Against IP/MAC Spoofing Attack

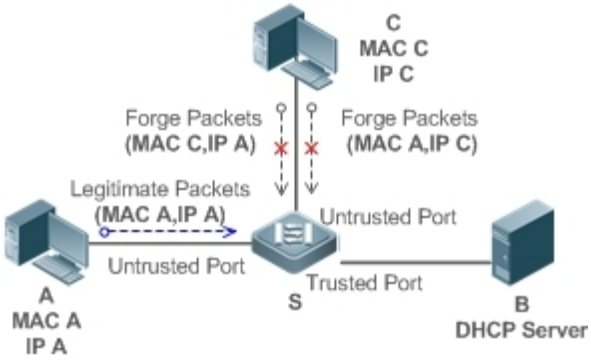
Scenario

Check the IP packets from DHCP untrusted ports. Forged IP packets will be filtered out based on the IP or IP-MAC field.

For example, in the following figure, the IP packets sent by DHCP clients are checked.

- The Source IP Address fields of IP packets should match DHCP-assigned IP addresses.
- The Source MAC Address fields of layer-2 packets should match the MAC addresses in DHCP request packets from clients.

Figure 18- 1



Remarks:	S is a network access server (NAS). A and C are user PCs. B is a DHCP server within the control area.

Deployment

- Enable DHCP Snooping on S to realize DHCP monitoring.
- Set all downlink ports on S as DHCP untrusted ports.
- Enable IP Source Guard on S to realize IP packet filtering.
- Enable IP–MAC match mode for IP Source Guard on S, filtering IP packets based on IP and MAC addresses.

18.3 Features

Basic Concepts

Source IP Address

Indicate the source IP address field of an IP packet.

Source MAC Address

Indicate the source MAC address field of an IP packet.

IP-based Filtering

Indicate a policy of IP packet filtering, where only the source IP addresses of all IP packets (except DHCP packets) passing through a port are checked. It is the default filtering policy of IP Source Guard.

IP-MAC based Filtering

A policy of IP packet filtering, where both the source IP addresses and source MAC addresses of all IP packets are checked, and only those user packets with these IP addresses and MAC addresses existing in the binding database are permitted.

Address Binding Database

As the basis of security control of the IP Source Guard function, the data in the address binding database comes from two ways: the DHCP Snooping binding database and static configuration. When IP Source Guard is enabled, the data of the DHCP Snooping binding database is synchronized to the address binding database of IP Source Guard, so that IP packets can be filtered strictly through IP Source Guard on a device with DHCP Snooping enabled.

Excluded VLAN

By default, when IP Source Guard is enabled on a port, it is effective to all the VLANs under the port. Users may specify excluded VLANs, within which IP packets are not checked and filtered, which means that such IP packets are not controlled by IP Source Guard. At most 32 excluded VLANs can be specified for a port.

Overview

Feature	Description
Checking Source Address Fields of Packets	Filter the IP packets passing through ports by IP-based or IP-MAC based filtering.

18.3.1 Checking Source Address Fields of Packets

Filter the IP packets passing through ports based on source IP addresses or on both source IP addresses and source MAC addresses to prevent malicious attack by forging packets. When there is no need to check and filter IP packets within a VLAN, an excluded VLAN can be specified to release such packets.

Working Principle

When IP Source Guard is enabled, the source addresses of packets passing through a port will be checked. The port can be a wired switching port, a layer-2 aggregate port (AP), or a layer-2 encapsulation sub-interface. Such packets will pass the port only when the source address fields of the packets match the set of the address binding records generated by DHCP Snooping, or the static configuration set by the administrator. There are two matching modes as below.

IP-based Filtering

Packets are allowed to pass a port only if the source IP address fields of them belong to the address binding database.

IP-MAC Based Filtering

Packets are allowed to pass a port only when both the layer-2 source MAC addresses and layer-3 source IP addresses of them match an entry in the address binding database.

Specifying Excluded VLAN

Packets within such a VLAN are allowed to pass a port without check or filtering.

Related Configuration

Enabling IP Source Guard on a Port

By default, the IP Source Guard is disabled on ports.

It can be enabled using the **ip verify source** command.

Usually IP Source Guard needs to work with DHCP Snooping. Therefore, DHCP Snooping should also be enabled. DHCP Snooping can be enabled at any time on FS devices, either before or after IP Source Guard is enabled.

Configuring a Static Binding

By default, legal users passing IP Source Guard check are all from the binding database of DHCP Snooping.

Bound users can be added using the **ip source binding** command.

Specifying an Excluded VLAN

By default, IP Source Guard is effective to all the VLANs under a port.

Excluded VLANs may be specified which are exempted from IP Source Guard using the **ip verify source exclude-vlan** command.

Excluded VLANs can be specified only after IP Source Guard is enabled on a port. Specified excluded VLANs will be deleted automatically when IP Source Guard is disabled on a port.

The above-mentioned port can be a wired switching port, a layer-2 AP port or a layer-2 encapsulation sub-interface..

18.4 Configuration

Configuration	Description and Command	
Configuring IP Source Guard	(Mandatory) It is used to enable IP Source Guard.	
	ip verify source	Enables IP Source Guard on a port.
	ip source binding	Configures a static binding.
	Ip verify source exclude-vlan	Specifies an excluded VLAN for IP Source Guard.

18.4.1 Configuring IP Source Guard

Configuration Effect

- Check the source IP addresses of input IP packets.

Notes

- When IP Source Guard is enabled, IP packets forwarding may be affected. In general case, IP Source Guard is enabled together with DHCP Snooping.
- IP Source Guard cannot be configured on the trusted ports controlled by DHCP Snooping.
- IP Source Guard cannot be configured on the global IP+MAC exclusive ports.
- IP Source Guard can be configured and enabled only on wired switch ports, Layer-2 AP ports, Layer-2 encapsulation sub-ports. In a wired access scenario, it is supposed to be configured in the interface configuration mode.

Configuration Steps

- Enable DHCP Snooping.
- Enable IP Source Guard.

Verification

Use the monitoring commands to display the address binding database of IP Source Guard.

Related Commands

Enabling IP Source Guard on a Port

Command	**ip verify source** [**port-security**]
Parameter Description	**port-security**: Enable IP-MAC based filtering.
Command	Interface configuration mode
Usage Guide	Detection of users based on IP address or both IP and MAC addresses can be realized by enabling IP Source Guard for a port.

Configuring a Static Binding

Command	**ip source binding** *mac-address* **{ vlan** *vlan-id* **}** *ip-address* **{ interface** *interface-id* **| ip-mac | ip-only }**
Parameter Description	**mac-address:** The MAC address of a static binding **vlan-id:** The VLAN ID of a static binding. It indicates the outer VLAN ID of a QINQ-termination user. **ip-address:** The IP address of a static binding *interface-id***:** The Port ID (PID) of a static binding **ip-mac:** IP-MAC based mode **ip-only:** IP-based mode
Configuration Mode	Global configuration mode
Usage Guide	Through this command, legitimate users can pass IP Source Guard detection instead of being controlled by DHCP.

Specifying an Exception VLAN for IP Source Guard

Command	**ip verify source exclude-vlan** *vlan-id*
Parameter Description	**vlan-id:** A VLAN ID exempted from IP Source Guard on a port
Command	Interface configuration mode
Usage Guide	By using this command, the specified VLANs under a port where IP Source Guard function is enabled can be exempted from check and filtering.

Configuration Example

Enabling IP Source Guard on Port 1

Configuration Steps	● Enable DHCP Snooping. ● Enable IP Source Guard.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ip verify source FS(config-if-GigabitEthernet 0/1)# end
Verification	Displays the address filtering table of IP Source Guard.
	FS# show ip verify source

Configuring a Static Binding

Configuration Steps	● Enable DHCP Snooping. ● Enable IP Source Guard. ● Configure a static binding.
	FS# configure terminal FS(config)# ip source binding 00d0.f801.0101 vlan 1 192.168.4.243 interface GigabitEthernet 0/3 FS(config)# end
Verification	Displays the address filtering table of IP Source Guard.
	FS# show ip verify source NO. INTERFACE FilterType FilterStatus IPADDRESS MACADDRESS VLAN TYPE ----- ---------- ---------- ---------- ---------- ---------- ---- ---------- 1 GigabitEthernet 0/3 UNSET Inactive-restrict-off 192.168.4.243 00d0.f801.0101 1 Static 2 GigabitEthernet 0/1 IP-ONLY Active Deny-All

Specifying an Excluded VLAN

Configuration Steps	● Enable DHCP Snooping. ● Enable IP Source Guard.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ip verify source FS(config-if-GigabitEthernet 0/1)# ip verify source exclude-vlan 1 FS(config-if)# end
Verification	Display the configuration of excluded VLANs specified on a port.

Configuration Steps	● Enable DHCP Snooping. ● Enable IP Source Guard.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ip verify source FS(config-if-GigabitEthernet 0/1)# ip verify source exclude-vlan 1 FS(config-if)# end
Verification	Display the configuration of excluded VLANs specified on a port.
	FS# show run

Common Errors

- Enable IP Source Guard on a trusted port under DHCP Snooping.
- Specify an excluded VLAN before IP Source Guard is enabled.

18.5 Monitoring

Displaying

Description	Command
Displays the address filtering table of IP Source Guard.	**show ip verify source** [**interface** *interface-id*]
Displays the address binding database of IP Source Guard.	**show ip source binding**

19 Configuring IPv6 Source Guard

19.1 Overview

IPv6 Source Guard binding allows IPv6 packets to be filtered by hardware so as to ensure that only the users having corresponding information in the IPv6 packet hardware filtering database can access the Internet, thus preventing users from configuring IP addresses without authorization or fabricating IPv6 packets.

19.2 Applications

Application	Description
Prevention of IPv6/MAC Spoofing	There are malicious users on a network who fabricate IPv6 packets to launch an attack.

19.2.1 Prevention of IPv6/MAC Spoofing

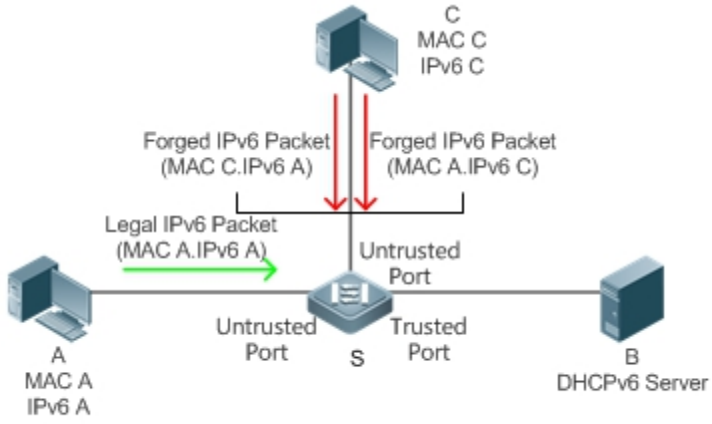
Scenario

When checking the IPv6 packets from the untrusted DHCPv6 ports, you may check IPv6 fields only or IPv6+MAC fields, thereby filtering fabricated IPv6 packets.

As shown in the following figure, IPv6 packets sent from the Dynamic Host Configuration Protocol version 6 (DHCPv6) client will be checked.

- The source address fields of IPv6 packets must match IPv6 addresses assigned by the DHCPv6 client.
- The source media access control (MAC) addresses of Layer-2 packets must match those assigned by DHCPv6 Snooping to hardware filtering records.

Figure 20- 1



Remarks	S is an access device. A and C are user PCs. B is a controlled DHCPv6 server.

Deployment

- Enable DHCPv6 Snooping on the access device S for DHCPv6 monitoring.

- Set all the downstream interfaces on the access device S as untrusted DHCPv6 ports.
- On the access device S, enable IPv6 Source Guard for IPv6 packet filtering.
- On the access device S, set the match mode of IPv6 Source Guard as IPv6+MAC for checking MAC fields and IPv6 fields of IPv6 packets.

19.3 Features

Basic Concepts

Source IPv6

Indicates the source IPv6 address fields of IPv6 packets

Source MAC

Indicates the source MAC address fields of Layer-2 packets

Source IPv6-based Filtering

The source IPv6-based filtering policy checks only the source IPv6 addresses of all IPv6 packets (except DHCP packets) passing through the interface. The source IPv6-based filtering policy is the default filtering policy of IPv6 Source Guard.

Source IPv6+Source MAC-based Filtering

The source IPv6-based filtering policy checks the source IPv6+source MAC of all IPv6 packets, and only the user packets saved in the database for binding user records are allowed to pass through.

Database for Binding User Records

The database for binding user records is the basis for IPv6 Source Guard security control. Currently, the data in the database binding user records come from the following two sources. One is the DHCPv6 Snooping binding database. After IPv6 Source Guard is enabled, the information in the DHCPv6 Snooping binding database is synchronized to the user binding database of IPv6 Source Guard so that IPv6 Source Guard can filter the IPv6 packets of the client on the device where DHCPv6 Snooping is enabled. The other is users' static configuration.

Overview

Feature	Description
Checking the Source Address Fields of Packets	Filters the IPv6 packets passing through the interface based on source IPv6 or source IPv6+source MAC.

19.3.1 Checking the Source Address Fields of Packets

Filter the IPv6 packets transiting the port based on source IPv6 or source IPv6+source MAC, thereby preventing malicious users from fabricating packets to launch an attack.

Working Principle

After IPv6 Source Guard is enabled, the device checks the source addresses of the packets passing through the port. The port can be a wired switch port, Layer-2 aggregate port (AP) or Layer-2 encapsulation sub interface. Only the packets whose source address fields match the user binding record set generated by DHCPv6 Snooping or the user set statically configured by the administrator can pass

through the port. There are two matching methods:

Source IPv6 Address-based Filtering

If IPv6 fields of a packet belong to the identity association in the user binding records, the packet is allowed to pass through the port.

IPv6+MAC Address-based Filtering

Only when Layer-2 MAC and Layer-3 IPv6 of a packet completely match a certain record in the set of authenticated users can the packet pass through the port.

Related Configuration

Enabling IPv6 Source Guard on a Port

By default, IPv6 Source Guard is disabled on a port.

IPv6 Source Guard of the port can be enabled or disabled by running the **ipv6 verify source** command.

Typically, DHCPv6 Snooping is used together with IPv6 Source Guard , so DHCPv6 Snooping needs to be enabled. Timing for enabling DHCPv6 Snooping is not limited on FS devices. You can enable DHCPv6 Snooping before or after IPv6 Source Guard is enabled.

Configuring Static IPv6 Source Guard Users

By default, all sets of authenticated users checked by IPv6 Source Guard are from the bound users of DHCPv6 Snooping.

Run the **ipv6 source binding** command to add extra user binding records.

19.4 Configuration

Configuration	Description and Command	
Configuring IPv6 Source Guard	(Mandatory) It is used to enable IPv6 Source Guard.	
	ipv6 verify source	Enables IPv6 Source Guard on a port.
	ipv6 source binding	Configure statically bound users.

19.4.1 Configuring IPv6 Source Guard

Configuration Effect

- Check the source IPv6 fields entered into IPv6 packets.

Notes

- IPv6 Source Guard is based on DHCPv6 Snooping; that is to say, interface-based IPv6 Source Guard takes effect only on the untrusted ports controlled by DHCPv6 Snooping. If configured on trusted ports or the interfaces on VLANs not controlled by DHCPv6 Snooping, the function will not take effect.

Configuration Steps

- Enable DHCPv6 Snooping.
- Enable IPv6 Source Guard.

Verification

Use the monitoring command provided by the device to view the user filtering entries of IPv6 Source Guard.

Related Commands

Enabling IPv6 Source Guard on a Port

Command	**ipv6 verify source** [**port-security**]
Parameter Description	**port-security**: Configures IPv6 Source Guard to perform IPv6+MAC-based detection.
Command Mode	Interface mode
Usage Guide	By enabling IPv6 Source Guard on a port through this command, you can detect users based on IPv6 or IPv6+MAC.

Adding Information on Static Users to Ipv6 Source Address Binding Database

Command	**ipv6 source binding** *mac-address* **vlan** *vlan-id ipv6-address* **{ interface** *interface-id* **| ip-mac | ip-only }**
Parameter Description	***mac-address***: Indicates the MAC address of a statically added user. ***vlan-id***: Indicates the VLAN ID of a statically added user. ***ipv6-address***: Indicates the IPv6 addresses of a statically added user. ***interface-id***: Indicates the wired access interface for a statically added user. **wlan-id**: Indicates the wireless access WLAN for a statically added user. **ip-mac**: Indicates that the global binding mode is IPv6+MAC binding mode. **ip-only**: Indicates that the global binding mode is IPv6 binding mode only.
Command Mode	Global configuration mode
Usage Guide	By running this command, some users can pass the check of IPv6 Source Guard without being controlled by DHCPv6.

Configuration Example

Enabling IPv6 Source Guard on a Port

Configuration Steps	● Enable DHCPv6 Snooping.
	FS(config)# ipv6 access-list v6-list FS(config-ipv6-nacl)# permit ipv6 fe80::/10 any FS(config-ipv6-nacl)# permit ipv6 ::/128 any FS(config-ipv6-nacl)# exit FS(config)# security global access-group v6-list
	● Enable IPv6 Source Guard.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ipv6 verify source FS(config-if-GigabitEthernet 0/1)# end
Verification	View the user filtering entries of IPv6 Source Guard.

Configuration Steps	● Enable DHCPv6 Snooping.
	FS(config)# ipv6 access-list v6-list FS(config-ipv6-nacl)# permit ipv6 fe80::/10 any FS(config-ipv6-nacl)# permit ipv6 ::/128 any FS(config-ipv6-nacl)# exit FS(config)# security global access-group v6-list
	● Enable IPv6 Source Guard.
	FS(config)# interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# ipv6 verify source FS(config-if-GigabitEthernet 0/1)# end
Verification	View the user filtering entries of IPv6 Source Guard.
	FS# show ipv6 source binding

Adding a Statically Bound User

Configuration Steps	● Enable DHCPv6 Snooping.(Omitted) ● Enable IPv6 Source Guard.(Omitted) ● Add a static user.
	FS# configure terminal FS(config)# ipv6 source binding 0001.0002.0006 vlan 1 2008::1 ip-mac FS(config)# end

Verification	View the user filtering entries of IPv6 Source Guard. FS# show ipv6 source binding Total number of bindings: 7 NO. Filter Type Filter Status IPv6 Address MACAddress VLAN Type Interface ------ ---------- ---------- ---------- ---------- ---- ---------- ---------- 1 IPv6+MAC Inactive-system-error 2000::127 0001.0002.0003 1 Static Global 2 IPv6-ONLY Active 2008::4 0001.0002.0004 1 DHCPv6-Snooping GigabitEthernet 0/5 3 IPv6-ONLY Active 2008::7 0001.0002.0007 1 Static Global 4 IPv6+MAC Active 2008::1 0001.0002.0006 1 Static Global 5 UNSET Inactive-restrict-off 2008::9 0001.0002.0009 1 DHCPv6-Snooping GigabitEthernet 0/1 6 IPv6-ONLY Active Deny-All GigabitEthernet 0/5

Common Errors

- IPv6 Source Guard is enabled on the trusted DHCPv6 Snooping port.

19.5 Monitoring

Displaying

Description	Command
Displays information on the IPv6 source address binding database.	**show ipv6 source binding**

20 Configuring Gateway-targeted ARP Spoofing Prevention

20.1 Overview

Gateway-targeted Address Resolution Protocol (ARP) spoofing prevention effectively prevents gateway-targeted ARP spoofing by checking on the logical port whether the source IP addresses of ARP packets (Sender IP fields of ARP packets) are the self-configured gateway IP addresses.

Protocols and Standards

RFC 826: Ethernet Address Resolution Protocol

20.2 Applications

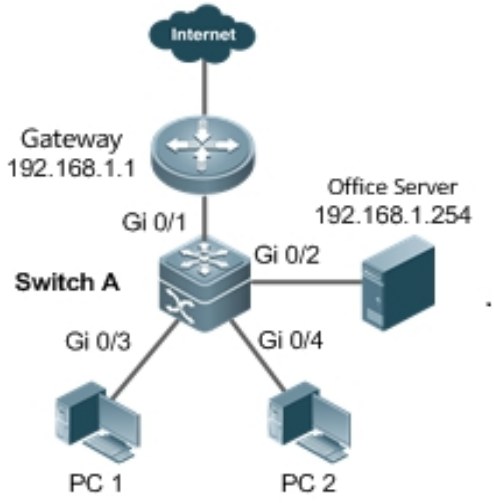
Application	Description
Typical Application of Gateway-targeted ARP Spoofing Prevention	Blocks ARP spoofing packets with forged gateway address and intranet server IP addresses to ensure that users can access the Internet.

20.2.1 Typical Application of Gateway-targeted ARP Spoofing Prevention

Scenario

- PC users access the office server through the access device Switch A, and connect to external networks through the gateway.
- If any users legally use forged gateway IP addresses or server IP addresses to perform ARP spoofing, the other users cannot access the Internet and the server.
- The ARP spoofing packets with forged gateway address and intranet server IP addresses must be blocked to ensure that users can access the Internet.

Figure 20- 1 Typical Topology of Gateway-targeted ARP Spoofing Prevention



Deployment

- On the access switch (Switch A), enable gateway-targeted spoofing prevention on the ports (Gi 0/3 and Gi 0/4 in this case) directly connected to the PC. The gateway addresses include intranet gateway address and intranet server address.

20.3 Features

Basic Concepts

ARP

ARP is a TCP/IP protocol that obtains physical addresses according to IP addresses. Its function is as follows: The host broadcasts ARP requests to all hosts on the network and receives the returned packets to determine physical addresses of the target IP addresses, and saves the IP addresses and hardware addresses in the local ARP cache, which can be directly queried in response to future requests. On the same network, all the hosts using the ARP are considered as mutually trustful to each other. Each host on the network can independently send ARP response packets; the other hosts receive the response packets and record them in the local ARP cache without detecting their authenticity. In this way, attackers can send forged ARP response packets to target hosts so that the messages sent from these hosts cannot reach the proper host or reach a wrong host, thereby causing ARP spoofing.

Gateway-targeted ARP Spoofing

When User A sends an ARP packet requesting the media access control (MAC) address of a gateway, User B on the same VLAN also receives this packet, and User B can send an ARP response packet, passing off the gateway IP address as the source IP address of the packet, and User B's MAC address as the source MAC address. This is called gateway-targeted ARP spoofing. After receiving the ARP response, User A regards User B's machine as the gateway, so all the packets sent from User A to the gateway during communication will be sent to User B. In this way, User A's communications are intercepted, thereby causing ARP spoofing.

Overview

Feature	Description
Gateway-targeted ARP Spoofing Prevention	Blocks ARP spoofing packets with forged gateway address and intranet server IP addresses to ensure that users can access the Internet.

20.3.1 Gateway-targeted ARP Spoofing Prevention

Working Principle

Gateway-targeted Spoofing Prevention

Gateway-targeted ARP spoofing prevention effectively prevents ARP spoofing aimed at gateways by checking on the logical port whether the source IP addresses of ARP packets are the self-configured gateway IP addresses. If an ARP packet uses the gateway address as the source IP address, the packet will be discarded to prevent users from receiving wrong ARP response packets. If not, the packet will not be handled. In this way, only the devices connected to the switch can send ARP packets, and the ARP response packets sent from the other PCs which pass for the gateway are filtered by the switch.

Related Configuration

Configuring Gateway-targeted Spoofing Prevention Addresses

- By default, no gateway-targeted ARP spoofing prevention address is configured.
- Run the **anti-arp-spoofing ip** command to configure the gateway-targeted ARP spoofing prevention addresses.

20.4 Configuration

Configuration	Description and Command	
Configuring Gateway-targeted Spoofing Prevention	Optional.	
	anti-arp-spoofing ip	Configures gateway-targeted ARP spoofing prevention on the logical port and specifies the gateway IP address.

20.4.1 Configuring Gateway-targeted Spoofing Prevention

Configuration Effect

Enable gateway-targeted ARP spoofing prevention.

Configuration Steps

Configuring Gateway-targeted Spoofing Prevention

- Gateway-targeted ARP spoofing prevention is mandatory. It must be enabled.

Verification

- Run the **show run** command to check configuration.
- Run the **show anti-arp-spoofing** command to display all data on gateway-targeted ARP spoofing prevention.

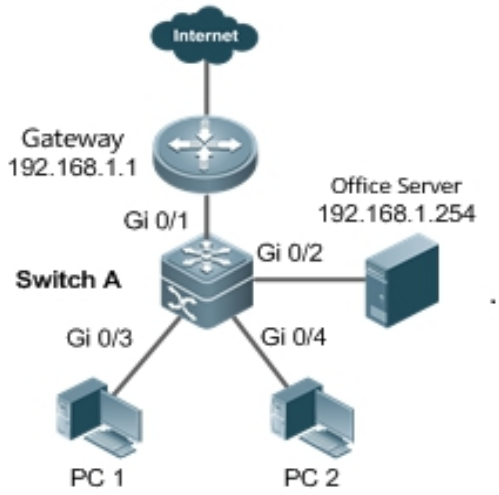
Related Commands

Configuring Gateway-targeted Spoofing Prevention

Command	**anti-arp-spoofing ip** *ip-address*
Parameter Description	*ip-address*: Indicates the IP address of the gateway.
Command Mode	Interface configuration mode
Usage Guide	Gateway-targeted ARP Spoofing prevention is supported only on Layer-2 ports.

Configuration Example

Configuring Gateway-targeted Spoofing Prevention

Scenario **Figure 20- 2**	
	PC users access the office server through the access device Switch A, and connect external networks through the gateway. If any users legally use forged gateway IP addresses or server IP addresses to perform ARP spoofing, the other users cannot access the Internet or the server. The ARP spoofing packets with forged gateway address and intranet server IP addresses must be blocked to ensure that users can access the Internet.
Configuration Steps	Enable gateway-targeted spoofing prevention on the port directly connected to the PC.
	SwitchA# configure terminal Enter configuration commands, one per line. End with CNTL/Z. SwitchA(config)#interface range gigabitEthernet 0/3-4 SwitchA(config-if-range)# anti-arp-spoofing ip 192.168.1.1 SwitchA(config-if-range)# anti-arp-spoofing ip 192.168.1.254
Verification	Run the **show anti-arp-spoofing** command to check for data on gateway-targeted ARP spoofing prevention.
	SwitchA#show anti-arp-spoofing NO PORT IP STATUS ----- ---------- ---------- ---------- 3 Gi0/3 192.168.1.1 active 4 Gi0/3 192.168.1.254 active 5 Gi0/4 192.168.1.1 active 6 Gi0/4 192.168.1.254 active

20.5 Monitoring

Displaying

Description	Command
Displays all data on gateway-targeted ARP spoofing prevention.	**show anti-arp-spoofing**

21 Configuring NFPP

21.1 Overview

Network Foundation Protection Policy (NFPP) provides guards for switches.

Malicious attacks are always found in the network environment. These attacks bring heavy burdens to switches, resulting in high CPU usage and operational troubles. These attacks are as follows:

Denial of Service (DoS) attacks may consume lots of memory, entries, or other resources of a switch, which will cause system service termination.

Massive attack traffic is directed to the CPU, occupying the entire bandwidth of the CPU. In this case, normal protocol traffic and management traffic cannot be processed by the CPU, causing protocol flapping or management failure. The forwarding in the data plane will also be affected and the entire network will become abnormal.

A great number of attack packets directed to the CPU consume massive CPU resources, making the CPU highly loaded and thereby influencing device management and performance.

NFPP can effectively protect the system from these attacks. Facing attacks, NFPP maintains the proper running of various system services with a low CPU load, thereby ensuring the stability of the entire network.

21.2 Applications

Application	Description
Attack Rate Limiting	Due to various malicious attacks such as ARP attacks and IP scanning attacks in the network, the CPU cannot process normal protocol and management traffics, causing protocol flapping or management failure. The NFPP attack rate limiting function is used to limit the rate of attack traffic or isolate attack traffic to recover the network.
CentralizedBandwidth Allocation	If normal service traffics are too large, you need to classify and prioritize the traffics. When a large number of packets are directed to the CPU, the CPU will be highly loaded, thereby causing device management or device running failure. The centralized bandwidth distribution function is used to increase the priority of such traffics so that switches can run stably.

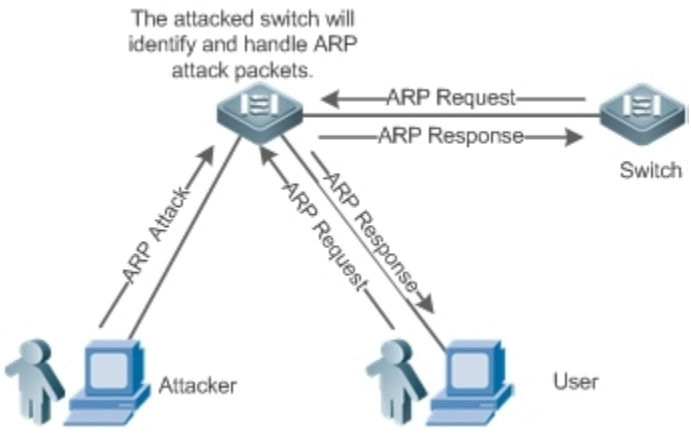
21.2.1 Attack Rate Limiting

Scenario

NFPP supports attack detection and rate limiting for various types of packets, including Address Resolution Protocol (ARP), Internet Control Message Protocol (ICMP), and Dynamic Host Configuration Protocol (DHCP) packets. It also allows users to define packet matching characteristics and corresponding attack detection and rate limiting policies. The attack rate limiting function takes effect based on types of packets. This section uses ARP packets as an example scenario to describe the application.

If an attacker floods ARP attack packets while CPU capability is insufficient, most of the CPU resources will be consumed for processing these ARP packets. If the rate of attacker's ARP packet rates exceeds the maximum ARP bandwidth specified in the CPU Protect Policy (CPP) of the switch, normal ARP packets may be dropped. As shown in Figure 22- 1, normal hosts will fail to access the network, and the switch will fail to send ARP replies to other devices.

Figure 22- 1



Deployment

- By default, the ARP attack detection and rate limiting function is enabled with corresponding policies configured. If the rate of an attacker's ARP packets exceeds the rate limit, the packets are discarded. If it exceeds the attack threshold, a monitoring user is generated and prompt information is exported.
- If the rate of an attacker's ARP packets exceeds the rate limit defined in CPP and affects normal ARP replies, you can enable attack isolation to discard ARP attack packets based on the hardware and recover the network.

For details about CPP-related configurations, see the *Configuring CPU Protection*.

To maximize the use of NFPP guard functions, modify the rate limits of various services in CPP based on the application environment or use the configurations recommended by the system. You can run the **show cpu-protect summary** command to display the configurations.

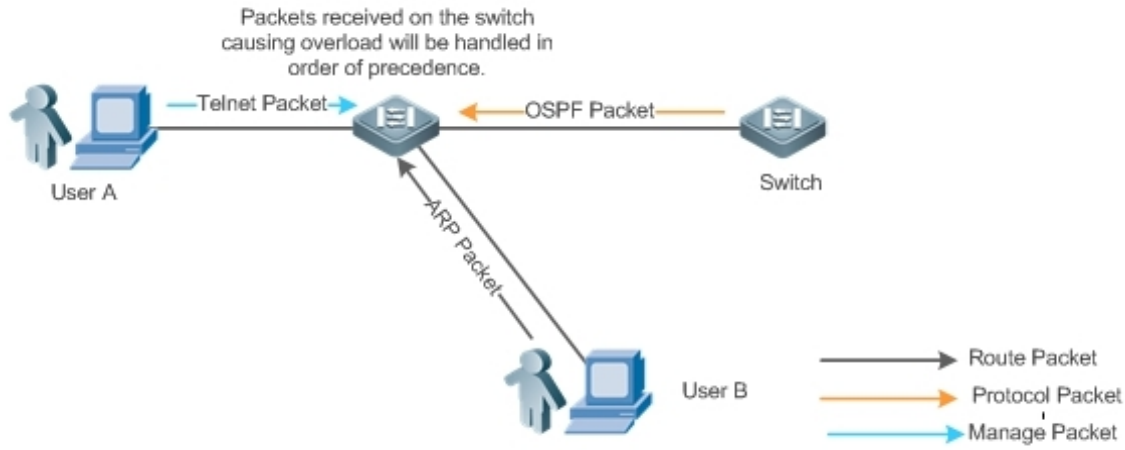
21.2.2 Centralized Bandwidth Allocation

Scenario

A switch classifies services defined in CPP into three types: Manage, Route, and Protocol. Each type of services has an independent bandwidth. Different types of services cannot share their bandwidths. Traffics with bandwidths exceeding the thresholds will be discarded. By such service classification, service packets are processed by orders of precedence.

As shown in Figure 22- 2, the switch receives a large number of Telnet packets, OSPF packets, and ARP packets, causing CPU overload. In this case, the CPU cannot process all packets, and a large quantity of packets are backlogged in the queue, causing various problems such as frequent Telnet disconnection, OSPF protocol flapping, and ARP access failure on hosts.

Figure 22- 2



Deployment

- By default, CPU centralized bandwidth allocation is enabled to assign an independent bandwidth and bandwidth ratio to each type of services. At the time, the CPU first processes Telnet packets to ensure uninterrupted connection of Telnet service, and then processes OSPF packets to maintain OSPF protocol stability, and finally processes ARP packets.
- If the preceding problems still occur in default configurations, you can accordingly adjust the bandwidths and bandwidth ratios of various types of services.

21.3 Features

Basic Concepts

ARP Guard

In local area networks (LANs), IP addresses are mapped to MAC addresses through ARP, which has a significant role in safeguarding network security. ARP-based DoS attacks mean that a large number of unauthorized ARP packets are sent to the gateway through the network, causing the failure of the gateway to provide services for normal hosts. To prevent such attacks, limit the rate of ARP packets and identify and isolate the attack source.

IP Guard

Many hacker attacks and network virus intrusions start from scanning active hosts in the network. Therefore, many scanning packets rapidly occupy the network bandwidth, causing network communication failure.

To solve this problem, FS Layer-3 switches provide IP guard function to prevent hacker scanning and Blaster Worm viruses and reduce the CPU load. Currently, there are mainly two types of IP attacks:

Scanning destination IP address changes: As the greatest threat to the network, this type of attacks not only consumes network bandwidth and increases device load but also is a prelude of most hacker attacks.

Sending IP packets to non-existing destination IP addresses at high rates: This type of attacks is mainly designed for consuming the CPU load. For a Layer-3 device, if the destination IP address exists, packets are directly forwarded by the switching chip without occupying CPU resources. If the destination IP address does not exist, IP packets are sent to the CPU, which then sends ARP requests to query the MAC address corresponding to the destination IP address. If too many packets are sent to the CPU, CPU resources will be consumed. This type of attack is less destructive than the former one.

To prevent the latter type of attack, limit the rate of IP packets and find and isolate the attack source.

ICMP Guard

ICMP is a common approach to diagnose network failures. After receiving an ICMP echo request from a host, the switch returns an ICMP echo reply. The preceding process requires the CPU to process the packets, thereby definitely consuming part of CPU resources. If an attacker sends a large number of ICMP echo requests to the destination device, massive CPU resources on the device will be consumed heavily, and the device may even fail to work properly. This type of attacks is called ICMP flood. To prevent this type of attacks, limit the rate of ICMP packets and find and isolate the attack source.

DHCP Guard

DHCP is widely used in LANs to dynamically assign IP addresses. It is significant to network security. Currently, the most common DHCP attack, also called DHCP exhaustion attack, uses faked MAC addresses to broadcast DHCP requests. Various attack tools on the Internet can easily complete this type of attack. A network attacker can send sufficient DHCP requests to use up the address space provided by the DHCP server within a period. In this case, authorized hosts will fail to request DHCP IP addresses and thereby fail to access the network. To prevent this type of attacks, limit the rate of DHCP packets and find and isolate the attack source.

DHCPv6 Guard

DHCP version 6 (DHCPv6) is widely used in LANs to dynamically assign IPv6 addresses. Both DHCP version 4 (DHCPv4) and DHCPv6 have security problems. Attacks to DHCPv4 apply also to DHCPv6. A network attacker can send a large number of DHCPv6 requests to use up the address space provided by the DHCPv6 server within a period. In this case, authorized hosts will fail to request IPv6 addresses and thereby fail to access the network. To prevent this type of attacks, limit the rate of DHCPv6 packets and find and isolate the attack source.

ND Guard

Neighbor Discovery (ND) is mainly used in IPv6 networks to perform address resolution, router discovery, prefix discovery, and redirection. ND uses five types of packets: Neighbor Solicitation (NS), Neighbor Advertisement (NA), Router Solicitation (RS), Router Advertisement (RA), and Redirect. These packets are called ND packets.

Self-Defined Guard

There are various types of network protocols, including routing protocols such as Open Shortest Path First (OSPF), and Routing Information Protocol (RIP). Various devices need to exchange packets through different protocols. These packets must be sent to the CPU and processed by appropriate protocols. Once the network device runs a protocol, it is like opening a window for attackers. If an attacker sends a large number of protocol packets to a network device, massive CPU resources will be consumed on the device, and what's worse, the device may fail to work properly.

Since various protocols are being continuously developed, protocols in use vary with the user environments. FS devices hereby provide self-defined guard. Users can customize and flexibly configure guard types to meet guard requirements in different user environments.

Overview

Feature	Description
Host-based Rate Limiting and Attack Identification	Limits the rate according to the host-based rate limit and identify host attacks in the network.
Port-based Rate Limiting and Attack Identification	Limits the rate according to the port-based rate limit and identify port attacks.
Monitoring Period	Monitors host attackers in a specified period.
Isolation Period	Uses hardware to isolate host attackers or port attackers in a specified period.
Trusted Hosts	Trusts a host by not monitoring it.

Centralized BandwidthAllocation	Classifies and prioritizes packets.

21.3.1 Host-based Rate Limiting and Attack Identification

Limit the rate of attack packets of hosts and identify the attacks.

Identify ARP scanning.

Identify IP scanning.

Working Principle

Hosts can be identified in two ways: based on the source IP address, VLAN ID, and port and based on the link-layer source MAC address, VLAN ID, and port. Each host has a rate limit and an attack threshold (also called alarm threshold). The rate limit must be lower than the attack threshold. If the attack packet rate exceeds the rate limit of a host, the host discards the packets beyond the rate limit. If the attack packet rate exceeds the attack threshold of a host, the host identifies and logs the host attacks, and sends traps.

ARP scanning attack may have occurred if ARP packets beyond the scanning threshold received in the configured period meet either of the following conditions:

- The link-layer source MAC address is fixed but the source IP address changes.
- The link-layer source MAC address and source IP address are fixed but the destination IP address continuously changes.

Among IP packets beyond the scanning threshold received in the configured period, if the source IP address remains the same while the destination IP address continuously changes, IP scanning attack may have occurred.

When NFPP detects a specific type of attack packets under a service, it sends a trap to the administrator. If the attack traffic persists, NFPP will not resend the alarm until 60 seconds later.

To prevent CPU resource consumption caused by frequent log printing, NFPP writes attack detection logs to the buffer, obtains them from the buffer at a specified rate, and prints them. NFPP does not limit the rate of traps.

Related Configuration

Use ARP guard as an example:

Configuring the Global Host-based Rate Limit, Attack Threshold, and Scanning Threshold

In NFPP configuration mode:

Run the **arp-guard rate-limit** {**per-src-ip** | **per-src-mac**} *pps* command to configure rate limits of hosts identified based on the source IP address, VLAN ID, and port and hosts identified based on the link-layer source MAC address, VLAN ID, and port.

Run the **arp-guard attack-threshold** {**per-src-ip** | **per-src-mac**} *pps* command to configure attack thresholds of hosts identified based on the source IP address, VLAN ID, and port and hosts identified based on the link-layer source MAC address, VLAN ID, and port.

Run the **arp-guard scan-threshold** *pkt-cnt* command to configure the ARP scanning threshold.

Configuring Host-based Rate Limit and Attack Threshold, and Scanning Threshold on an Interface

In interface configuration mode:

Run the **nfpp arp-guard policy** {**per-src-ip | per-src-mac**} *rate-limit-pps attack-threshold-pps* command to configure rate limits and attack thresholds of hosts identified based on the source IP address, VLAN ID, and port and hosts identified based on the link-layer source MAC address, VLAN ID, and port on an interface.

Run the **nfpp arp-guard scan-threshold** *pkt-cnt* command to configure the scanning threshold on an interface.

Only ARP guard and IP guard support anti-scanning at present.

21.3.2 Port-based Rate Limiting and Attack Identification

Working Principle

Each port has a rate limit and an attack threshold. The rate limit must be lower than the attack threshold. If the packet rate exceeds the rate limit on a port, the port discards the packets. If the packet rate exceeds the attack threshold on a port, the port logs the attacks and sends traps.

Related Configuration

Use ARP guard as an example:

Configuring the Global Port-based Rate Limit and Attack Threshold

In NFPP configuration mode:

Run the **arp-guard rate-limit per-port** *pps* command to configure the rate limit of a port.

Run the **arp-guard attack-threshold per-port** *pps* command to configure the attack threshold of a port.

Configuring Port-based Rate Limit and Attack Threshold on an Interface

In interface configuration mode:

Run the **nfpp arp-guard policy per-port** *rate-limit-pps attack-threshold-pps* command to configure the rate limit and attack threshold of a port.

21.3.3 Monitoring Period

Working Principle

The monitoring user provides information about attackers in the current system. If the isolation period is 0 (that is, not isolated), the guard module automatically performs software monitoring on attackers in the configured monitoring period. If the isolation period is set to a non-zero value, the guard module automatically isolates the hosts monitored by software.

During software monitoring, if the isolation period is set to a non-zero value, the guard module automatically isolates the attacker and sets the timeout period as the isolation period.

The monitoring period is valid only when the isolation period is 0.

Related Configuration

Use ARP guard as an example:

Configuring the Global Monitoring Period

In NFPP configuration mode:

Run the **arp-guard monitor-period** *seconds* command to configure the monitoring period.

21.3.4 Isolation Period

Working Principle

Isolation is performed by the guard policies after attacks are detected. Isolation is implemented using the filter of the hardware to ensure that these attacks will not be sent to the CPU, thereby ensuring proper running of the device.

Hardware isolation supports two modes: host-based and port-based isolation. At present, only ARP guard supports port-based hardware isolation.

A policy is configured in the hardware to isolate attackers. However, hardware resources are limited. When hardware resources are used up, the system prints logs to notify the administrator.

Related Configuration

Use ARP guard as an example:

Configuring the Global Isolation Period

In NFPP configuration mode:

Run the **arp-guard isolate-period** [*seconds* | **permanent**] command to configure the isolation period. If the isolation period is set to 0, isolation is disabled. If it is set to a non-zero value, the value indicates the isolation period. If it is set to **permanent**, ARP attacks are permanently isolated.

Configuring the Isolation Period on an Interface

In interface configuration mode:

Run the **nfpp arp-guard isolate-period** [*seconds* | **permanent**] command to configure the isolation period. If the isolation period is set to 0, isolation is disabled. If it is set to a non-zero value, the value indicates the isolation period. If it is set to **permanent**, ARP attacks are permanently isolated.

Enabling Isolate Forwarding

In NFPP configuration mode:

Run the **arp-guard isolate-forwarding enable** command to enable isolate forwarding.

Enabling Port-based Ratelimit Forwarding

In NFPP configuration mode:

Run the **arp-guard ratelimit-forwarding enable** command to enable port-based ratelimit forwarding.

At present, only ARP guard supports the configuration of isolate forwarding and ratelimit forwarding.

21.3.5 Trusted Hosts

Working Principle

If you do not want to monitor a host, you can run related commands to trust the host. This trusted host will be allowed to send packets to the CPU.

Related Configuration

Use IP anti-scanning as an example:

Configuring Trusted Hosts

In NFPP configuration mode:

Run the **ip-guard trusted-host** *ip mask* command to trust a host.

Run the **trusted-host** {*mac mac_mask* | *ip mask* | *IPv6/prefixlen*} command to trust a host for a self-defined guard.

21.3.6 Centralized Bandwidth Allocation

Working Principle

Services defined in CPP are classified into three types: Manage, Route, and Protocol. (For details, see the following table.) Each type of service has an independent bandwidth. Different types of services cannot share their bandwidths. Traffics exceeding the bandwidth thresholds are discarded. By such service classification, service packets are processed by orders of precedence.

NFPP allows the administrator to flexibly assign bandwidth for three types of packets based on the actual network environment so that Protocol and Manage packets can be first processed. Prior processing of Protocol packets ensures proper running of protocols, and prior processing of Manage packets ensures proper management for the administrator, thereby ensuring proper running of important device functions and improving the guard capability of the device.

After classified rate limiting, all types of packets are centralized in a queue. When one type of service is processed inefficiently, packets of this service will be backlogged in the queue and may finally use up resources of the queue. NFPP allows the administrator to configure the percentages of these three types of packets in the queue. When the queue length occupied by one type of packets exceeds the value of the total queue length multiplied by the percentage of this packet type, the excessive packets will be discarded. This efficiently prevents one type of packets from exclusively occupying queue resources.

Packet Type	Service Type Defined in CPP
Protocol	tp-guard, dot1x, rldp, rerp, slow-packet, bpdu, dhcps, gvrp, ripng, dvmrp, igmp, ospf, rip, vrrp, ospf3, dhcp-relay-s, dhcp-relay-c, option82
Route	unknown-ipmc, unknown-ipmcv6, ttl1, ttl0, udp-helper, ip4-packet-other, ip6-packet-other, non-ip-packet-other, arp
Manage	ip4-packet-local, ip6-packet-local

For the definitions of service types, see the Configuring CPU Protection.

Related Configuration

Configuring the Maximum Bandwidth of Specified Packets

In global configuration mode:

Run the **cpu-protect sub-interface** { **manage** | **protocol**|**route**} **pps** *pps_value* command to configure the maximum bandwidth of specified packets.

Configuring the Maximum Percentage of Specified Packets in the Queue

In global configuration mode:

Run the **cpu-protect sub-interface** { **manage** | **protocol** | **route**} **percent** *percent_value* command to configure the maximum percentage of specified packets in the queue.

21.4 Configuration

Configuration	Description and Command	
Configuring ARP Guard	**arp-guard enable**	Enables ARP guard globally.
	arp-guard isolate-period	Configures the global ARP-guard isolation period.

Configuration	Description and Command	
	arp-guard isolate-forwarding enable	Enables ARP-guard isolate forwarding.
	arp-guard ratelimit-forwarding enable	Enables APR-guard ratelimit forwarding.
	arp-guard monitor-period	Configures the global ARP-guard monitoring period.
	arp-guard monitored-host-limit	Configures the maximum number of ARP-guard monitored hosts.
	arp-guard rate-limit	Configures the global ARP-guard rate limit.
	arp-guard attack-threshold	Configures the global ARP-guard attack threshold.
	arp-guard scan-threshold	Configures the global ARP-guard scanning threshold.
	nfpp arp-guard enable	Enables ARP guard on an interface.
	nfpp arp-guard policy	Configures the APR-guard rate limit and attack threshold on an interface.
	nfpp arp-guard scan-threshold	Configures the APR-guard scanning threshold on an interface.
	nfpp arp-guard isolate-period	Configures the APR-guard isolation period on an interface.
Configuring IP Guard	**ip-guard enable**	Enables IP guard globally.
	ip-guard isolate-period	Configures the global IP-guard isolation period.
	ip-guard monitor-period	Configures the global IP-guard monitoring period.
	ip-guard monitored-host-limit	Configures the maximum number of IP-guard monitored hosts.
	ip-guard rate-limit	Configures the global IP-guard rate limit.
	ip-guard attack-threshold	Configures the global IP-guard attack threshold.
	ip-guard scan-threshold	Configures the global IP-guard scanning threshold.
	ip-guard trusted-host	Configures IP-guard trusted hosts.
	nfpp ip-guard enable	Enables IP guard on an interface.
	nfpp ip-guard policy	Configures the IP-guard rate limit and attack threshold on an interface.
	nfpp ip-guard scan-threshold	Configures the IP-guard scanning threshold on an interface.
	nfpp ip-guard isolate-period	Configures the IP-guard isolation period on an interface.
Configuring ICMP Guard	**icmp-guard enable**	Enables ICMP guard globally.
	icmp-guard isolate-period	Configures the global ICMP-guard isolation period.
	icmp-guard monitor-period	Configures the global ICMP-guard monitoring period.
	icmp-guard monitored-host-limit	Configures the maximum number of ICMP-guard monitored hosts.
	icmp-guard rate-limit	Configures the global ICMP-guard rate limit.
	icmp-guard attack-threshold	Configures the global ICMP-guard attack threshold.
	icmp-guard trusted-host	Configures ICMP-guard trusted hosts.
	nfpp icmp-guard enable	Enables ICMP guard on an interface.
	nfpp icmp-guard policy	Configures the ICMP-guard rate limit and attack threshold on an interface.

Configuration	Description and Command	
	nfpp icmp-guard isolate-period	Configures the ICMP-guard isolation period on an interface.
Configuring DHCP Guard	**dhcp-guard enable**	Enables DHCP guard globally.
	dhcp-guard isolate-period	Configures the global DHCP-guard isolation period.
	dhcp-guard monitor-period	Configures the global DHCP-guard monitoring period.
	dhcp-guard monitored-host-limit	Configures the maximum number of DHCP-guard monitored hosts.
	dhcp-guard rate-limit	Configures the global DHCP-guard rate limit.
	dhcp-guard attack-threshold	Configures the global DHCP-guard attack threshold.
	nfpp dhcp-guard enable	Enables DHCP guard on an interface.
	nfpp dhcp-guard policy	Configures the DHCP-guard rate limit and attack threshold on an interface.
	nfpp dhcp-guard isolate-period	Configures the DHCP-guard isolation period on an interface.
Configuring DHCPv6 Guard	**dhcpv6-guard enable**	Enables DHCPv6 guard globally.
	dhcpv6-guard isolate-period	Configures the global DHCPv6-guard isolation period.
	dhcpv6-guard monitor-period	Configures the global DHCPv6-guard monitoring period.
	dhcpv6-guard monitored-host-limit	Configures the maximum number of DHCPv6-guard monitored hosts.
	dhcpv6-guard rate-limit	Configures the global DHCPv6-guard rate limit.
	dhcpv6-guard attack-threshold { **per-src-mac | per-port**} *pps*	Configures the global DHCPv6-guard attack threshold.
	nfpp dhcpv6-guard enable	EnablesDHCPv6 guard on an interface.
	nfpp dhcpv6-guard policy	Configures the DHCPv6-guard rate limit and attack threshold on an interface.
	nfpp dhcpv6-guard isolate-period	Configures the DHCPv6-guard isolation period on an interface.
Configuring ND Guard	**nd-guard enable**	Enables ND guard globally.
	nd-guard ratelimit-forwarding enable	Enables ND-guard ratelimit forwarding.
	nd-guard rate-limit per-port	Configures the global ND-guard rate limit.
	nd-guard attack-threshold per-port	Configures the global ND-guard attack threshold.
	nfpp nd-guard enable	Enables ND guard on an interface.
	nfpp nd-guard policy per-port	Configures the ND-guard rate limit and attack threshold on an interface.
Configuring a Self-Defined Guard	**define**	Configures the name of a self-defined guard.
	match	Configures **match** fields of a self-defined guard.
	global-policy	Configures the global rate limit and attack threshold of a self-defined guard.
	isolate-period	Configures the global isolation period of a self-defined guard.

Configuration	Description and Command	
	monitor-period	Configures the global monitoring period of a self-defined guard.
	monitored-host-limit	Configures the maximum number of monitored hosts of a self-defined guard.
	trusted-host	Configures trusted hosts of a self-defined guard.
	define *name* **enable**	Enables a self-defined guard globally.
	nfpp define *name* **enable**	Enables a self-defined guard on an interface.
	nfpp define	Configures the rate limit and attack threshold of a self-defined guard on an interface.
Configuring NFPP Logging	**log-buffer entries**	Configures the log buffer size.
	log-buffer logs	Configures the log buffer rate.
	logging vlan	Configures VLAN-based logging filtering.
	logging interface	Configures interface-based logging filtering.
	logging enable	Enables log printing.

21.4.1 Configuring ARP Guard

Configuration Effect

- ARP attacks are identified based on hosts or ports. Host-based ARP attack identification supports two modes: identification based on the source IP address, VLAN ID, and port and identification based on the link-layer source MAC address, VLAN ID, and port. Each type of attack identification has a rate limit and an attack threshold. If the ARP packet rate exceeds the rate limit, the packets beyond the rate limit are discarded. If the ARP packet rate exceeds the attack threshold, the system prints alarm information and sends traps. In host-based attack identification, the system also isolates the attack source.
- ARP guard can also detect ARP scanning attacks. ARP scanning attacks indicate that the link-layer source MAC address is fixed but the source IP address changes, or that the link-layer source MAC address and source IP address are fixed but the destination IP address continuously changes. Due to the possibility of false positive, hosts possibly performing ARP scanning are not isolated and are provided for the administrator's reference only.
- Configure ARP-guard isolation to assign hardware-isolated entries against host attacks so that attack packets are neither sent to the CPU nor forwarded.

Notes

- For a command that is configured both in NFPP configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in NFPP configuration mode.
- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.
- ARP guard prevents only ARP DoS attacks to the switch, but not ARP spoofing or ARP attacks in the network.
- For trusted ports configured for Dynamic ARP Inspection (DAI), ARP guard does not take effect, preventing false positive of ARP traffic over the trusted ports. For details about DAI trusted ports, see the Configuring Dynamic ARP Inspection.

Configuration Steps

Enabling ARP Guard

- (Mandatory) ARP guard is enabled by default.

- This function can be enabled in NFPP configuration mode or interface configuration mode.
- If ARP guard is disabled, the system automatically clears monitored hosts, scanned hosts, and isolated entries on ports.

Configuring the ARP-Guard Isolation Period

- (Optional) ARP-guard isolation is disabled by default.
- If the packet traffic of attackers exceeds the rate limit defined in CPP, you can configure the isolation period to discard packets and therefore to save bandwidth resources.
- The isolation period can be configured in NFPP configuration mode or interface configuration mode.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Enabling ARP-Guard Isolate Forwarding

- (Optional) ARP-guard isolate forwarding is enabled by default.
- To make isolation valid only at the management plane instead of the forwarding plane, you can enable this function.
- This function can be enabled in NFPP configuration mode.

Enabling ARP-Guard Ratelimit Forwarding

- (Optional) This function is enabled by default.
- If the port-based isolation entry takes effect, you can enable this function to pass some of the packets while not discarding all of them.
- This function can be enabled in NFPP configuration mode.

Configuring the ARP-Guard Monitoring Period

- (Mandatory) The default ARP-guard monitoring period is 600 seconds.
- If the ARP-guard isolation period is configured, it is directly used as the monitoring period, and the configured monitoring period will lose effect.
- The monitoring period can be configured in NFPP configuration mode.

Configuring the Maximum Number of ARP-Guard Monitored Hosts

- (Mandatory) The maximum number of ARP-guard monitored hosts is 20,000 by default.
- Set the maximum number of ARP-guard monitored hosts reasonably. As the number of monitored hosts increases, more CPU resources are used.
- The maximum number of ARP-guard monitored hosts can be configured in NFPP configuration mode.
- If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted.
- If the table of monitored hosts is full, the system prints the log "% NFPP_ARP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the ARP-Guard Attack Threshold

- Mandatory.
- To achieve the best ARP-guard effect, you are advised to configure the host-based rate limit and attack threshold based on the following order: Source IP address-based rate limit < Source IP address-based attack threshold <Source MAC address-based rate limit <Source MAC address-based attack threshold.
- The attack threshold can be configured in NFPP configuration mode or interface configuration mode.
- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.
- If the memory cannot be allocated to detected attackers, the system prints the log "%NFPP_ARP_GUARD-4-NO_MEMORY: Failed to alloc memory." to notify the administrator.
- Source MAC address-based rate limiting takes priority over source IP address-based rate limiting while the latter takes priority over port-based rate limiting.

Configuring the ARP-Guard Scanning Threshold

- Mandatory.
- The scanning threshold can be configured in NFPP configuration mode or interface configuration mode.
- The ARP scanning table stores only the latest 256 records. When the ARP scanning table is full, the latest record will overwrite the earliest record.
- ARP scanning attack may have occurred if ARP packets received within 10 seconds meet either of the following conditions:
  - The link-layer source MAC address is fixed but the source IP address changes.
  - The link-layer source MAC address and source IP address are fixed but the destination IP address continuously changes, and the change times exceed the scanning threshold.

Verification

When a host in the network sends ARP attack packets to a switch configured with ARP guard, check whether these packets can be sent to the CPU.

- If the packets exceed the attack threshold or scanning threshold, an attack log is displayed.
- If an isolated entry is created for the attacker, an isolation log is displayed.

Related Commands

Enabling ARP Guard Globally

Command	**arp-guard enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ARP-Guard Isolation Period

Command	**arp-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. **permanent**: Indicates permanent isolation.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Enabling ARP-Guard Isolate Forwarding

Command	**arp-guard isolate-forwarding enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Enabling ARP-Guard Ratelimit Forwarding

Command	**arp-guard ratelimit-forwarding enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ARP-Guard Monitoring Period

Command	**arp-guard monitor-period** *seconds*
Parameter Description	*seconds*: Indicates the monitoring period in the unit of second. The value ranges from 180 to 86,400.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Maximum Number of ARP-Guard Monitored Hosts

Command	**arp-guard monitored-host-limit** *number*
Parameter Description	*number*: Indicates the maximum number of monitored hosts, ranging from 1 to 4,294,967,295.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ARP-Guard Rate Limit

Command	**arp-guard rate-limit** {**per-src-ip** |**per-src-mac** | **per-port**} *pps*
Parameter Description	**per-src-ip**: Limits the rate of each source IP address. **per-src-mac**: Limits the rate of each source MAC address.

	per-port: Limits the rate of each port. *pps*: Indicates the rate limit, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ARP-Guard Attack Threshold

Command	**arp-guard attack-threshold** {**per-src-ip** | **per-src-mac** | **per-port**} *pps*
Parameter Description	**per-src-ip**: Configures the attack threshold of each source IP address. **per-src-mac**: Configures the attack threshold of each source MAC address. **per-port**: Configures the attack threshold of each port. *pps*: Indicates the attack threshold, ranging from 1 to 19,999. The unit is packets per second (pps).
Command Mode	NFPP configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuring the Global ARP-Guard Scanning Threshold

Command	**arp-guard scan-threshold** *pkt-cnt*
Parameter Description	*pkt-cnt*: Indicates the scanning threshold, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Enabling ARP Guard on an Interface

Command	**nfpp arp-guard enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	ARP guard configured in interface configuration mode takes priority over that configured in NFPP configuration mode.

Configuring the ARP-Guard Isolation Period on an Interface

Command	**nfpp arp-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the ARP-Guard Rate Limit and Attack Threshold on an Interface

Command	**nfpp arp-guard policy** {**per-src-ip** | **per-src-mac** | **per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	**per-src-ip**: Configures the rate limit and attack threshold of each source IP address. **per-src-ip**: Configures the rate limit and attack threshold of each source MAC address. **per-port**: Configures the rate limit and attack threshold of each port. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuring the ARP-Guard Scanning Threshold on an Interface

Command	**nfpp arp-guard scan-threshold** *pkt-cnt*
Parameter Description	*pkt-cnt*: Indicates the scanning threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

CPU Protection Based on ARP Guard

Scenario	● ARP host attacks exist in the system, and some hosts fail to properly establish ARP connection. ● ARP scanning exists in the system, causing a very high CPU utilization rate.
Configuration Steps	● Set the host-based attack threshold to 5 pps. ● Set the ARP scanning threshold to 10 pps. ● Set the isolation period to 180 pps.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#arp-guard rate-limit per-src-mac 5 FS (config-nfpp)#arp-guard attack-threshold per-src-mac 10 FS (config-nfpp)#arp-guard isolate-period 180
Verification	● Run the **show nfpp arp-guard summary** command to display the configuration.
	(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.) Interface Status Isolate-period Rate-limit Attack-threshold Scan-threshold Global Disable 180 4/5/100 8/10/200 15 Maximum count of monitored hosts: 1000 Monitor period: 600s
	● Run the **show nfpp arp-guard hosts** command to display the monitored hosts.
	If col_filter 1 shows '*', it means "hardware do not isolate host".

	VLAN interface IP address MAC address remain-time(s) ---- --------- ---------- ---------- ---------- 1 Gi0/43 5.5.5.16 - 175 Total: 1 host
	● Run the **show nfpp arp-guard scan** command to display the scanned hosts.
	VLAN interface IP address MAC address timestamp ---- --------- ---------- ---------- --------- 1 Gi0/5 - 001a.a9c2.4609 2013-4-30 23:50:32 1 Gi0/5 192.168.206.2 001a.a9c2.4609 2013-4-30 23:50:33 1 Gi0/5 - 001a.a9c2.4609 2013-4-30 23:51:33 1 Gi0/5 192.168.206.2 001a.a9c2.4609 2013-4-30 23:51:34 Total: 4 record(s)

Common Errors

N/A

21.4.2 Configuring IP Guard

Configuration Effect

- IP attacks are identified based on hosts or physical interfaces. In host-based IP attack identification, IP attacks are identified based on the source IP address, VLAN ID, and port. Each type of attack identification has a rate limit and an attack threshold. If the IP packet rate exceeds the rate limit, the packets beyond the rate limit are discarded. If the IP packet rate exceeds the attack threshold, the system prints alarm information and sends traps. In host-based attack identification, the system also isolates the attack source.
- IP guard can also detect IP scanning attacks. IP anti-scanning applies to IP packet attacks as follows: the destination IP address continuously changes but the source IP address remains the same, and the destination IP address is not the IP address of the local device.
- Configure IP guard isolation to assign hardware-isolated entries against host attacks so that attack packets are neither sent to the CPU nor forwarded.
- IP anti-scanning applies to IP packet attacks where the destination IP address is not the local IP address. The CPP limits the rate of IP packets where the destination IP address is the local IP address.

Notes

- For a command that is configured both in NFPP configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in NFPP configuration mode.
- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.

Configuration Steps

Enabling IP Guard

- (Mandatory) IP guard is enabled by default.

- This function can be enabled in NFPP configuration mode or interface configuration mode.
- If IP guard is disabled, the system automatically clears monitored hosts.

Configuring the IP-Guard Isolation Period

- (Optional) IP-guard isolation is disabled by default.
- If the packet traffic of attackers exceeds the rate limit defined in CPP, you can configure the isolation period to discard packets and therefore to save bandwidth resources.
- The isolation period can be configured in NFPP configuration mode or interface configuration mode.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the IP-Guard Monitoring Period

- (Mandatory) The default IP-guard monitoring period is 600 seconds.
- If the IP-guard isolation period is configured, it is directly used as the monitoring period, and the configured monitoring period will lose effect.
- The monitoring period can be configured in NFPP configuration mode.

Configuring the Maximum Number of IP-Guard Monitored Hosts

- (Mandatory) The maximum number of IP-guard monitored hosts is 20,000 by default.
- Set the maximum number of IP-guard monitored hosts reasonably. As the number of monitored hosts increases, more CPU resources are used.
- The maximum number of IP-guard monitored hosts can be configured in NFPP configuration mode.
- If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20,000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted.
- If the table of monitored hosts is full, the system prints the log "% NFPP_IP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the IP-Guard Attack Threshold

- Mandatory.
- The attack threshold can be configured in NFPP configuration mode or interface configuration mode.
- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.
- If the memory cannot be allocated to detected attackers, the system prints the log "%NFPP_IP_GUARD-4-NO_MEMORY: Failed to alloc memory." to notify the administrator.
- Source IP address-based rate limiting takes priority over port-based rate limiting.

Configuring the IP-Guard Scanning Threshold

- Mandatory.
- The scanning threshold can be configured in NFPP configuration mode or interface configuration mode.
- ARP scanning attack may have occurred if ARP packets received within 10 seconds meet the following conditions:
  - The source IP address remains the same.
  - The destination IP address continuously changes and is not the local IP address, and the change times exceed the scanning threshold.

Configuring IP-Guard Trusted Hosts

- (Optional) No IP-guard trusted host is configured by default.
- For IP guard, you can only configure a maximum of 500 IP addresses not to be monitored.
- Trusted hosts can be configured in NFPP configuration mode.
- If any entry matching a trusted host (IP addresses are the same) exists in the table of monitored hosts, the system automatically deletes this entry.
- If the table of trusted hosts is full, the system prints the log "%ERROR: Attempt to exceed limit of 500 trusted hosts." to notify the administrator.
- If a trusted host cannot be deleted, the system prints the log "%ERROR: Failed to delete trusted host 1.1.1.0 255.255.255.0." to notify the administrator.
- If a host cannot be trusted, the system prints the log "%ERROR: Failed to add trusted host 1.1.1.0 255.255.255.0." to notify the administrator.
- If the host to trust already exists, the system prints the log "%ERROR: Trusted host 1.1.1.0 255.255.255.0 has already been configured." to notify the administrator.
- If the host to delete from the trusted table does not exist, the system prints the log "%ERROR: Trusted host 1.1.1.0 255.255.255.0 is not found." to notify the administrator.
- If the memory cannot be allocated to a trusted host, the system prints the log "%ERROR: Failed to alloc memory." to notify the administrator.

Verification

When a host in the network sends IP attack packets to a switch configured with IP guard, check whether these packets can be sent to the CPU.

- If the rate of packets from untrusted hosts exceeds the attack threshold or scanning threshold, an attack log is displayed.
- If an isolated entry is created for the attacker, an isolation log is displayed.

Related Commands

Enabling IP Guard Globally

Command	**ip-guard enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode

Usage Guide	N/A

Configuring the Global IP-Guard Isolation Period

Command	**ip-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. **permanent**: Indicates permanent isolation.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global IP-Guard Monitoring Period

Command	**ip-guard monitor-period** *seconds*
Parameter Description	*seconds*: Indicates the monitoring period in the unit of second. The value ranges from 180 to 86,400.
Command Mode	NFPP configuration mode
Usage Guide	If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the Maximum Number of IP-Guard Monitored Hosts

Command	**ip-guard monitored-host-limit** *number*
Parameter Description	*number*: Indicates the maximum number of monitored hosts, ranging from 1 to 4,294,967,295.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global IP-Guard Rate Limit

Command	**ip-guard rate-limit** {**per-src-ip** | **per-port**} *pps*
Parameter Description	**per-src-ip**: Limits the rate of each source IP address. **per-port**: Limits the rate of each port. *pps*: Indicates the rate limit, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global IP-Guard Attack Threshold

Command	**ip-guard attack-threshold** {**per-src-ip** | **per-port**} *pps*
Parameter Description	**per-src-ip**: Configures the attack threshold of each source IP address. **per-port**: Configures the attack threshold of each port. *pps*: Indicates the attack threshold, ranging from 1 to 19,999. The unit is pps.
Command Mode	NFPP configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuring the Global IP-Guard Scanning Threshold

Command	**ip-guard scan-threshold** *pkt-cnt*
Parameter Description	*pkt-cnt*: Indicates the scanning threshold, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring IP-Guard Trusted Hosts

Command	**ip-guard trusted-host** *ip mask*
Parameter Description	*ip*: Indicates the IP address. *mask*: Indicates the mask of an IP address. **all**: Used with **no** to delete all trusted hosts.
Command Mode	NFPP configuration mode
Usage Guide	If you do not want to monitor a host, you can run this command to trust the host. This trusted host can send IP packets to the CPU, without any rate limiting or alarm reporting.

Enabling IP Guard on an Interface

Command	**nfpp ip-guard enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	IP guard configured in interface configuration mode takes priority over that configured in NFPP configuration mode.

Configuring the IP-Guard Isolation Period on an Interface

Command	**nfpp ip-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the IP-Guard Rate Limit and Attack Threshold on an Interface

Command	**nfpp ip-guard policy** {**per-src-ip | per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	**per-src-ip**: Configures the attack threshold of each source IP address. **per-port**: Configures the attack threshold of each port. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode

Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuring the IP-Guard Scanning Threshold on an Interface

Command	**nfpp ip-guard scan-threshold** *pkt-cnt*
Parameter Description	*pkt-cnt*: Indicates the scanning threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

CPU Protection Based on IP Guard

Scenario	● IP host attacks exist in the system, and packets of some hosts cannot be properly routed and forwarded. ● IP scanning exists in the system, causing a very high CPU utilization rate. ● Packet traffic of some hosts is very large in the system, and these packets need to pass through.
Configuration Steps	● Configure the host-based attack threshold. ● Configure the IP scanning threshold. ● Set the isolation period to a non-zero value. ● Configure trusted hosts.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#ip-guard rate-limit per-src-ip 20 FS (config-nfpp)#ip-guard attack-threshold per-src-ip 30 FS (config-nfpp)#ip-guard isolate-period 180 FS (config-nfpp)#ip-guard trusted-host 192.168.201.46 255.255.255.255
Verification	● Run the **show nfpp ip-guard summary** command to display the configuration.
	(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.) Interface Status Isolate-period Rate-limit Attack-threshold Scan-threshold Global Disable 180 20/-/100 30/-/200 100 Maximum count of monitored hosts: 1000 Monitor period: 600s
	● Run the **show nfpp ip-guard hosts** command to display the monitored hosts.
	If col_filter 1 shows '*', it means "hardware do not isolate host". VLAN interface IP address Reason remain-time(s) ---- --------- ---------- ------ ---------- 1 Gi0/5 192.168.201.47 ATTACK 160

	Total: 1 host
	● Run the **show nfpp ip-guard trusted-host** command to display the trusted hosts.
	IP address mask ---------- ---- 192.168.201.46 255.255.255.255 Total: 1 record(s)

Common Errors

N/A

21.4.3 Configuring ICMP Guard

Configuration Effect

- ICMP attacks are identified based on hosts or ports. In host-based attack identification, ICMP attacks are identified based on the source IP address, VLAN ID, and port. Each type of attack identification has a rate limit and an attack threshold. If the ICMP packet rate exceeds the rate limit, the packets beyond the rate limit are discarded. If the ICMP packet rate exceeds the attack threshold, the system prints alarm information and sends traps. In host-based attack identification, the system also isolates the attack source.
- Configure ICMP guard isolation to assign hardware-isolated entries against host attacks so that attack packets are neither sent to the CPU nor forwarded.

Notes

- For a command that is configured both in NFPP configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in NFPP configuration mode.
- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.

Configuration Steps

Enabling ICMP Guard

- (Mandatory) ICMP guard is enabled by default.
- This function can be enabled in NFPP configuration mode or interface configuration mode.
- If ICMP guard is disabled, the system automatically clears monitored hosts.

Configuring the ICMP-Guard Isolation Period

- (Optional) ICMP-guard isolation is disabled by default.
- If the packet traffic of attackers exceeds the rate limit defined in CPP, you can configure the isolation period to discard packets and therefore to save bandwidth resources.
- The isolation period can be configured in NFPP configuration mode or interface configuration mode.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the ICMP-Guard Monitoring Period

- (Mandatory) The default ICMP-guard monitoring period is 600 seconds.

- If the ICMP-guard isolation period is configured, it is directly used as the monitoring period, and the configured monitoring period will lose effect.
- The monitoring period can be configured in NFPP configuration mode.

Configuring the Maximum Number of ICMP-Guard Monitored Hosts

- (Mandatory) The maximum number of ICMP-guard monitored hosts is 20,000 by default.
- Set the maximum number of ICMP-guard monitored hosts reasonably. As the number of actually monitored hosts increases, more CPU resources are used.
- The maximum number of ICMP-guard monitored hosts can be configured in NFPP configuration mode.
- If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted.
- If the table of monitored hosts is full, the system prints the log "% NFPP_ICMP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the ICMP-Guard Attack Threshold

- Mandatory.
- The attack threshold can be configured in NFPP configuration mode or interface configuration mode.
- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.
- If the memory cannot be allocated to detected attackers, the system prints the log "%NFPP_ ICMP_GUARD-4-NO_MEMORY: Failed to alloc memory." to notify the administrator.
- Source IP address-based rate limiting takes priority over port-based rate limiting.

Configuring ICMP-Guard Trusted Hosts

- (Optional) No ICMP-guard trusted host is configured by default.
- For ICMP guard, you can only configure a maximum of 500 IP addresses not to be monitored.
- Trusted hosts can be configured in NFPP configuration mode.
- If any entry matching a trusted host (IP addresses are the same) exists in the table of monitored hosts, the system automatically deletes this entry.
- If the table of trusted hosts is full, the system prints the log "%ERROR: Attempt to exceed limit of 500 trusted hosts." to notify the administrator.
- If a trusted host cannot be deleted, the system prints the log "%ERROR: Failed to delete trusted host 1.1.1.0 255.255.255.0." to notify the administrator.
- If a host cannot be trusted, the system prints the log "%ERROR: Failed to add trusted host 1.1.1.0 255.255.255.0." to notify the administrator.

- If the host to trust already exists, the system prints the log "%ERROR: Trusted host 1.1.1.0 255.255.255.0 has already been configured." to notify the administrator.
- If the host to delete from the trusted table does not exist, the system prints the log "%ERROR: Trusted host 1.1.1.0 255.255.255.0 is not found." to notify the administrator.
- If the memory cannot be allocated to a trusted host, the system prints the log "%ERROR: Failed to alloc memory." to notify the administrator.

Verification

When a host in the network sends ICMP attack packets to a switch configured with ICMP guard, check whether these packets can be sent to the CPU.

- If the rate of packets from an untrusted host exceeds the attack threshold, an attack log is displayed.
- If an isolated entry is created for the attacker, an isolation log is displayed.

Related Commands

Enabling ICMP Guard Globally

Command	**icmp-guard enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ICMP-Guard Isolation Period

Command	**icmp-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	NFPP configuration mode
Usage Guide	The attacker isolation period falls into two types: global isolation period and port-based isolation period (local isolation period). For a port, if the port-based isolation period is not configured, the global isolation period is used; otherwise, the port-based isolation period is used.

Configuring the Global ICMP-Guard Monitoring Period

Command	**icmp-guard monitor-period** *seconds*
Parameter Description	*seconds*: Indicates the monitoring period in the unit of second. The value ranges from 180 to 86,400.
Command Mode	NFPP configuration mode
Usage Guide	If the isolation period is 0, the system performs software monitoring on detected attackers. The timeout period is the monitoring period. During software monitoring, if the isolation period is set to a non-zero value, the system automatically performs hardware isolation against monitored attackers and sets the timeout period as the monitoring period. The

	monitoring period is valid only when the isolation period is 0. If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the Maximum Number of ICMP-Guard Monitored Hosts

Command	**icmp-guard monitored-host-limit** *number*
Parameter Description	*number*: Indicates the maximum number of monitored hosts, ranging from 1 to 4,294,967,295.
Command Mode	NFPP configuration mode
Usage Guide	If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted. If the table of monitored hosts is full, the system prints the log "% NFPP_ICMP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the Global ICMP-Guard Rate Limit

Command	**icmp-guard rate-limit** {**per-src-ip** | **per-port**} *pps*
Parameter Description	**per-src-ip**: Limits the rate of each source IP address. **per-port**: Limits the rate of each port. *pps*: Indicates the rate limit, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ICMP-Guard Attack Threshold

Command	**icmp-guard attack-threshold** {**per-src-ip** | **per-port**} *pps*
Parameter Description	**per-src-ip**: Configures the attack threshold of each source IP address. **per-port**: Configures the attack threshold of each port. *pps*: Indicates the attack threshold, ranging from 1 to 19,999. The unit is pps.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring ICMP-Guard Trusted Hosts

Command	**icmp-guard trusted-host** *ip mask*
Parameter Description	*ip*: Indicates the IP address. *mask*: Indicates the mask of an IP address. **all**: Used with **no** to delete all trusted hosts.
Command Mode	NFPP configuration mode
Usage Guide	If you do not want to monitor a host, you can run this command to trust the host. This trusted host can send ICMP packets to the CPU, without any rate limiting or alarm reporting. You can configure the mask so that no host in one

	network segment is monitored. You can configure a maximum of 500 trusted hosts.

Enabling ICMP Guard on an Interface

Command	**nfpp icmp-guard enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	ICMP guard configured in interface configuration mode takes priority over that configured in NFPP configuration mode.

Configuring the ICMP-Guard Isolation Period on an Interface

Command	**nfpp icmp-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the ICMP-Guard Rate Limit and Attack Threshold on an Interface

Command	**nfpp icmp-guard policy** {**per-src-ip | per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	**per-src-ip**: Configures the rate limit and attack threshold of each source IP address. **per-port**: Configures the rate limit and attack threshold of each port. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuration Example

CPU Protection Based on ICMP Guard

Scenario	● ICMP host attacks exist in the system, and some hosts cannot successfully ping devices. ● Packet traffic of some hosts is very large in the system, and these packets need to pass through.
Configuration Steps	● Configure the host-based attack threshold. ● Set the isolation period to a non-zero value. ● Configure trusted hosts.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#icmp-guard rate-limit per-src-ip 20 FS (config-nfpp)#icmp-guard attack-threshold per-src-ip 30

	FS (config-nfpp)#icmp-guard isolate-period 180 FS (config-nfpp)#icmp-guard trusted-host 192.168.201.46 255.255.255.255
Verification	● Run the **show nfpp icmp-guard summary** command to display the configuration.
	(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.) Interface Status Isolate-period Rate-limit Attack-threshold Global Disable 180 20/-/400 30/-/400 Maximum count of monitored hosts: 1000 Monitor period: 600s
	● Run the **show nfpp icmp-guard hosts** command to display the monitored hosts.
	If col_filter 1 shows '*', it means "hardware do not isolate host". VLAN interface IP address remain-time(s) ---- --------- ---------- ---------- 1 Gi0/5 192.168.201.47 160 Total: 1 host
	● Run the **show nfpp icmp-guard trusted-host** command to display the trusted hosts.
	IP address mask ---------- ---- 192.168.201.46 255.255.255.255 Total: 1 record(s)

Common Errors

N/A

21.4.4 Configuring DHCP Guard

Configuration Effect

- DHCP attacks are identified based on hosts or ports. In host-based attack identification, DHCP attacks are identified based on the link-layer source IP address, VLAN ID, and port. Each type of attack identification has a rate limit and an attack threshold. If the DHCP packet rate exceeds the rate limit, the packets beyond the rate limit are discarded. If the DHCP packet rate exceeds the attack threshold, the system prints alarm information and sends traps. In host-based attack identification, the system also isolates the attack source.
- Configure DHCP guard isolation to assign hardware-isolated entries against host attacks so that attack packets are neither sent to the CPU nor forwarded.

Notes

- For a command that is configured both in NFPP configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in NFPP configuration mode.

- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.
- For trusted ports configured for DHCP snooping, DHCP guard does not take effect, preventing false positive of DHCP traffic on the trusted ports. For details about trusted ports of DHCP snooping, see "Configuring Basic Functions of DHCP Snooping" in the Configuring DHCP Snooping.

Configuration Steps

Enabling DHCP Guard

- (Mandatory) DHCP guard is enabled by default.
- This function can be enabled in NFPP configuration mode or interface configuration mode.
- If DHCP guard is disabled, the system automatically clears monitored hosts.

Configuring the DHCP-Guard Isolation Period

- (Optional) DHCP-guard isolation is disabled by default.
- If the packet traffic of attackers exceeds the rate limit defined in CPP, you can configure the isolation period to discard packets and therefore to save bandwidth resources.
- The isolation period can be configured in NFPP configuration mode or interface configuration mode.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the DHCP-Guard Monitoring Period

- (Mandatory) DHCP-guard monitoring is enabled by default.
- If the DHCP-guard isolation period is configured, it is directly used as the monitoring period, and the configured monitoring period will lose effect.
- The monitoring period can be configured in NFPP configuration mode.

Configuring the Maximum Number of DHCP-Guard Monitored Hosts

- (Mandatory) The maximum number of DHCP-guard monitored hosts is 20,000 by default.
- Set the maximum number of DHCP-guard monitored hosts reasonably. As the number of monitored hosts increases, more CPU resources are used.
- The maximum number of DHCP-guard monitored hosts can be configured in NFPP configuration mode.
- If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted.
- If the table of monitored hosts is full, the system prints the log "% NFPP_DHCP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the DHCP-Guard Attack Threshold

- Mandatory.
- The attack threshold can be configured in NFPP configuration mode or interface configuration mode.

- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.
- If the memory cannot be allocated to detected attackers, the system prints the log "%NFPP_DHCP_GUARD-4-NO_MEMORY: Failed to alloc memory." to notify the administrator.
- Source MAC address-based rate limiting takes priority over port-based rate limiting.

Verification

When a host in the network sends DHCP attack packets to a switch configured with DHCP guard, check whether these packets can be sent to the CPU.

- If the parameter of the packets exceeds the attack threshold, an attack log is displayed.
- If an isolated entry is created for the attacker, an isolation log is displayed.

Related Commands

Enabling DHCP Guard Globally

Command	**dhcp-guard enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global DHCP-Guard Isolation Period

Command	**dhcp-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	NFPP configuration mode
Usage Guide	The attacker isolation period falls into two types: global isolation period and port-based isolation period (local isolation period). For a port, if the port-based isolation period is not configured, the global isolation period is used; otherwise, the port-based isolation period is used.

Configuring the Global DHCP-Guard Monitoring Period

Command	**dhcp-guard monitor-period** *seconds*
Parameter Description	*seconds*: Indicates the monitoring period in the unit of second. The value ranges from 180 to 86,400.
Command Mode	NFPP configuration mode
Usage Guide	If the isolation period is 0, the system performs software monitoring on detected attackers. The timeout period is the monitoring period. During software monitoring, if the isolation period is set to a non-zero value, the system automatically

	performs hardware isolation against monitored attackers and sets the timeout period as the monitoring period. The monitoring period is valid only when the isolation period is 0. If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the Maximum Number of DHCP-Guard Monitored Hosts

Command	**dhcp-guard monitored-host-limit** *number*
Parameter Description	*number*: Indicates the maximum number of monitored hosts, ranging from 1 to 4,294,967,295.
Command Mode	NFPP configuration mode
Usage Guide	If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted. If the table of monitored hosts is full, the system prints the log "% NFPP_DHCP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the Global DHCP-Guard Rate Limit

Command	**dhcp-guard rate-limit** {**per-src-mac** | **per-port**} *pps*
Parameter Description	**per-src-mac**: Limits the rate of each source MAC address. **per-port**: Limits the rate of each port. *pps*: Indicates the rate limit, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global DHCP-Guard Attack Threshold

Command	**dhcp-guard attack-threshold** {**per-src-mac** | **per-port**} *pps*
Parameter Description	**per-src-mac**: Configures the attack threshold of each source MAC address. **per-port**: Configures the attack threshold of each port. *pps*: Indicates the attack threshold, ranging from 1 to 19,999. The unit is pps.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Enabling DHCP Guard on an Interface

Command	**nfpp dhcp-guard enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	DHCP guard configured in interface configuration mode takes priority over that configured in NFPP configuration mode.

Configuring the DHCP-Guard Isolation Period on an Interface

Command	**nfpp dhcp-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the DHCP-Guard Rate Limit and Attack Threshold on an Interface

Command	**nfpp dhcp-guard policy** {**per-src-mac | per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	**per-src-ip**: Configures the rate limit and attack threshold of each source IP address. **per-port**: Configures the rate limit and attack threshold of each port. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuration Example

CPU Protection Based on DHCP Guard

Scenario	● DHCP host attacks exist in the system, and some hosts fail to request IP addresses.
Configuration Steps	● Configure the host-based attack threshold. ● Set the isolation period to a non-zero value.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#dhcp-guard rate-limit per-src-mac 8 FS (config-nfpp)#dhcp-guard attack-threshold per-src-mac 16 FS (config-nfpp)#dhcp-guard isolate-period 180
Verification	● Run the **show nfpp dhcp-guard summary** command to display the configuration.
	(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.) Interface Status Isolate-period Rate-limit Attack-threshold Global Disable 180 -/8/150 -/16/300 Maximum count of monitored hosts: 1000 Monitor period: 600s
	● Run the **show nfpp dhcp-guard hosts** command to display the monitored hosts.
	If col_filter 1 shows '*', it means "hardware do not isolate host". VLAN interface MAC address remain-time(s)

	---- --------- ---------- ---------- *1 Gi0/5 001a.a9c2.4609 160 Total: 1 host

Common Errors

N/A

21.4.5 Configuring DHCPv6 Guard

Configuration Effect

- DHCPv6 attacks are identified based on hosts or ports. In host-based attack identification, DHCPv6 attacks are identified based on the link-layer source IP address, VLAN ID, and port. Each type of attack identification has a rate limit and an attack threshold. If the DHCPv6 packet rate exceeds the rate limit, the packets beyond the rate limit are discarded. If the DHCPv6 packet rate exceeds the attack threshold, the system prints alarm information and sends traps.
- In host-based attack identification, the system also isolates the attack source.

Notes

- For a command that is configured both in NFPP configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in NFPP configuration mode.
- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.
- For trusted ports configured for DHCPv6 snooping, DHCPv6 guard does not take effect, preventing false positive of DHCPv6 traffic on the trusted ports. For details about trusted ports of DHCPv6 snooping, see "Configuring Basic Functions of DHCPv6 Snooping" in the Configuring DHCPv6 Snooping.

Configuration Steps

↘ Enabling DHCPv6 Guard

- (Mandatory) DHCPv6 guard is enabled by default.
- DHCPv6 guard can be enabled in NFPP configuration mode or interface configuration mode.
- If DHCPv6 guard is disabled, the system automatically clears monitored hosts.

↘ Configuring the DHCPv6-Guard Isolation Period

- (Optional) DHCPv6-guard isolation is disabled by default.
- If the packet traffic of attackers exceeds the rate limit defined in CPP, you can configure the isolation period to discard packets and therefore to save bandwidth resources.
- The DHCPv6-guard isolation period can be configured in NFPP configuration mode or interface configuration mode.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

↘ Configuring the DHCPv6-Guard Monitoring Period

- (Mandatory) The default DHCPv6-guard monitoring period is 600 seconds.
- If the DHCPv6-guard isolation period is configured, it is directly used as the monitoring period, and the configured monitoring

period does not take effect.

- The DHCPv6-guard monitoring period can be configured in NFPP configuration mode.

Configuring the Maximum Number of DHCPv6-Guard Monitored Hosts

- (Mandatory) The maximum number of DHCPv6-guard monitored hosts is 20,000 by default.
- Set the maximum number of DHCPv6-guard monitored hosts reasonably. As the number of monitored hosts increases, more CPU resources are used.
- The maximum number of DHCPv6-guard monitored hosts can be configured in NFPP configuration mode.
- If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted.
- If the table of monitored hosts is full, the system prints the log "% NFPP_DHCPV6_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the DHCPv6-Guard Attack Threshold

- Mandatory.
- The DHCPv6-guard attack threshold can be configured in NFPP configuration mode or interface configuration mode.
- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.
- If the memory cannot be allocated to detected attackers, the system prints the log "%NFPP_DHCPV6_GUARD-4-NO_MEMORY: Failed to alloc memory." to notify the administrator.
- Source MAC address-based rate limiting takes priority over port-based rate limiting.

Verification

When a host in the network sends DHCPv6 attack packets to a switch configured with DHCPv6 guard, check whether these packets can be sent to the CPU.

- If the parameter of the packets exceeds the attack threshold, an attack log is displayed.
- If an isolated entry is created for the attacker, an isolation log is displayed.

Related Commands

Enabling DHCPv6 Guard Globally

Command	**dhcpv6-guard enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global DHCPv6-Guard Isolation Period

Command	**dhcpv6-guardisolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	NFPP configuration mode
Usage Guide	The attacker isolation period falls into two types: global isolation period and port-based isolation period (local isolation period). For a port, if the port-based isolation period is not configured, the global isolation period is used; otherwise, the port-based isolation period is used.

Configuring the Global DHCPv6-Guard Monitoring Period

Command	**dhcpv6-guard monitor-period** *seconds*
Parameter Description	*seconds*: Indicates the monitoring period in the unit of second. The value ranges from 180 to 86,400.
Command Mode	NFPP configuration mode
Usage Guide	If the isolation period is 0, the system performs software monitoring on detected attackers. The timeout period is the monitoring period. During software monitoring, if the isolation period is set to a non-zero value, the system automatically performs hardware isolation against monitored attackers and sets the timeout period as the monitoring period. The monitoring period is valid only when the isolation period is 0. If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the Maximum Number of DHCPv6-Guard Monitored Hosts

Command	**dhcpv6-guard monitored-host-limit** *number*
Parameter Description	*number*: Indicates the maximum number of monitored hosts, ranging from 1 to 4,294,967,295.
Command Mode	NFPP configuration mode
Usage Guide	If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted. If the table of monitored hosts is full, the system prints the log "% NFPP_DHCPV6_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 monitored hosts." to notify the administrator.

Configuring the Global DHCPv6-Guard Rate Limit

Command	**dhcpv6-guardrate-limit** { **per-src-mac | per-port**} *pps*
Parameter Description	**per-src-mac**: Limits the rate of each source MAC address. **per-port**: Limits the rate of each port. *pps*: Indicates the rate limit, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode

Usage Guide	N/A

Configuring the Global DHCPv6-Guard Attack Threshold

Command	**dhcpv6-guard attack-threshold** { **per-src-mac** | **per-port**} *pps*
Parameter Description	**per-src-mac**: Configures the attack threshold of each source MAC address. **per-port**: Configures the attack threshold of each port. *pps*: Indicates the attack threshold, ranging from 1 to 19,999. The unit is pps.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Enabling DHCPv6 Guard on an Interface

Command	**nfpp dhcpv6-guard enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	DHCPv6 guard configured in interface configuration mode takes priority over that configured in NFPP configuration mode.

Configuring the DHCPv6-Guard Isolation Period on an Interface

Command	**nfpp dhcpv6-guard isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the DHCP-Guard Rate Limit and Attack Threshold on an Interface

Command	**nfpp dhcpv6-guard policy** {**per-src-mac** | **per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	**per-src-ip**: Configures the rate limit and attack threshold of each source IP address. **per-port**: Configures the rate limit and attack threshold of each port. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuration Example

CPU Protection Based on DHCPv6 Guard

Scenario	• DHCPv6 host attacks exist in the system, and DHCPv6 neighbor discovery fails on some hosts.
Configuration	• Configure the host-based attack threshold.

Steps	● Set the isolation period to a non-zero value.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#dhcpv6-guard rate-limit per-src-mac 8 FS (config-nfpp)#dhcpv6-guard attack-threshold per-src-mac 16 FS (config-nfpp)#dhcpv6-guard isolate-period 180
Verification	● Run the **show nfpp dhcpv6-guard summary** command to display the configuration.
	(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.) Interface Status Isolate-period Rate-limit Attack-threshold Global Disable 180 -/8/150 -/16/300 Maximum count of monitored hosts: 1000 Monitor period: 600s
	● Run the **show nfpp dhcpv6-guard hosts** command to display the monitored hosts.
	If col_filter 1 shows '*', it means "hardware do not isolate host". VLAN interface MAC address remain-time(s) ---- --------- ---------- ---------- *1 Gi0/5 001a.a9c2.4609 160 Total: 1 host

Common Errors

N/A

21.4.6 Configuring ND Guard

Configuration Effect

- AR ND guard classifies ND packets into three types based on their purposes: 1. NS and NA; 2. RS; 3. RA and Redirect. Type 1 packets are used for address resolution. Type 2 packets are used by hosts to discover the gateway. Type 3 packets are related to routing: RAs are used to advertise the gateway and prefix while Redirect packets are used to advertise a better next hop.
- At present, only port-based ND packet attack identification is supported. You can configure the rate limits and attack thresholds for these three types of packets respectively. If the ND packet rate exceeds the rate limit, the packets beyond the rate limit are discarded. If the ND packet rate exceeds the attack threshold, the system prints logs and sends traps.

Notes

- For a command that is configured both in NFPP configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in NFPP configuration mode.
- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.

Configuration Steps

Enabling ND Guard

- (Mandatory) ND guard is enabled by default.
- This function can be enabled in NFPP configuration mode or interface configuration mode.

Enabling ND-Guard Ratelimit Forwarding

- (Optional) This function is enabled by default.
- If the port-based isolation entry takes effect, you can enable this function to pass some of the packets while not discarding all of them.
- This function can be enabled in NFPP configuration mode.

Configuring the ND-Guard Attack Threshold

- Mandatory.
- The ND-guard attack threshold can be enabled in NFPP configuration mode or interface configuration mode.
- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.
- If memories cannot assigned to detected attackers, the system prints the log "%NFPP_ND_GUARD-4-NO_MEMORY: Failed to alloc memory." to notify the administrator.

Verification

When a host in the network sends ND attack packets to a switch configured with ND guard, check whether these packets can be sent to the CPU.

- If the parameter of the packets exceeds the attack threshold, an attack log is displayed.

Related Commands

Enabling ND Guard Globally

Command	**nd-guard enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Enabling ND-Guard Ratelimit Forwarding

Command	**nd-guard ratelimit-forwarding enable**
Parameter Description	N/A
Command	NFPP configuration mode

Mode	
Usage Guide	N/A

Configuring the Global ND-Guard Rate Limit

Command	**nd-guard rate-limit per-port** [**ns-na** | **rs** | **ra-redirect**] *pps*
Parameter Description	**ns-na**: Indicates NSs and NAs. **rs**: Indicates RSs. **ra-redirect**: Indicates RAs and Redirect packets. *pps*: Indicates the rate limit, ranging from 1 to 19,999.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Global ND-Guard Attack Threshold

Command	**nd-guard attack-threshold per-port**[**ns-na** | **rs** | **ra-redirect**] *pps*
Parameter Description	**ns-na**: Indicates NSs and NAs. **rs**: Indicates RSs. **ra-redirect**: Indicates RAs and Redirect packets. *pps*: Indicates the attack threshold, ranging from 1 to 19,999. The unit is pps.
Command Mode	NFPP configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Enabling ND Guard on an Interface

Command	**nfpp nd-guard enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	ND guard configured in interface configuration mode takes priority over that configured in NFPP configuration mode.

Configuring the ND-Guard Rate Limit and Attack Threshold on an Interface

Command	**nfpp nd-guard policy per-port** [**ns-na** | **rs** | **ra-redirect**] *rate-limit-pps attack-threshold-pps*
Parameter Description	**ns-na**: Indicates NSs and NAs. **rs**: Indicates RSs. **ra-redirect**: Indicates RAs and Redirect packets. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuration Example

CPU Protection Based on ND Guard

Scenario	● ND host attacks exist in the system, and neighbor discovery fails on some hosts.
Configuration Steps	● Configure the host-based attack threshold.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)# nd-guard rate-limit per-port ns-na 30 FS (config-nfpp)# nd-guard attack-threshold per-port ns-na 50
Verification	● Run the **show nfpp nd-guard summary** command to display the configuration.
	(Format of column Rate-limit and Attack-threshold is NS-NA/RS/RA-REDIRECT.) Interface Status Rate-limit Attack-threshold Global Disable 30/15/15

Common Errors

N/A

21.4.7 Configuring a Self-Defined Guard

Configuration Effect

- Configure a self-defined guard to resolve network attack problems in special scenarios.

Notes

- For a command that is configured both in self-defined guard configuration mode and interface configuration mode, the configuration in interface configuration mode takes priority over that configured in self-defined guard configuration mode.
- Isolation is disabled by default. If isolation is enabled, attackers will occupy hardware entries of the security module.
- A self-defined guard takes priority over basic guards. When configuring the match fields of self-defined guards, see the Configuration Guide.

Configuration Steps

Configuring the Guard Name

- (Mandatory) Configure the name of a self-defined guard to create the self-defined guard.
- The guard name must be unique, and the match fields and values c must be different from those of ARP, ICMP, DHCP, IP, and DHCPv6 guards. If the parameters you want to configure already exist, a message is displayed to indicate the configuration failure.

Configuring the Match Fields

- Mandatory.
- Self-defined packets are classified based on the following fields: etype (Ethernet link-layer type), smac (source MAC address), dmac (destination MAC address), protocol (IPv4/IPv6 protocol number), sip (source IPv4/IPv6 address), dip (destination IPv4/IPv6 address), sport (source transport-layer port), and dport (destination transport-layer port).

- **protocol** is valid only when the value of **etype** is **ipv4** or **ipv6**. **src-ip** and **dst-ip** are valid only when the value of **etype** is **ipv4**. **src-ipv6** and **dst-ipv6** are valid only when the value of **etype** is **ipv6**. **src-port** and **dst-port** are valid only when the value of **protocol** is **tcp** or **udp**.
- If the **match** fields and values of a self-defined guard are totally the same as those of an existing guard, the system prints the log "%ERROR: the match type and value are the same with define name (name of an existing guard)." to notify the administrator of the configuration failure.
- If **protocol** is configured but **etype** is IPv4 or IPv6 in the **match** policy, the system prints the log "%ERROR: protocol is valid only when etype is IPv4(0x0800) or IPv6(0x86dd)."
- If **src-ip** and **dst-ip** are configured but **etype** is not IPv4 in the **match** policy, the system prints the log "%ERROR: IP address is valid only when etype is IPv4(0x0800)."
- If **src-ipv6** and **dst-ipv6** are configured but **etype** is not IPv6 in the **match** policy, the system prints the log "%ERROR: IPv6 address is valid only when etype is IPv6(0x86dd)."
- If **src-port** and **dst-port** are configured but **protocol** is not TCP or UDP in the **match** policy, the system prints the log "%ERROR: Port is valid only when protocol is TCP(6) or UDP(17)."
- The following table lists guard policies corresponding to some common network protocols. The rate limits and attack thresholds listed below can meet the requirements in most network scenarios and are for reference only. You can configure valid rate limits and attack thresholds based on actual scenarios.

Protocol	match	policy per-src-ip	policy per-src-mac	policy per-port
RIP	etype 0x0800 protocol 17 dst-port 520	rate-limit 100 attatch-threshold 150	Not applicable to this policy	rate-limit 300 attatch-threshold 500
RIPng	etype 0x86dd protocol 17 dst-port 521	rate-limit 100 attatch-threshold 150	Not applicable to this policy	rate-limit 300 attatch-threshold 500
BPDU	dst-mac 0180.c200.0000	Not applicable to this policy	rate-limit 20 attatch-threshold 40	rate-limit 100 attatch-threshold 100
RERP	dst-mac 01d0.f800.0001	Not applicable to this policy	rate-limit 20 attatch-threshold 40	rate-limit 100 attatch-threshold 100
REUP	dst-mac 01d0.f800.0007	Not applicable to this policy	rate-limit 20 attatch-threshold 40	rate-limit 100 attatch-threshold 100
OSPFv2	etype 0x0800 protocol 89	rate-limit 800 attatch-threshold 1200	Not applicable to this policy	rate-limit 2000 attatch-threshold 3000
OSPFv3	etype 0x86dd protocol 89	rate-limit 800 attatch-threshold 1200	Not applicable to this policy	rate-limit 2000 attatch-threshold 3000
VRRP	etype 0x0800 protocol 112	rate-limit 64 attatch-threshold 100	Not applicable to this policy	rate-limit 1024 attatch-threshold 1024
IPv6 VRRP	etype 0x86dd protocol 112	rate-limit 64 attatch-threshold 100	Not applicable to this policy	rate-limit 1024 attatch-threshold 1024
SNMP	etype 0x0800 protocol 17 dst-port 161	rate-limit 1000 attatch-threshold 1200	Not applicable to this policy	rate-limit 2000 attatch-threshold 3000

Protocol	match	policy per-src-ip	policy per-src-mac	policy per-port
RSVP	etype 0x0800 protocol 46	rate-limit 800 attatch-threshold 1200	Not applicable to this policy	rate-limit 1200 attatch-threshold 1500
LDP (UDP hello)	etype 0x0800 protocol 17 dst-port 646	rate-limit 10 attatch-threshold 15	Not applicable to this policy	rate-limit 100 attatch-threshold 150

- To contain as many existing protocol types as possible and facilitate expansion of new protocol types, self-defined guards allow hosts to freely combine type fields of packets. If the configuration is inappropriate, the network may become abnormal. Therefore, the network administrator needs to have a good knowledge of network protocols. As a reference, the following table lists valid configurations of currently known protocols for common self-defined guard policies. For other protocols not listed in the table, configure them with caution.

Configuring the Global Rate Limit and Attack Threshold

- (Mandatory) If these parameters are not configured, the self-defined guard cannot be enabled.
- You must configure one of the per-src-ip, per-src-mac, and per-port fields. Otherwise, the policy cannot take effect.
- per-src-ip is valid only when etype is IPv4 or IPv6.
- The rate limit configured based on the source MAC address, VLAN ID, and port takes priority over that configured based on the source IP address, VLAN ID, and port.
- The port-based host identification policy of a self-defined guard must be consistent with the global port-based host identification policy.
- If the **per-src-ip** policy is not configured globally but configured for a port, the system prints the log "%ERROR: name (name of a self-defined guard) has not per-src-ip policy." to notify the administrator of the configuration failure.
- If the **per-src-mac** policy is not configured globally but configured for a port, the system prints the log "%ERROR: name (name of a self-defined guard) has not per-src-mac policy." to notify the administrator of the configuration failure.
- If the memory cannot be allocated to detected attackers, the system prints the log "%NFPP_DEFINE_GUARD-4-NO_MEMORY: Failed to allocate memory." to notify the administrator.
- If the configured rate limit is greater than the attack threshold, the system prints the log "%ERROR: rate limit is higher than attack threshold 500pps." to notify the administrator.
- If the configured attack threshold is less than the rate limit, the system prints the log "%ERROR: attack threshold is smaller than rate limit 300pps." to notify the administrator.

Configuring the Global Isolation Period

- (Optional) Isolation is disabled by default.
- If the packet traffic of attackers exceeds the rate limit defined in CPP, you can configure the isolation period to discard packets and therefore to save bandwidth resources.
- The isolation period can be configured in self-defined guard configuration mode or interface configuration mode.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the Global Monitoring Period

- (Mandatory) The default monitoring period is 600 seconds.

- If the isolation period is configured, it is directly used as the monitoring period, and the configured monitoring period will lose effect.
- The monitoring period can be configured in self-defined guard configuration mode.
- If the isolation period is 0, the system performs software monitoring on detected attackers. The timeout period is the monitoring period. During software monitoring, if the isolation period is set to a non-zero value, the system automatically performs hardware isolation against monitored attackers and sets the timeout period as the monitoring period. The monitoring period is valid only when the isolation period is 0.
- If the isolation period is changed to 0, attackers under the corresponding port is deleted, instead of being monitored.

Configuring the Maximum Number of Monitored Hosts

- (Mandatory) The maximum number of monitored hosts is 20,000 by default.
- Set the maximum number of monitored hosts reasonably. As the number of monitored hosts increases, more CPU resources are used.
- The maximum number of monitored hosts can be configured in self-defined guard configuration mode.
- If the number of monitored hosts reaches 20,000 (default value) and the administrator sets the maximum number lower than 20,000, the system does not delete monitored hosts but prints the log "%ERROR: The value that you configured is smaller than current monitored hosts 20000, please clear a part of monitored hosts." This information notifies the administrator that the configuration does not take effect and that some monitored hosts need to be deleted.
- If the table of monitored hosts is full, the system prints the log "% NFPP_DEFINE-4-SESSION_LIMIT: Attempt to exceed limit of name's 20000 monitored hosts." to notify the administrator.

Configuring Trusted Hosts

- (Optional) No trusted host is configured by default.
- You can configure a maximum of 500 trusted IP address or MAC address for a self-defined guard.
- Trusted hosts can be configured in self-defined guard configuration mode.
- If you do not want to monitor a host, you can run the following commands to trust the host. This trusted host can send ICMP packets to the CPU, without any rate limiting or alarm reporting. You can configure the mask so that no host in one network segment is monitored.
- You must configure the **match** type before configuring trusted hosts. If the packet type is IPv4 in the **match** policy, you are not allowed to configure trusted IPv6 addresses. If the packet type is IPv6 in the match policy, you are not allowed to configure trusted IPv4 addresses.
- If the **match** type is not configured, the system prints the log "%ERROR: Please configure match rule first."
- If a trusted IPv4 host is added but **etype** is not IPv4 in the **match** policy, the system prints the log "%ERROR: Match type can't support IPv4 trusted host."
- If a trusted IPv6 host is added but **etype** is not IPv6 in the **match** policy, the system prints the log "%ERROR: Match type can't support IPv6 trusted host."
- If the table of trusted hosts is full, the system prints the log "%ERROR: Attempt to exceed limit of 500 trusted hosts." to notify the administrator.
- If any entry matching a trusted host (IP addresses are the same) exists in the table of monitored hosts, the system automatically

deletes this entry.

- If a trusted host cannot be deleted, the system prints the log "%ERROR: Failed to delete trusted host 1.1.1.0 255.255.255.0." to notify the administrator.
- If a host cannot be trusted, the system prints the log "%ERROR: Failed to add trusted host 1.1.1.0 255.255.255.0." to notify the administrator.
- If the host to trust already exists, the system prints the log "%ERROR: Trusted host 1.1.1.0 255.255.255.0 has already been configured." to notify the administrator.
- If the host to delete from the trusted table does not exist, the system prints the log "%ERROR: Trusted host 1.1.1.0 255.255.255.0 is not found." to notify the administrator.
- If the memory cannot be allocated to a trusted host, the system prints the log "%ERROR: Failed to allocate memory." to notify the administrator.

Enabling a Self-Defined Guard

- Mandatory.
- You have to configure at least one policy between host-based self-defined guard policy and port-based self-defined guard policy. Otherwise, the self-defined guard cannot be enabled.
- If a self-defined guard is disabled, the system automatically clears monitored hosts.
- Self-defined guards can be configured in self-defined guard configuration mode or interface configuration mode.
- If a self-defined guard policy is not completely configured, the self-defined guard cannot be enabled and a prompt is displayed to notify hosts of the missing policy configurations.
- If the name of a self-defined guard does not exist, the system prints the log "%ERROR: The name is not exist."
- If the match type is not configured for a self-defined guard, the system prints the log "%ERROR: name (name of the self-defined guard) doesn't match any type."
- If no policy is configured for a self-defined guard, the system prints the log "%ERROR: name (name of the self-defined guard) doesn't specify any policy."

Verification

When a host in the network sends packets to a switch configured with a self-defined NFPP guard, check whether these packets can be sent to the CPU.

- If the rate of packets from an untrusted host exceeds the attack threshold, an attack log is displayed.
- If an isolated entry is created for the attacker, an isolation log is displayed.

Related Commands

Configuring the Name of a Self-defined Guard

Command	**define** *name*
Parameter Description	**name**: Indicates the name of a self-defined guard.
Command Mode	NFPP configuration mode

Usage Guide	N/A

Configuring Match Fields of a Self-defined Guard

Command	**match** [**etype***type*] [**src-mac***smac* [**src-mac-mask***smac_mask*]] [**dst-mac***dmac* [**dst-mac-mask***dst_mask*]] [**protocol***protocol*] [**src-ip***sip* [**src-ip-mask***sip-mask*]] [**src-ipv6***sipv6* [**src-ipv6-masklen***sipv6-masklen*]] [**dst-ip***dip*[**dst-ip-mask***dip-mask*]] **[dst-ipv6***dipv6* [**dst-ipv6-masklen***dipv6-masklen*]][**src-port***sport*] [**dst-port** *dport*]
Parameter Description	*type*: Indicates the type of Ethernet link-layer packets. *smac*: Indicates the source MAC address. *smac_mask*: Indicates the mask of the source MAC address. *dmac*: Indicates the destination MAC address. *dst_mask*: Indicates the mask of the destination MAC address. *protocol*: Indicates the protocol number of IPv4/IPv6 packets. *sip*: Indicates the source IPv4 address. *sip-mask*: Indicates the mask of the source IPv4 address. *sipv6*: Indicates the source IPv6 address. *sipv6-masklen*: Indicates the mask length of the source IPv6 address. *dip*: Indicates the destination IPv4 address. *dip-mask*: Indicates the mask of the destination IPv4 address. *dipv6*: Indicates the destination IPv6 address. *dipv6-masklen*: Indicates the mask length of the destination IPv6 address. *sport*: Indicates the ID of the source transport-layer port. *dsport*: Indicates the ID of the destination transport-layer port.
Command Mode	Self-defined guard configuration mode
Usage Guide	Create a new self-defined guard and specify the packet fields matched by this guard.

Configuring the Global Rate Limit and Attack Threshold of a Self-defined Guard

Command	**global-policy** {**per-src-ip** | **per-src-mac** | **per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	**per-src-ip**: Collects rate statistics for host identification based on the source IP address, VLAN ID, and port. **per-src-mac**: Collects rate statistics for host identification based on the source MAC address, VLAN ID, and port. **per-port**: Collects rate statistics based on each packet receiving port. *rate-limit-pps*: Indicates the rate limit. *attack-threshold-pps*: Indicates the attack threshold.
Command Mode	Self-defined guard configuration mode
Usage Guide	Before creating a self-defined guard type, you must specify rate statistic classification rules for this type, namely, source IP address-based host identification, source MAC address-based host identification, host-based self-defined packet rate statistics, or port-based rate statistics, and specify the rate limits and attack thresholds for the specified rules.

Configuring the Global Isolation Period of a Self-defined Guard

Command	**isolate-period** [*seconds* | **permanent**]
Parameter Description	*seconds*: Indicates the isolation period in the unit of second. It can be set to 0 or any value from 30 to 86,400. The value 0 indicates no isolation. **permanent**: Indicates permanent isolation.

Command Mode	Self-defined guard configuration mode
Usage Guide	If the isolation period is not 0, a host is isolated and its packets of the self-defined guard type are discarded when the packet rate of the self-defined guard exceeds the attack threshold.

Configuring the Global Monitoring Period of a Self-defined Guard

Command	**monitor-period** *seconds*
Parameter Description	*seconds*: Indicates the monitoring period in the unit of second. The value ranges from 180 to 86,400.
Command Mode	Self-defined guard configuration mode
Usage Guide	N/A

Configuring the Maximum Number of Monitored Hosts of a Self-defined Guard

Command	**monitored-host-limit** *number*
Parameter Description	*number*: Indicates the maximum number of monitored hosts, ranging from 1 to 4,294,967,295.
Command Mode	Self-defined guard configuration mode
Usage Guide	N/A

Configuring Trusted Hosts of a Self-defined Guard

Command	**trusted-host** {*mac mac_mask* | *ip mask* | *IPv6/prefixlen*}
Parameter Description	*mac*: Indicates the MAC address. *mac_mask*: Indicates the mask of an MAC address. *ip*: Indicates the IP address. *mask*: Indicates the mask of an IP address. *IPv6/prefixlen*: Indicates the IPv6 address and its mask length. **all**: Used with **no** to delete all trusted hosts.
Command Mode	Self-defined guard configuration mode
Usage Guide	N/A

Enabling a Self-Defined Guard Globally

Command	**define** *name* **enable**
Parameter Description	*name*: Indicates the name of a self-defined guard.
Command Mode	NFPP configuration mode
Usage Guide	The configuration takes effect only after you have configured **match**, **rate-count**, **rate-limit**, and **attack-threshold**. Otherwise, the configuration fails.

Enabling a Self-defined Guard on an Interface

Command	**nfpp define** *name* **enable**
Parameter Description	*name*: Indicates the name of a self-defined guard.
Command Mode	Interface configuration mode
Usage Guide	The self-defined name must exist. The configuration takes effect only after you have configured **match**, **rate-count**, **rate-limit**, and **attack-threshold**. Otherwise, the configuration fails.

Configuring the Rate Limit and Attack Threshold of a Self-defined Guard on an Interface

Command	**nfpp define** *name* **policy** {**per-src-ip** | **per-src-mac**| **per-port**} *rate-limit-pps attack-threshold-pps*
Parameter Description	*name*: Indicates the name of a self-defined guard. **per-src-ip**: Configures the rate limit and attack threshold of each source IP address. **per-src-mac**: Configures the rate limit and attack threshold of each source MAC address. **per-port**: Configures the rate limit and attack threshold of each port. *rate-limit-pps*: Indicates the rate limit, ranging from 1 to 19,999. *attack-threshold-pps*: Indicates the attack threshold, ranging from 1 to 19,999.
Command Mode	Interface configuration mode
Usage Guide	The attack threshold must be equal to or greater than the rate limit.

Configuration Example

CPU Protection Based on a Self-Defined Guard

Scenario	● Basic guards cannot protect the system with RIP attacks.
Configuration Steps	● Configure a self-defined guard, with the key fields matching RIP packets. ● Configure the rate limit. ● Configure the isolation period. ● Configure trusted hosts.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#define rip FS (config-nfpp-define)#match etype 0x0800 protocol 17 dst-port 520 FS (config-nfpp-define)#global-policy per-src-ip 100 150 FS (config-nfpp-define)# isolate-period 180 FS (config-nfpp-define)#trusted-host 192.168.201.46 255.255.255.255 FS (config-nfpp-define)#exit FS (config-nfpp)#define rip enable
Verification	● Run the **show nfpp define summary rip** command to display the configuration.
	Define rip summary: match etype 0x800 protocol 17 dst-port 520

	Maximum count of monitored hosts: 1000 Monitor period:600s (Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.) Interface Status Isolate-period Rate-limit Attack-threshold Global Enable 180 100/-/- 150/-/-
	● Run the **show nfpp define trusted-host rip** command to display the trusted hosts.
	Define rip: IP trusted host number is 1: IP address IP mask ---------- ------- 192.168.201.46 255.255.255.255 Total: 1 record(s)Global Enable 180 100/-/- 150/-/-
	● Run the **show nfpp define hosts rip** command to display the monitored hosts.
	If col_filter 1 shows '*', it means "hardware do not isolate host". VLAN interface IP address remain-time(s) ---- --------- ---------- ---------- 1 Gi0/5 192.168.201.47 160 Total: 1 host

Common Errors

N/A

21.4.8 Configuring Centralized Bandwidth Allocation

Configuration Effect

- Configure centralized bandwidth allocation so that Manage and Protocol packets are first processed when the network is busy.

Notes

- The following condition must be met: Valid percentage range of a type of packets ≤ 100% – Percentage of the sum of the other two types

Configuration Steps

Configuring the Maximum Bandwidth of Specified Packets

- (Mandatory) Manage, Route, and Protocol packets share the same default bandwidth.

Configuring the Maximum Percentage of Specified Packets in the Queue

- (Mandatory) By default, Manage packets occupy 30% of the bandwidth, Route packets occupy 25%, and Protocol packets occupy

45%.

Verification

Send a large number of protocol packets such as OSPF packets to a switch, causing high CPU utilization.

- When the host pings the switch, the pinging must be successful and no packet is lost.

Related Commands

Configuring the Maximum Bandwidth of Specified Packets

Command	**cpu-protect sub-interface** { **manage** | **protocol**|**route**} **pps** *pps_value*
Parameter Description	*pps_value*: Indicates the rate limit, ranging from 1 to 100,000.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring the Maximum Percentage of Specified Packets in the Queue

Command	**cpu-protect sub-interface** { **manage** | **protocol** | **route**} **percent** *percent_value*
Parameter Description	*percent_value*: Indicates the percentage of a type of packets in the queue, ranging from 1 to 100.
Command Mode	Global configuration mode
Usage Guide	The following condition must be met: Valid percentage range of a type of packets ≤ 100% – Percentage of the sum of the other two types

Configuration Example

Prioritizing Packets Sent to the CPU Through Centralized Bandwidth Allocation

Scenario	• Various types of mass packets exist in the network and belong to different centralized types.
Configuration Steps	• Configure the maximum bandwidth of specified packets. • Configure the maximum percentage of specified packets in the queue.
	FS# configure terminal FS(config)# cpu-protect sub-interface manage pps 5000 FS(config)# cpu-protect sub-interface manage percent 25
Verification	N/A

Common Errors

N/A

21.4.9 Enabling/Disabling All Guards

Configuration Effect

- Use the (no) all-guard enable command to enable or disable all attack guards so that you do not need to disable or enable them

one by one.

Notes

- Only basic guards (ARP, ICMP, IP, DHCP, DHCPv6, and ND) are applied.
- Only the global configuration is applied. Interface-based guard configuration remains the same.
- After the command is executed, basic guards are displayed by using the **show running-config** command.
- The **no all-guard enable** command just packs the **no** commands of all basic guards together. After you run the disabling command, the **no** commands of all basic guards are displayed under the **show running-config** command. After you run the enabling command, the default conditions are displayed under the **show running-config** command.

Configuration Steps

Running (no) all-guard enable in Global Configuration Mode

Verification

When a host sends a large number of packets corresponding to basic guards to a switch, such as ARP/ICMP packets, NFPP guard detection takes effect by default.

- Run the **no all-guard enable** command. With the **show cpu-protect** command used, NFPP ratelimit failure is displayed. With the **show nfpp xx-guard host** command used, no attacker is displayed. With the **show nfpp xx-guard summary** command used, the "disabled" status of guards is displayed.

Configuration Example

Prioritizing Packets Sent to the CPU Through Centralized Bandwidth Allocation

Scenario	● N/A
Configuration Steps	● N/A
	FS(config)#show running-config | begin nfpp nfpp log-buffer enable arp-guard rate-limit per-port 201 arp-guard attack-threshold per-port 210 ! FS(config)# nfpp FS(config-nfpp)#no all-guard enable FS(config-nfpp)#show running-config | begin nfpp nfpp log-buffer enable no arp-guard enable arp-guard rate-limit per-port 201

	arp-guard attack-threshold per-port 210 no icmp-guard enable no ip-guard enable no dhcp-guard enable no dhcpv6-guard enable no nd-guard enable ! FS(config-nfpp)#all-guard enable FS(config-nfpp)#show running-config | begin nfpp nfpp log-buffer enable arp-guard rate-limit per-port 201 arp-guard attack-threshold per-port 210 ! no service password-encryption !
Verification	N/A

Common Errors

N/A

21.4.10 Configuring NFPP Logging

Configuration Effect

- NFPP obtains a log from the dedicated log buffer at a certain rate, generates a system message, and clears this log from the dedicated log buffer.

Notes

- Logs are continuously printed in the log buffer, even if attacks have stopped.

Configuration Steps

Configuring the Log Buffer Size

- Mandatory.
- If the log buffer is full, new logs replace the old ones.
- If the log buffer overflows, subsequent logs replace the previous ones with all attributes marked with a hyphen (-) is displayed in the log buffer. The administrator needs to increase the log buffer size or the system message generation rate.

Configuring the Log Buffer Rate

- Mandatory.
- The log buffer rate depends on two parameters: the time period and the number of system messages generated in the time period.
- If both of the preceding two parameters are set to 0, system messages are immediately generated for logs but are not stored in the log buffer.

Enabling Log Filtering

- (Optional) Log filtering is disabled by default.
- Logs can be filtered based on an interface or VLAN.
- If log filtering is enabled, logs not meeting the filtering rule are discarded.

Enabling Log Printing

- (Mandatory) Logs are stored in the buffer by default.
- If you want to monitor attacks in real time, you can configure logs to be printed on the screen to export the log information in real time.

Verification

Check whether the configuration takes effect based on the log configuration and the number and interval of printed logs.

Related Commands

Configuring the Log Buffer Size

Command	**log-buffer entries** *number*
Parameter Description	*number*: Indicates the buffer size in the unit of the number of logs, ranging from 0 to 1,024.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring the Log Buffer Rate

Command	**log-buffer logs** *number_of_message* **interval** *length_in_seconds*
Parameter Description	*number_of_message*: Ranges from 0 to 1,024. The value 0 indicates that all logs are recorded in the log buffer and no system message is generated. *length_in_seconds*: Ranges from 0 to 86,400 (1 day). The value 0 indicates that logs are not recorded in the log buffer but system messages are instantly generated. This also applies to *number_of_message* and *length_in_seconds*. *number_of_message/length_in_second* indicates the system message generation rate.
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuring VLAN-based Log Filtering

Command	**logging vlan** *vlan-range*
Parameter Description	*vlan-range*: Records logs in a specified VLAN range. The value format is 1-3,5 for example.

Command Mode	NFPP configuration mode
Usage Guide	Run this command to filter logs so that only logs in the specified VLAN range are recorded. Between interface-based log filtering and VLAN-based log filtering, if either rule is met, logs are recorded in the log buffer.

Configuring Interface-based Log Filtering

Command	**logging interface** *interface-id*
Parameter Description	*interface-id*: Records logs of a specified interface.
Command Mode	NFPP configuration mode
Usage Guide	Run this command to filter logs so that only logs of the specified interface are recorded. Between interface-based log filtering and VLAN-based log filtering, if either rule is met, logs are recorded in the log buffer.

Enabling Log Printing

Command	**log-buffer enable**
Parameter Description	N/A
Command Mode	NFPP configuration mode
Usage Guide	N/A

Configuration Example

Configuring NFPP Logging

Scenario	● If attackers are too many, log printing will affect the usage of user interfaces, which requires restriction.
Configuration Steps	● Configure the log buffer size. ● Configure the log buffer rate. ● Configure VLAN-based log filtering.
	FS# configure terminal FS(config)# nfpp FS (config-nfpp)#log-buffer entries 1024 FS (config-nfpp)#log-buffer logs 3 interval 5 FS (config-nfpp)#logging interface vlan 1
Verification	● Run the **show nfpp log summary** command to display the configuration.
	Total log buffer size : 1024 Syslog rate : 3 entry per 5 seconds Logging: VLAN 1
	● Run the **show nfpp log buffer** command to display logs in the log buffer.

```
Protocol VLAN Interface   IP address          MAC address        Reason              Timestamp
-------- ---- ---------   ----------          -----------        ------              ---------
ARP        1     Gi0/5        192.168.206.2      001a.a9c2.4609     SCAN                2013-5-1 5:4:24
```

21.5 Monitoring

Clearing

Description	Command
Clears the ARP-guard scanning table.	**clear nfpp arp-guard scan**
Clears ARP-guard monitored hosts.	**clear nfpp arp-guard hosts**
Clears IP-guard monitored hosts.	**clear nfpp ip-guard hosts**
Clears ND-guard monitored hosts.	**clear nfpp nd-guard hosts**
Clears ICMP-guard monitored hosts.	**clear nfpp icmp-guard hosts**
Clears DHCP-guard monitored hosts.	**clear nfpp dhcp-guard hosts**
Clears DHCPv6-guard monitored hosts.	**clear nfpp dhcpv6-guard hosts**
Clears self-defined guard monitored hosts.	**clear nfpp define** *name* **hosts**
Clears NFPP logs.	**clear nfpp log**

Displaying

Description	Command
Displays ARP-guard configuration.	**show nfpp arp-guard summary**
Displays ARP-guard monitored hosts.	**show nfpp arp-guard hosts**
Displays the ARP-guard scanning table.	**show nfpp arp-guard scan**
Displays IP-guard configuration.	**show nfpp ip-guard summary**
Displays IP-guard monitored hosts.	**show nfpp ip-guard hosts**
Displays the IP-guard scanning table.	**show nfpp ip-guard trusted-host**
Displays ICMP-guard configuration.	**show nfpp icmp-guard summary**
Displays ICMP-guard monitored hosts.	**show nfpp icmp-guard hosts**
Displays the ICMP-guard scanning table.	**show nfpp icmp-guard trusted-host**
Displays DHCP-guard configuration.	**show nfpp dhcp-guard summary**
Displays DHCP-guard monitored hosts.	**show nfpp dhcp-guard hosts**
Displays DHCPv6-guard configuration.	**show nfpp dhcpv6-guard summary**
Displays DHCPv6-guard monitored hosts.	**show nfpp dhcpv6-guard hosts**
Displays ND-guard configuration.	**show nfpp nd-guard summary**
Displays self-defined guard configuration.	**show nfpp define summary** [*name*]
Displays the monitored hosts.	**show nfpp define hosts** *name*
Displays the trusted hosts.	**show nfpp define trusted-host** *name*
Displays NFPP logs.	**show nfpp log summary**
Displays the NFPP log buffer.	**show nfpp log buffer** [**statistics**]

22 Configuring DoS Protection

22.1 Overview

Denial of Service (DoS) attacks refer to attacks that cause DoS and aim to put computers or networks out of service.

DoS attacks are diversified in types and can be implemented in many ways, but have one common purpose, that is, prevent victim hosts or networks cannot receive, respond, or process external requests in time. In particular, on a layer-2 (L-2) network, DoS attack packets can be spread in the entire broadcast domain. If hackers maliciously initiate DoS attacks, some operating systems (OSs) may collapse. FS products supports the following anti DoS attack functions:

- Denying land attacks
- Denying invalid TCP packets
- Denying invalid layer-4 (L4) ports

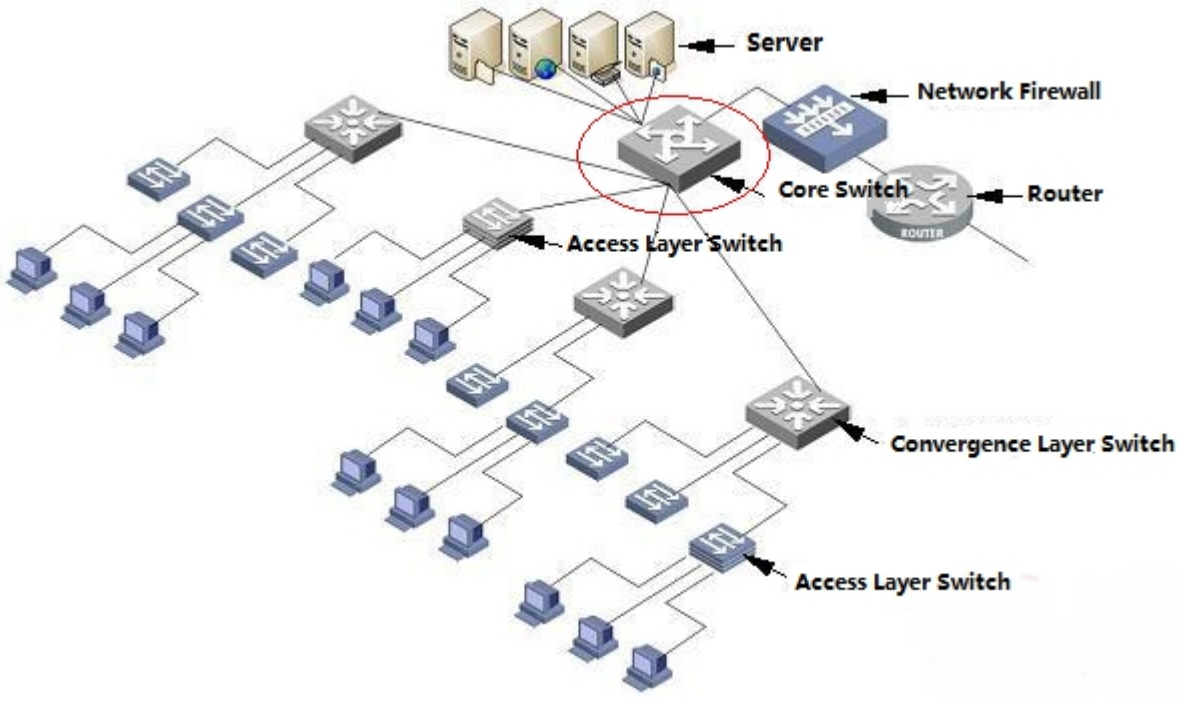
22.2 Applications

Application	Description
Protecting Servers Against DoS Attacks	On a campus network, configure the anti DoS attack function on the devices connected to servers to effectively reduce the negative impacts brought by DoS attacks to servers.

22.2.1 Protecting Servers Against DoS Attacks

As show in Figure 23- 1, servers are connected to the core switch. The anti DoS attack function is configured on the core switch to prevent malicious DoS attacks and ensure that servers can provide services normally.

Figure 23- 1



Deployment

Enable the function of denying land attacks on the core switch to protect servers against land attacks.

Enable the function of denying invalid TCP packets on the core switch to protect servers against invalid TCP packets.

Enable the function of denying invalid L4 ports on the core switch to protect servers against attacks caused by invalid L4 ports.

22.3 Features

Overview

Feature	Description
Denying Land Attacks	Drop packets with the same source and destination IP addresses or the same L4 source and destination port IDs on the device to prevent these packets from attacking OSs on the network.
Denying Invalid TCP Packets	Drop invalid TCP packets on the device to prevent invalid TCP packets from attacking OSs on the network. (For details about the definition of invalid TCP packets, see "Denying Invalid TCP Packets".
Denying Invalid L4 Ports	Drop packets with the same L4 source and destination port IDs on the device to prevent these packets from attacking OSs on the network.

22.3.1 Denying Land Attacks

This function protects servers against land attacks.

Working Principle

In a land attack, the attacker sets the source and destination IP addresses or the L4 source and destination port IDs in a SYN packet to the same address of the target host. Consequently, the attacked host will be trapped in an infinite loop or even collapse when attempting to set up a TCP connection with itself.

If the function of denying land attacks is enabled, the device checks packets based on characteristics of land packets (that is, SYN packets with the same source and destination IP addresses), and drops invalid packets.

Related Configuration

Enabling the Function of Denying Land Attacks

By default, the function of denying land attacks is disabled.

Run the **ip deny land** command to enable or disable the function of denying land attacks.

22.3.2 Denying Invalid TCP Packets

This function protects servers against invalid TCP packets.

Working Principle

There are several flag fields in the TCP packet header:

- SYN: Connection establishment flag. The TCP SYN packet is used to set this flag to 1 to request establishment of a connection.
- ACK: Acknowledgement flag. In a TCP connection, this field must be available in every flag (except the first packet, that is, the TCP SYN packet) as the acknowledgement of the previous packet.
- FIN: Finish flag. When a host receives the TCP packet with the FIN flag, the host disconnects the TCP connection.
- RST: Reset flag. When the IP protocol stack receives a TCP packet that contains a non-existent destination port, it responds with a

packet with the RST flag.

- PSH: This flag notifies the protocol stack to submit TCP data to the upper-layer program for processing as soon as possible.

In invalid TCP packets, flag fields are set improperly so that the processing resources of hosts are exhausted or even the system collapses. The following lists several common methods for setting flag fields in invalid TCP packets:

- TCP packets with both the SYN and FIN flags

Normally, a TCP packet cannot contain both the SYN and FIN flags. In addition, RFC does not stipulate how the IP protocol stack should process such invalid packets containing both the SYN and FIN flags. Therefore, the protocol stack of each OS may process such packets in different ways when receiving these packets. Attackers can use this feature to send packets containing both the SYN and FIN flags to identify the OS type and initiate attacks on this OS.

- TCP packets without any flag

Normally, a TCP packet contains at least one of the five flags, including SYN, FIN, ACK, RST, and PSH. The first TCP packet (TCP SYN packet) must contain the SYN flag, and the subsequent packets contain the ACK flag. Based on such assumptions, some protocol stack does not specify the method for processing TCP packets without any flag, and therefore may collapse if such protocol stack receives TCP packets without any flag. Attackers use this feature to initiate attacks on target hosts.

- TCP packets with the FIN flag but without the ACK flag

Normally, except the first packet (TCP SYN packet), all other packets, including the packets with the FIN flag, contain the ACK flag. Some attackers may send TCP packets with the FIN flag but without the ACK flag to the target hosts, causing breakdown of the target hosts.

- TCP packets with the SYN flag and the source port ID set to a value between 0 and 1,023

Port IDs 0 to 1,023 are known port IDs allocated by the Internet Assigned Numbers Authority (IANA). In most systems, these port IDs can be used only by the system (or root) processes or programs run by privileged users. These ports (0–1023) cannot be used as the source port IDs in the first TCP packets (with the SYN flag) sent by clients.

If the function of denying invalid TCP packets is enabled, the device checks packets based on characteristics of invalid TCP packets, and drops invalid TCP packets.

Related Configuration

Enabling the Function of Denying Invalid TCP Packets

By default, the function of denying invalid TCP packets is disabled.

Run the **ip deny invalid-tcp** command to enable or disable the function of denying invalid TCP packets.

22.3.3 Denying Invalid L4 Ports

This function protects servers against invalid L4 ports.

Working Principle

Attackers sends packets in which the IP address of the target host is the same as the L4 port ID of the host to the host target. As a result, the target host sends TCP connection setup requests to itself. Under such attacks, resources of the target host will soon be exhausted and the system will collapse.

If the function of denying invalid L4 ports is enabled, the device checks the L4 source port ID and destination port ID in the packets. If they are the same, the device drops the packets.

Related Configuration

Enabling the Function of Denying Invalid L4 Ports

By default, the function of denying invalid L4 ports is disabled.

Run the **ip deny invalid-l4port** command to enable or disable the function of denying invalid L4 ports.

22.4 Configuration

Configuration Item	Description and Command	
Configuring the Function of Denying Land Attacks	Optional.	
	ip deny land	Enables the function of denying land attacks.
Configuring the Function of Denying Invalid TCP Packets	Optional.	
	ipdeny invalid-tcp	Enables the function of denying invalid TCP packets.
Configuring the Function of Denying Invalid L4 Ports	Optional.	
	ip deny invalid-l4port	Enables the function of denying invalid L4 ports.

22.4.1 Configuring the Function of Denying Land Attacks

Configuration Effect

Enable the function of denying land attacks. Then, the device checks packets based on characteristics of land packets, and drops land packets.

Configuration Steps

Enabling the Function of Denying Land Attacks

- Mandatory.
- Perform this configuration on a device connected to a server.

Verification

- Run the **showipdenyland** command to display the status of the function of denying land attacks.
- After this function is enabled, construct a land attack packet and confirm that this packet cannot be forwarded.

Related Commands

Configuring the Function of Denying Land Attacks

Command	**[no] ip deny land**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Enabling the Function of Denying Land Attacks

Configuration Steps	● Enable the function of denying land attacks in global configuration mode.
	FS# configure terminal FS(config)# ip deny land FS(config)# end
Verification	Run the **showipdenyland** command to display the status of the function of denying land attacks. The following example shows how to display the status of the function of denying land attacks: FS#show ip deny land DoS Protection Mode State ---------- ----- protect against land attack On

22.4.2 Configuring the Function of Denying Invalid TCP Packets

Configuration Effect

Enable the function of denying invalid TCP packets. Then, the device checks packets based on characteristics of invalid TCP packets, and drops invalid TCP packets.

Configuration Steps

Enables the Function of Denying Invalid TCP Packets

- Mandatory.
- Perform this configuration on a device connected to a server.

Verification

- Run the **show ip deny invalid-tcp** command to display the status of the function of denying invalid TCP packets.
- After this function is enabled, construct an invalid TCP packet and confirm that this packet cannot be forwarded.

Related Commands

Configuring the Function of Denying Invalid TCP Packets

Command	**[no] ip deny invalid-tcp**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Enabling the Function of Denying Invalid TCP Packets

Configuration Steps	● Enable the function of denying invalid TCP packets in global configuration mode.
	FS# configure terminal FS(config)# ip deny invalid-tcp FS(config)# end
Verification	Run the **show ip deny invalid-tcp** command to display the status of the function of denying invalid TCP packets.
	The following example shows how to display the status of the function of denying invalid TCP packets: FS#show ip deny invalid-tcp DoS Protection Mode State ---------- ----- protect against invalid tcp attack On

22.4.3 Configuring the Function of Denying Invalid L4 Ports

Configuration Effect

Enable the function of denying invalid L4 ports. Then, the device checks the L4 source port ID and destination port ID in the packets. If they are the same, the device drops the packets.

Configuration Steps

Enabling the Function of Denying Invalid L4 Ports

- Mandatory.
- Perform this configuration on a device connected to a server.

Verification

- Run the **show ip deny invalid-l4port** command to display the status of the function of denying invalid L4 ports.
- After this function is enabled, construct a packet in which the L4 source port ID is the same as the destination port ID and confirm that this packet cannot be forwarded.

Related Commands

Configuring the Function of Denying Invalid L4 Ports

Command	**[no] ip deny invalid-l4port**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Enabling the Function of Denying Invalid L4 Ports

Configuration Steps	● Enable the function of denying invalid L4 ports in global configuration mode.
	FS# configure terminal FS(config)# ip deny invalid-l4port FS(config)# end
Verification	Run the **show ip deny invalid-l4port** command to display the status of the function of denying invalid L4 ports. The following example shows how to display the status of the function of denying invalid L4 ports: FS#show ip deny invalid-l4port DoS Protection Mode State ---------- ----- protect against invalid l4port attack On

22.5 Monitoring

Displaying

Description	Command
Displays the status of the function of denying land attacks.	**Showipdeny land**
Displays the status of the function of denying invalid TCP packets.	**show ip deny invalid-tcp**
Displays the status of the function of denying invalid L4 ports.	**show ip deny invalid-l4port**
Displays the status of all antiDoS attack functions.	**show ip deny**

Chapter8 ACL & QoS Configuration

1. Configuring ACL
2. Configuring QoS

1 Configuring the ACL

1.1 Overview

Access control list (ACL) is also called access list or firewall. It is even called packet filtering in some documents. The ACL defines rules to determine whether to forward or drop data packets arriving at a network interface.

ACLs are classified by function into two types:

- Security ACLs: Used to control data flows that are allowed to pass through a network device.
- Quality of service (QoS) ACLs: Used to classify and process data flows by priority.

ACLs are configured for a lot of reasons. Major reasons include:

- Network access control: To ensure network security, rules are defined to limit access of users to some services (for example, only access to the WWW and email services is permitted, and access to other services such as Telnet is prohibited), or to allow users to access services in a specified period of time, or to allow only specified hosts to access the network.
- QoS: QoS ACLs are used to preferentially classify and process important data flows. For details about the use of QoS ALCs, see the configuration manual related to QoS.

1.2 Applications

Application	Description
Access Control of an Enterprise Network	On an enterprise network, the network access rights of each department, for example, access rights of servers and use permissions of chatting tools (such as QQ and MSN), must be controlled according to requirements.

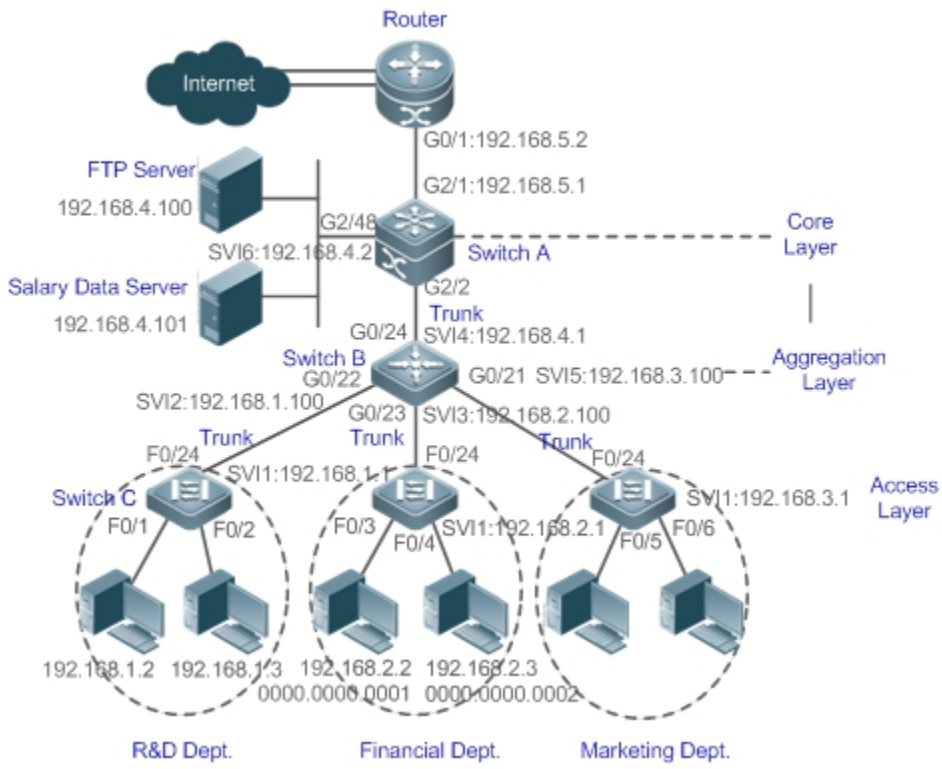
1.2.1 Access Control of an Enterprise Network

Scenario

Internet viruses can be found everywhere. Therefore, it is necessary to block ports that are often used by viruses to ensure security of an enterprise network as follows:

- Allow only internal PCs to access the server.
- Prohibit PCs of a non-financial department from accessing PCs of the financial department, and prohibit PCs of a non-R&D department from accessing PCs of the R&D department.
- Prohibit the staff of the R&D department from using chatting tools (such as QQ and MSN) during working hours from 09:00 to 18:00.

Figure 1- 1



Remarks	Switch C at the access layer:It is connected to PCs of each department and to Switch B at the aggregation layer through the gigabit optical fiber (trunk mode). Switch B at the aggregation layer:Multiple virtual local area networks (VLANs) are divided. One VLAN is defined for one department. These VLANs are connected to Switch A at the core layer through the 10-gigabit optical fiber (trunk mode). Switch A at the core layer:It is connected to various servers, such as the File Transfer Protocol (FTP) server and Hypertext Transfer Protocol (HTTP) server, and to the Internet through firewalls.

Deployment

- Configure an extended ACL on the port G2/1 to filter data packets, thus protecting the network against the viruses. This port is located on a core-layer device (Switch A) and used to connect Switch A to the uplink port G2/1 of a router.
- Allow only internal PCs to access servers, and prohibit external PCs from accessing servers. Define and apply the extended IP ACLs on G2/2 or switch virtual interface (SVI) 2 that is used to connect Switch A to an aggregation layer device or server.
- Prohibit mutual access between specified departments. Define and apply the extended IP ACLs on G0/22 and G0/23 of Switch B.
- Configure and apply the time-based extended IP ACLs on SVI 2 of Switch B to prohibit the R&D department from using chatting tools (such as QQ and MSN) in a specified period of time.

1.3 Features

Basic Concepts

↘ ACL

ACLs include basic ACLs and dynamic ACLs.

You can select basic or dynamic ACLs as required. Generally, basic ACLs can meet the security requirements. However, experienced hackers may use certain software to access the network by means of IP address spoofing. If dynamic ACLs are used, users are requested to pass identify authentication before accessing the network, which prevents hackers from intruding the network. Therefore, you can use dynamic ACLs in some sensitive areas to guarantee network security.

IP address spoofing is an inherent problem of all ACLs, including dynamic ACLs. Hackers may use forged IP addresses to access the network during the validity period of authenticated user identities. Two methods are available to resolve this problem. One is to set the idle time of user access to a smaller value, which increases the difficulty in intruding networks. The other is to encrypt network data using the IPSec protocol, which ensures that all data is encrypted when arriving at a device.

ACLs are generally configured on the following network devices:

- Devices between the internal network and the external network (such as the Internet)
- Devices on the border of two network segments
- Devices connected to controlled ports

ACL statements must be executed in strict compliance with their sequence in the ACL. Comparison starts from the first statement. Once the header of a data packet matches a statement in the ACL, the subsequent statements are ignored and no longer checked.

Input/Output ACLs, Filtering Field Template, and Rules

When receiving a packet on an interface, the device checks whether the packet matches any access control entry (ACE) in the input ACL of this interface. Before sending a packet through a interface, the device checks whether the packet matches any ACE in the output ACL of this interface.

When different filtering rules are defined, all or only some rules may be applied simultaneously. If a packet matches an ACE, this packet is processed according to the action policy (permit or deny) defined in this ACE. ACEs in an ACL identify Ethernet packets based on the following fields in the Ethernet packets:

Layer 2 (L2) fields:

- 48-bit source MAC address (containing all 48 bits)
- 48-bit destination MAC address (containing all 48 bits)
- 16-bit L2 type field

Layer 3 (L3) fields:

- Source IP address field (All source IP address values can be specified, or the subnet can be used to define a type of data flows.)
- Destination IP address field (All destination IP address values can be specified, or the subnet can be used to define a type of data flows.)
- Protocol type field

Layer 4 (L4) fields:

- Either a TCP source or destination port is specified, or both are specified, or the range of the source or destination port is specified.
- Either a UDP source or destination port is specified, or both are specified, or the range of the source or destination port is specified.

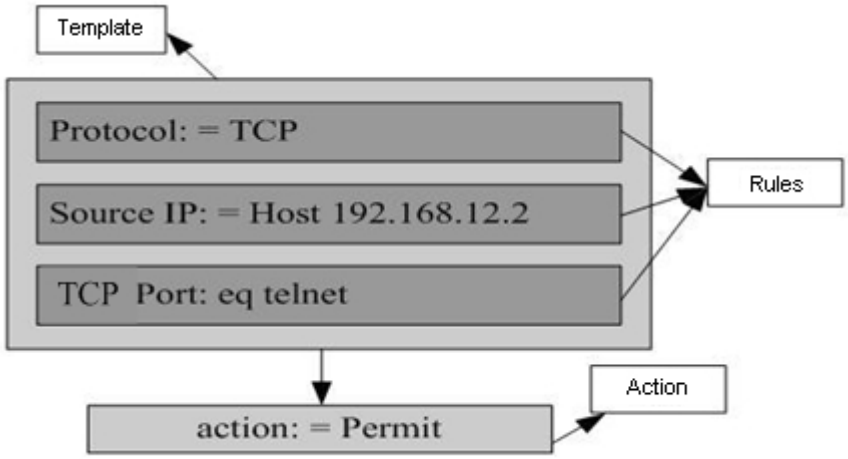
Filtering fields refer to the fields in packets that can be used to identify or classify packets when an ACE is generated. A filtering field template is a combination of these fields. For example, when an ACE is generated, packets are identified and classified based on the destination IP address field in each packet; when another ACE is generated, packets are identified and classified based on the source IP address field and UDP source port field in each packet. The two ACEs use different filtering field templates.

Rules refer to values of fields in the filtering field template of an ACE.For example, the content of an ACE is as follows:

```
permit tcp host 192.168.12.2 any eq telnet
```

In this ACE, the filtering field template is a combination of the following fields:source IP address field, IP protocol field, and TCP destination port field. The corresponding values (rules) are as follows:source IP address = Host 192.168.12.2; IP protocol = TCP; TCP destination port = Telnet.

Figure 1- 2 Analysis of the ACE: permit tcp host 192.168.12.2 any eq telnet



A filtering field template can be a combination of L3 and L4 fields, or a combination of multiple L2 fields. The filtering field template of a standard or an extended ACL, however, cannot be a combination of L2 and L3 fields, a combination of L2 and L4 fields, or a combination of L2, L3, and L4 fields. To use a combination of L2,L3, and L4 fields, you can use the expert ACLs.

An SVI associated with ACLs in the outgoing direction supports the IP standard, IP extended, MAC extended, and expert ACLs.

If an MAC extended or expert ACL is configured to match the destination MAC address and is applied to the outgoing direction of the SVI, the related ACE can be configured but cannot take effect. If an IP extended or expert ACL is configured to match the destination IP address, but the destination IP address is not in the subnet IP address range of the associated SVI, the configured ACL cannot take effect. For example, assume that the address of VLAN 1 is **192.168.64.1 255.255.255.0**, an IP extended ACL is created, and the ACE is **deny udp any 192.168.65.1 0.0.0.255 eq 255**. If this ACL is applied to the outgoing interface of VLAN 1, the ACL cannot take effect because the destination IP address is not in the subnet IP address range of VLAN 1. If the ACE is **deny udp any 192.168.64.1 0.0.0.255 eq 255**, the ACL can take effect because the destination IP address is in the subnet IP address range of VLAN 1.

On a switch, if ACLs are applied to the outgoing direction of a physical port or an aggregate port (AP), the ACLs can filter only well-known packets (unicast or multicast packets), but not unknown unicast packets. That is, for unknown or broadcast packets, ACLs configured in the outgoing direction of a port does not take effect.

On a switch, if the input ACL and DOT1X, global IP+MAC binding, port security, and IP source guard are shared among all ports, the permit and default deny ACEs do not take effect, but other deny ACEs take effect.

On a switch, if the input ACL and QoS are shared, the permit ACEs do not take effect, other deny ACEs take effect, and the default deny ACE takes effect after the QoS ACE takes effect.

On a switch, you can run the **norgos-security compatible** command to make the permit and deny ACEs take effect at the same time when the port-based input ACL and DOT1X, global IP+MAC binding, port security, and IP source guard are shared.

If ACEs are added to an ACL and then the switch is restarted after an ACL is applied to the incoming direction of multiple SVIs, the ACL may fail to be configured on some SVIs due to the limited hardware capacity.

If an expert ACL is configured and applied to the outgoing direction of an interface, and some ACEs in this ACL contain the L3 matching information (e.g. the IP address and L4 port), non-IP packets sent to the device from this interface cannot be controlled by the permit and deny ACEs in this ACL.

If ACEs of an ACL (IP ACL or expert extended ACL) are configured to match non-L2 fields (such as SIP and DIP), the ACL does not take effect on tagged MPLS packets.

ACL Logging

To allow users better learn the running status of ACLs on a device, you can determine whether to specify the ACL logging option as required when adding ACEs. If this option is specified, logs are output when packets matching ACEs are found. ACL logs are displayed based on ACEs. That is, the device periodically displays ACEs with matched packets and the number of matched packets. An example of the log is as follows:

```
*Sep    9 16:23:06: %ACL-6-MATCH: ACL 100 ACE 10 permit icmp any any, match 78 packets.
```

To control the amount of logs and output frequency, you can configure the log update interval respectively for the IPv4 ACL and the IPv6 ACL.

An ACE containing the ACL logging option consumes more hardware resources. If all configured ACEs contain this option, the ACE capacity of a device will be reduced by half.

By default, the log update interval is 0, that is, no log is output. After the ACL logging option is specified in an ACE, you need to configure the log update interval to output related logs.

For an ACE containing the ACL logging option, if no packet is matched in the specified interval, no packet matching log related to this ACE will be output. If matched packets are found in the specified interval, packet matching logs related to this ACE will be output when the interval expires. The number of matched packets is the total number of packets that match the ACE during the specified interval, that is, the period from the previous log output to the current log output.

Only switches support the ACL logging function.

ACL Packet Matching Counters

To implement network management, users may want to know whether an ACE has any matched packets and how many packets are matched. ACLs provide the ACE-based packet matching counters. You can enable or disable packet matching counters for all ACEs in an ACL, which can be an IP ACL, MAC ACL, expert ACL, or IPv6 ACL. In addition, you can run the **clear counters access-list** [*acl-id* | *acl-name*] command to reset ACL counters for a new round of statistics.

Enabling ACL counters requires more hardware entries. In an extreme case, this will reduce by half the number of ACEs that can be configured on a device.

Only switches support the ACL packet matching counters.

Overview

Feature	Description
IP ACL	Control incoming or outgoing IPv4 packets of a device based on the L3 or L4 information in the IPv4 packet header.
MAC Extended ACL	Control incoming or outgoing L2 packets of a device based on the L2 information in the Ethernet packet header.
Expert Extended ACL	Combine the IP ACL and MAC extended ACL into an expert extended ACL, which controls (permits or denies) incoming or outgoing packets of a device using the same rule based on the L2, L3, and L4 information in the packet header.
IPv6 ACL	Control incoming or outgoing IPv6 packets of a device based on the L3 or L4 information in the IPv6 packet header.
ACL80	Customize the matching fields and mask for scenarios where fixed matching fields cannot meet the requirements.
ACL Redirection	Redirect incoming packets of a device that match ACEs to a specified outgoing interface.
Global Security ACL	Make an ACL take effect in the incoming direction of all interfaces, instead of applying the ACL on every interface.
Security Channel	Allow packets to bypass the check of access control applications, such as DOT1X and Web authentication, to meet requirements of some special scenarios.
SVI Router ACL	Enable users in the same VLAN to communicate with each other.
ACL Logging	Output ACL packet matching logs at a specified interval according to requirements. The logs help users learn the packet matching result of a specified ACE.

1.3.1 IP ACL

The IP ACL implements refined control on incoming and outgoing IPv4 packets of a device. You can permit or deny the entry of specific IPv4 packets to a network according to actual requirements to control access of IP users to network resources.

Working Principle

Define a series of IP access rules in the IP ACL, and then apply the IP ACL either in the incoming or outgoing direction of an interface or globally. The device checks whether the incoming or outgoing IPv4 packets match the rules and accordingly forwards or blocks these packets.

To configure an IP ACL, you must specify a unique name or ID for the ACL of a protocol so that the protocol can uniquely identify each ACL. The following table lists the protocols that can use IDs to identify ACLs and the range of IDs.

Protocol	ID Range
Standard IP	1–99, 1300–1999
Extended IP	100–199, 2000–2699

Basic ACLs include the standard IP ACLs and extended IP ACLs. Typical rules defined in an ACL contain the following matching fields:

- Source IP address
- Destination IP address
- IP protocol number
- L4 source port ID or ICMP type
- L4 destination port ID or ICMP code

The standard IP ACL (ID range: 1–99, 1300–1999) is used to forward or block packets based on the source IP address, whereas the extended IP ACL (ID range: 100–199, 2000–2699) is used to forward or block packets based on a combination of the preceding matching fields.

For an individual ACL, multiple independent ACL statements can be used to define multiple rules. All statements reference the same ID or name so that these statements are bound with the same ACL. However, more statements mean that it is increasingly difficult to read and understand the ACL.

For routing products, the ICMP code matching field in an ACL rule is ineffective for ICPM packets whose ICPM type is 3. If the ICPM code of ICMP packets to be matched is configured in an ACL rule, the ACL matching result of incoming ICMP packets of a device whose ICPM type is 3 may be different from the expected result.

Implicit "Deny All Traffic" Rule Statement

At the end of every IP ACL is an implicit "deny all traffic" rule statement. Therefore, if a packet does not match any rule, the packet will be denied.

For example:

```
access-list 1 permit host 192.168.4.12
```

This ACL permits only packets sent from the source host 192.168.4.12, and denies packets sent from all other hosts. This is because the following statement exists at the end of this ACL: **access-list 1 deny any**.

If the ACL contains only the following statement:

```
access-list 1 deny host 192.168.4.12
```

Packets sent from any host will be denied when passing through this port.

When defining an ACL, you must consider the routing update packets. As the implicit "deny all traffic" statement exists at the end of an ACL, all routing update packets may be blocked.

Input Sequence of Rule Statements

Every new rule is added to the end of an ACL and in front of the default rule statement. The input sequence of statements in an ACL is very important. It determines the priority of each statement in the ACL. When determining whether to forward or block packets, a device

compares packets with rule statements based on the sequence that rule statements are created. After locating a matched rule statement, the device does not check any other rule statement.

If a rule statement is created and denies all traffic, all subsequent statements will not be checked.

For example:

```
access-list 101 deny ip any any
access-list 101 permittcp 192.168.12.0 0.0.0.255    eqtelnetany
```

The first rule statement denies all IP packets. Therefore, Telnet packets from the host on the network 192.168.12.0/24 will be denied. After the device finds that packets match the first rule statement, it does not check the subsequent rule statements any more.

Related Configuration

Configuring an IP ACL

By default, no IP ACL is configured on a device.

Run the **ip access-list { standard | extended } {*acl-name* | *acl-id*}** command in global configuration mode to create a standard or an extended IP ACL and enter standard or extended IP ACL mode.

Adding ACEs to an IP ACL

By default, a newly created IP ACL contains an implicit ACE that denies all IPv4 packets. This ACE is hidden from users, but takes effect when the ACL is applied to an interface. That is, all IPv4 packets will be discarded. Therefore, if you want the device to receive or send some specific IPv4 packets, add some ACEs to the ACL.

For a standard IP ACL, add ACEs as follows:

- No matter whether the standard IP ACL is a named or number ACL, you can run the following command in standard IP ACL mode to add an ACE:

[*sn*] { **permit** | **deny** } {**host***source*| **any** | *sourcesource-wildcard* } [**time-range***time-range-name*] [**log**]

- For a numbered standard IP ACL, you can also run the following command in global configuration mode to add an ACE:

access-list *acl-id* { **permit** | **deny** } {**host***source*| **any** | *sourcesource-wildcard* } [**time-range***tm-rng-name*][**log**]

For an extended IP ACL, you can add ACEs as follows:

- No matter whether the extended IP ACL is a named or numbered ACL, you can run the following command in extended IP ACL mode to add an ACE:

[*sn*] { **permit** | **deny** } *protocol*{**host***source*| **any** | *sourcesource-wildcard* } {**host***destination* | **any** | *destination destination-wildcard* }[[**precedence***precedence* [**tos** *tos]*] | **dscp***dscp*] [**fragment**] [**time-range***time-range-name*] [**log**]

- For a numbered extended IP ACL, you can also run the following command in global configuration mode to add an ACE:

access-list *acl-id* { **permit** | **deny** } *protocol*{**host***source*| **any** | *sourcesource-wildcard* } {**host***destination* | **any** | *destination destination-wildcard* }[[**precedence***precedence* [**tos** *tos]*] | **dscp***dscp*] [**fragment**] [**time-range***time-range-name*] [**log**]

Applying an IP ACL

By default, the IP ACL is not applied to any interface/VXLAN, that is, the IP ACL does not filter incoming or outgoing IP packets of the device.

Run the **ip access-group { *acl-id* | *acl-name* } { in| out }[reflect]** command in interface/VXLAN configuration mode to apply a standard or an extended IP ACL to a specified interface/VXLAN. By default, a reflexive ACL is disabled on a router. You can run the **reflect** command to enable the reflexive ACL. The working principle of the reflexive ACL is as follows:

a. A temporary ACL is automatically generated based on the L3 and L4 information of the traffic originated by the internal network. The temporary ACL is created according to the following principles: The IP protocol number remains unchanged, the source and destination IP addresses are swapped, and the TCP/UDP source and destination ports are also swapped.

b. The router allows traffic to enter the internal network only when the L3 and L4 information of the returned traffic exactly matches that of the temporary ACL previously created based on the outgoing traffic.

1.3.2 MAC Extended ACL

The MAC extended ACL implements refined control on incoming and outgoing packets based on the L2 header of packets. You can permit or deny the entry of specific L2 packets to a network, thus protecting network resources against attacks or control users' access to network resources.

Working Principle

Define a series of MAC access rules in the MAC extended ACL, and then apply the ACL to the incoming or outgoing direction of an interface. The device checks whether the incoming or outgoing packets match the rules and accordingly forwards or blocks these packets.

To configure an MAC extended ACL, you must specify a unique name or ID for this ACL to uniquely identify the ACL. The following table lists the range of IDs that identify MAC extended ACLs.

Protocol	ID Range
MAC extended ACL	700–799

Typical rules defined in an MAC extended ACL include:

- Source MAC address
- Destination MAC address
- Ethernet protocol type

The MAC extended ACL (ID range: 700–799) is used to filter packets based on the source or destination MAC address and the Ethernet type in the packets.

For an individual MAC extended ACL, multiple independent ACL statements can be used to define multiple rules. All statements reference the same ID or name so that these statements are bound with the same ACL. However, more statements mean that it is increasingly difficult to read and understand the ACL.

If ACEs in an MAC extended ACL are not defined specifically for IPv6 packets, that is, the Ethernet type is not specified or the value of the Ethernet type field is not 0x86dd, the MAC extended ACL does not filter IPv6 packets. If you want to filter IPv6 packets, use the IPv6 extended ACL.

Implicit "Deny All Traffic" Rule Statement

At the end of every MAC extended ACL is an implicit "deny all traffic" rule statement. Therefore, if a packet does not match any rule, the packet will be denied.

For example:

```
access-list 700 permit host 00d0.f800.0001 any
```

This ACL permits only packets from the host with the MAC address 00d0.f800.0001, and denies packets from all other hosts. This is because the following statement exists at the end of this ACL: **access-list 700 deny any any**.

Related Configuration

Configuring an MAC Extended ACL

By default, no MAC extended ACL is configured on a device.

Run the **mac access-list extended {*acl-name* | *acl-id*}** command in global configuration mode to create an MAC extended ACL and enter MAC extended ACL mode.

Adding ACEs to an MAC Extended ACL

By default, a newly created MAC extended ACL contains an implicit ACE that denies all L2 packets. This ACE is hidden from users, but takes effect when the ACL is applied to an interface. That is, all L2 packets will be discarded. Therefore, if you want the device to receive or send some specific L2 packets, add some ACEs to the ACL.

You can add ACEs to an MAC extended ACL as follows:

- No matter whether the MAC extended ACL is a named or numbered ACL, you can run the following command in MAC extended ACL mode to add an ACE:

[*sn*] { **permit** | **deny** } {**any** | **host** *src-mac-addr* | *src-mac-addrmask*}{**any** | **host** *dst-mac-addr* | *dst-mac-addrmask*} [*ethernet-type*] [**cos***cos*] [**inner***cos*] [**time-range***tm-rng-name*]

- For a numbered MAC extended ACL, you can also run the following command in global configuration mode to add an ACE:

access-list *acl-id* { **permit** | **deny** } {**any** | **host** *src-mac-addr* | *src-mac-addrmask* }{**any** | **host** *dst-mac-addr* | *dst-mac-addrmask* } [*ethernet-type*] [**cos***cos*] [**inner***cos*] [**time-range***time-range-name*]

Applying an MAC Extended ACL

By default, the MAC extended ACL is not applied to any interface, that is, the created MAC extended ACL does not filter incoming or outgoing L2 packets of a device.

Run the **mac access-group { *acl-id* | *acl-name* } { in| out }** command in interface/VXLAN configuration mode to apply an MAC extended ACL to a specified interface/VXLAN.

1.3.3 Expert Extended ACL

You can create an expert extended ACL to match the L2 and L3 information in packets using the same rule. The expert extended ACL can be treated as a combination and enhancement of the IP ACL and the MAC extended ACL because the expert extended ACL can contain ACEs in both the IP ACL and the MAC extended ACL. In addition, the VLAN ID can be specified in the expert extended ACL to filter packets.

Working Principle

Define a series of access rules in the expert extended ACL, and then apply the ACL in the incoming or outgoing direction of an interface. The device checks whether incoming or outgoing packets match the rules and accordingly forwards or blocks these packets.

To configure an expert extended ACL, you must specify a unique name or ID for this ACL so that the protocol can uniquely identify each ACL. The following table lists the ID range of the expert extended ACL.

Protocol	ID Range
Expert extended ACL	2700–2899

When an expert extended ACL is created, defined rules can be applied to all packets. The device determines whether to forward or block packets by checking whether packets match these rules.

Typical rules defined in an expert extended ACL include:

- All information in the basic ACL and MAC extended ACL
- VLAN ID

The expert extended ACL (ID range: 2700–2899) is a combination of the basic ACL and MAC extended ACL, and can filter packets based on the VLAN ID.

For an individual expert extended ACL, multiple independent statements can be used to define multiple rules. All statements reference the same ID or name so that these statements are bound with the same ACL.

If rules in an expert extended ACL are not defined specifically for IPv6 packets, that is, the Ethernet type is not specified or the value of the Ethernet type field is not 0x86dd, the expert extended ACL does not filter IPv6 packets. If you want to filter IPv6 packets, use the IPv6 extended ACL.

Implicit "Deny All Traffic" Rule Statement

At the end of every expert extended ACL is an implicit "deny all traffic" rule statement. Therefore, if a packet does not match any rule, the packet will be denied.

For example:

```
access-list 2700permit 0x0806 any any any any any
```

This ACL permits only ARP packets whose Ethernet type is 0x0806, and denies all other types of packets. This is because the following statement exists at the end of this ACL: **access-list 2700 deny any any any any**.

Related Configuration

Configuring an Expert Extended ACL

By default, no expert extended ACL is configured on a device.

Run the **expert access-list extended {*acl-name* | *acl-id* }** command in global configuration mode to create an expert extended ACL and enter expert extended ACL mode.

Adding ACEs to an Expert Extended ACL

By default, a newly created expert extended ACL contains an implicit ACE that denies all packets. This ACE is hidden from users, but takes effect when the ACL is applied to an interface. That is, all L2 packets will be discarded. Therefore, if you want the device to receive or send some specific L2 packets, add some ACEs to the ACL.

You can add ACEs to an expert extended ACL as follows:

- No matter whether the expert extended ACL is a named or numbered ACL, you can run the following command in expert extended ACL mode to add an ACE:

[*sn*] { **permit** | **deny** } [*protocol*| [*ethernet-type*] [**cos** [*out*] [**inner** *in*]]] [[**VID** [*out*] [**inner** *in*]]] { *sourcesource-wildcard* | **host***source* | **any** } { **host** *source-mac-address* | **any** } { *destination destination-wildcard* | **host***destination* | **any** } { **host** *destination-mac-address* | **any** } [**precedence***precedence*] [**tos** *tos*] [**fragment**] [**range***lowerupper*] [**time-range***time-range-name*]]

- For a numbered expert extended ACL, you can also run the following command in expert extended ACL mode to add an ACE:

access-list *acl-id* { **permit** | **deny** } [*protocol*| [*ethernet-type*] [**cos** [*out*] [**inner** *in*]]] [[**VID** [*out*] [**inner** *in*]]] { *sourcesource-wildcard* | **host***source* | **any** } { **host** *source-mac-address* | **any** } { *destination destination-wildcard* | **host***destination* | **any** } { **host** *destination-mac-address* | **any** } [**precedence***precedence*] [**tos** *tos*] [**fragment**] [**range***lowerupper*] [**time-range***time-range-name*]]

Applying an Expert Extended ACL

By default, the expert extended ACL is not applied to any interface, that is, the created expert extended ACL does not filter incoming or outgoing L2 or L3 packets of a device.

Run the **expert access-group { *acl-id* | *acl-name* } { in| out }** command in interface/VXLAN configuration mode to apply an expert extended ACL to a specified interface/VXLAN.

1.3.4 IPv6 ACL

The IPv6 ACL implements refined control on incoming and outgoing IPv6 packets of a device. You can permit or deny the entry of specific IPv6 packets to a network according to actual requirements to control access of IPv6 users to network resources.

Working Principle

Define a series of IPv6 access rules in the IPv6 ACL, and then apply the ACL in the incoming or outgoing direction of an interface. The device checks whether the incoming or outgoing IPv6 packets match the rules and accordingly forwards or blocks these packets.

To configure an IPv6 ACL, you must specify a unique name for this ACL.

Unlike the IP ACL, MAC extended ACL, and expert extended ACL, you can specify only a name but not an ID for the IPv6 ACL created.

Only one IP ACL, or one MAC extended ACL, or one expert extended ACL can be applied to the incoming or outgoing direction of an interface. Besides, one more IPv6 ACL can be applied.

Implicit "Deny All Traffic" Rule Statement

At the end of every IPv6 ACL is an implicit "deny all IPv6 traffic" rule statement. Therefore, if a packet does not match any rule, the packet will be denied.

For example:

```
ipv6 access-list ipv6_acl
```

```
10 permit ipv6 host 200::1 any
```

This ACL permits only IPv6 packets from the source host 200::1, and denies IPv6 packets from all other hosts. This is because the following statement exists at the end of this ACL: deny ipv6 any any.

Although the IPv6 ACL contains the implicit "deny all IPv6 traffic" rule statement by default, it does not filter ND packets.

Input Sequence of Rule Statements

Every new rule is added to the end of an ACL and in front of the default rule statement. The input sequence of statements in an ACL is very important. It determines the priority of each statement in the ACL. When determining whether to forward or block packets, a device compares packets with rule statements based on the sequence that rule statements are created. After locating a matched rule statement, the device does not check any other rule statement.

If a rule statement is created and permits all IPv6 traffic, all subsequent statements will not be checked.

For example:

```
ipv6 access-list ipv6_acl
10 permit ipv6 any any
20 deny ipv6 host 200::1 any
```

As the first rule statement permits all IPv6 packets, all IPv6 packets sent from the host 200::1 does not match the subsequent deny rule with the serial number of 20, and therefore will not be denied. After the device finds that packets match the first rule statement, it does not check the subsequent rule statements any more.

Related Configuration

Configuring an IPv6 ACL

By default, no IPv6 ACL is configured on a device.

Run the **ipv6 access-list *acl-name*** command in global configuration mode to create an IPv6 ACL and enter IPv6 ACL mode.

Adding ACEs to an IPv6 ACL

By default, a newly created IPv6 ACL contains an implicit ACE that denies all IPv6 packets. This ACE is hidden from users, but takes effect when the ACL is applied to an interface. That is, all IPv6 packets will be discarded. Therefore, if you want the device to receive or send some specific IPv6 packets, add some ACEs to the ACL.

Run the following command in IPv6 ACL mode to add an ACE:

[*sn*] {**permit** | **deny** } *protocol*{*src-ipv6-prefix/prefix-len* | **host***src-ipv6-addr* | **any**} {*dst-ipv6-pfix/pfix-len* | **host** *dst-ipv6-addr* |**any**} [**range***lower upper*] [**dscp***dscp*] [**flow-label** *flow-label*] [**fragment]** [**time-range***tm-rng-name*][**log**]

Applying an IPv6 ACL

By default, the IPv6 ACL is not applied to any interface, that is, the IPv6 ACL does not filter incoming or outgoing IPv6 packets of a device.

Run the **ipv6 traffic-filter *acl-name* { in| out }** command in interface/VXLAN configuration mode to apply an IPv6 ACL to a specified interface/VXLAN.

1.3.5 ACL80

ACL80 refers to the expert advanced ACL, and is also called custom ACL. It filters packets based on the first 80 bytes of every packet. Among these 80 bytes, the SMAC, DMAC, SIP, DIP, and ETYPE fields in a packet are mandatory, and you can specify the remaining 16 bytes.

Working Principle

A packet consists of a number of bytes. ACL80 allows you to match the specified 16 bytes by bit in the first 80 bytes of a packet. Any bit of a 16-byte field can be set to a value (**0** or **1**), indicating whether the bit is compared. When any byte is filtered, three factors are considered: content of the matching field, mask of the matching field, and the start position for matching. Bits of the matching field content are in one-to-one mapping relationship with bits of the matching field mask. The filtering rule specifies the value of the field to be filtered. The filtering field template specifies whether the corresponding field in the filtering rule should be filtered. (**1** indicates that

the bit specified in the filtering rule should be matched; **0** indicates that the bit specified in the filtering rule is not matched.) Therefore, when it is required to match a specific bit, you must set the corresponding bit to 1 in the filtering field template. For example, if the bit is set to **0** in the filtering field template, no bit is matched no matter which bit is specified in the filtering rule.

For example,

```
FS(config)#expert access-list advanced name
FS(config-exp-dacl)#permit 00d0f8123456 ffffffffffff 0
FS(config-exp-dacl)#deny 00d0f8654321 ffffffffffff 6
```

The custom ACL matches any byte of the first 80 bytes in a L2 data frame according to user' definition, and filters packets accordingly. To properly use a custom ACL, you must have an in-depth understanding about the structure of a L2 data frame. The following shows the first 64 bytes of a L3 data frame (every letter represents a hexadecimal number, and every two letters represent one byte):

AA AA AA AA AA AA BB BB BB BB BB BB CC CC DD DD

DD DD EE FF GG HH HH HH II II JJ KK LL LL MM MM

NN NN OO PP QQ QQ RR RR RR RR SS SS SS SS TT TT

UU UU VV VV VV VV WW WW WW WW XY ZZ aa aa bb bb

The following table describes the meaning and offset of each letter:

Letter	Meaning	Offset	Letter	Meaning	Offset
A	Destination MAC address	0	O	Time To Live (TTL) field	34
B	Source MAC address	6	P	Protocol number	35
C	VLAN tag field	12	Q	IP checksum	36
D	Data frame length	16	R	Source IP address	38
E	Destination service access point (DSAP) field	18	S	Destination IP address	42
F	Source service access point (SSAP) field	19	T	TCP source port	46
G	Cntl field	20	U	TCP destination port	48
H	Org Code field	21	V	Serial number	50
I	Encapsulated data type	24	W	Acknowledgment field	54
J	IP version number	26	XY	IP header length and reserved bit	58
K	TOS field	27	Z	Reserved bit and flags bit	59
L	IP packet length	28	a	Windows size field	60
M	ID	30	b	Miscellaneous	62
N	Flags field	32			

In the above table, the offset of each field is the offset of this field in the tagged 802.3 SNAP packet. In a custom ACL, you can use the rule mask and offset jointly to extract any byte from the first 80 bytes of a data frame, compare the byte with the rule customized in the ACL, and then filter matched data frames for further processing. Customized rules may be some fixed attributes of data. For example, to obtain all TCP packets, you can define the rule as "06", rule mask as "FF", and offset as "35". Then, the device can use the rule mask and offset jointly to extract the content of TCP protocol number field in a received data frame, and compare the extracted content with the rule to obtain all TCP packets.

Only switches support the ACL80.

The ACL80 supports filtering of the Ethernet, 803.3 SNAP, and 802.3 LLC packets. If the values of the fields from DSAP to cntl are set to AAAA03, the ACL is used to filter the 803.3 SNAP packets. If the values of the fields from DSAP to cntl are set to E0E003, the ACL is used to filter the 803.3 LLC packets. The value of the cntl field cannot be configured to filter Ethernet packets.

The ACL80 can be configured to compare packets with any of the 16 bytes. If the 16 bytes are already used, no ACE can be configured to compare packets with fields in any other bytes.

Related Configuration

Configuring an Expert Advanced ACL

By default, no expert advanced ACL is configured on a device.

Run the **expert access-list advanced *acl-name*** command in global configuration mode to create an expert advanced ACL and enter expert advanced ACL mode.

Adding ACEs to an Expert Advanced ACL

By default, a newly created expert advanced ACL contains an implicit ACE that denies all packets. This ACE is hidden from users, but takes effect when the ACL is applied to an interface. That is, all L2 packets will be discarded. Therefore, if you want the device to receive or send some specific L2 packets, add some ACEs to the ACL.

- Run the [*sn*] { **permit** | **deny** } *hex hex-mask offset* command in expert advanced ACL mode to add an ACE to the expert advanced ACL.

Applying an Expert Advanced ACL

By default, the expert advanced ACL is not applied to any interface, that is, the created expert advanced ACL does not filter incoming or outgoing packets of a device.

Run the **expert access-group *acl-name* { in| out }** command in interface configuration mode to apply an expert advanced ACL to a specified interface.

1.3.6 ACL Redirection

ACL redirection allows a device to analyze received packets and redirect the packets to a specified port for forwarding. To analyze specific incoming packets of a device, you can configure the ACL redirection function to redirect packets meeting rules to a specified port and capture packets on this port for analysis.

Working Principle

Bind different ACL policy to an interface and specify an output destination interface for each policy. When receiving packets on this interface, the device searches ACL policies bound to this interface one by one. If packets match criteria described in a certain policy, the device forwards packets on the destination interface specified by the policy, thus redirecting packets based on traffic.

Only switches support the ACL redirection function.

ACL redirection takes effect only in the incoming direction of an interface.

Related Configuration

Configuring an ACL

Before configuring ACL redirection, configure an ACL. For details about how to configure an ACL, see the earlier descriptions about ACL configuration.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Configuring ACL Redirection

By default, ACL redirection is not configured on a device.

Run the **redirect destinationinterface** *interface-name* **acl** {*acl-id* | ***acl-name*** } **in** command in interface configuration mode to configure ACL redirection.

You can configure the ACL redirection function only on an Ethernet interface, AP, or SVI.

1.3.7 Global Security ACL

To meet the requirements of security deployment, the port-based ACL is often configured to filter out virus packets and obtain packets with certain characteristics, for example, packets that attack the TCP port. Various virus packets exist in a global network environment, and the identification features of virus packets under each port are identical or similar. Therefore, an ACL is generally created. After the deny ACE for matching virus signatures is added to the ACL, the port-based ACL is applied to each port on the switch to filter out virus packets.

For two reasons, it is not convenient to use the port-based ACLs in antivirus scenarios such as virus filtering. The first reason is that the port-based ACL must be configured on every port, which results in repeated configuration, poor operation performance, and over-consumption of ACL resources. The second reason is that the access control function of the ACL is weakened. As the port-based ACL is used for virus filtering, basic functions of the ACL, such as route update restriction and network access restriction, cannot be used properly. The global security ACL can be used for global antivirus deployment and defense without affecting the port-based ACL. By running only one command, you can make the global security ACL takes effect on all L2 interfaces. In contrast, the port-based ACL must be configured on every interface.

Working Principle

The global security ACL takes effect on all L2 interfaces. When both the global security ACL and the port-based ACL are configured, both take effect. Packets that match the global security ACL are directly filtered out as virus packets. Packets that do not match the global security ACL are still controlled by the port-based ACL. You can disable the global security ACL on some ports so that these ports are not controlled by the global security ACL.

The global security ACL is mainly used for virus filtering. Therefore, in an ACL associated with the global security ACL, only the deny ACEs take effect, and the permit ACEs do not take effect.

Unlike the secure ACL applied to a port, the global security ACL does not contain the default "deny all traffic" ACE, that is, all packets that do not match the ACL are permitted.

A global secure ACL can take effect either on a L2 port or a routed port. That is, it takes effect on all the following types of ports: access port, trunk port, hibird port, routed port, and AP (L2 or L3). The global secure ACL does not take effect on an SVI.

You can disable the global security ACL on an individual physical port or AP, but not on a member port of an AP.

The global secure ACL supports only the associated IP standard ACL, IP extended ACL, MAC extended ACL and Expert extended ACL.

Related Configuration

Configuring an ACL

Before configuring the global security ACL, configure an ACL. For details about how to configure an ACL, see the earlier descriptions about ACL configuration.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL.

Configuring a Global Security ACL

By default, no global security ACL is configured on a device.

Run the **ip access-group** ***acl-id*** { **in** | **out** } command in global configuration mode to enable the global security ACL.

Configuring an Exclusive Interface of the Global Security ACL

By default, no exclusive interface is configured for the global security ACL on a device.

Run the **no global ip access-group** command in interface configuration to disable the global security ACL on a specified interface.

1.3.8 Security Channel

In some application scenarios, packets meeting some characteristics may need to bypass the checks of access control applications. For example, before DOT1X authentication, users are allowed to log in to a specified website to download the DOT1X authentication client. The security channel can be used for this purpose. When the security channel configuration command is executed to apply a secure ACL globally or to an interface or VXLAN, this ACL becomes a security channel.

Working Principle

The security channel is also an ACL, and can be configured globally or for a specified interface or VXLAN. When arriving at an interface, packets are check on the security channel. If meeting the matching conditions of the security channel, packets directly enters a switch without undergoing the access control, such as port security, Web authentication, 802.1x, and IP+MAC binding check. A globally applied security channel takes effect on all interfaces except exclusive interfaces.

The deny ACEs in an ACL that is applied to a security channel do not take effect. In addition, this ACL does not contain an implicit "deny all traffic" rule statement at the end of the ACL. If packets do not meet matching conditions of the security channel, they are checked according to the access control rules in compliance with the relevant process.

You can configure up to eight exclusive interfaces for the global security channel. In addition, you cannot configure interface-based security channel on these exclusive interfaces.

If a security channel is applied to an interface while a global security channel exists, this global security channel does not take effect on this interface.

If both port-based migratable authentication mode and security channel are applied to an interface, the security channel does not take effect.

An IPv6 ACL cannot be configured as a security channel.

Only switches support the security channel.

Related Configuration

Configuring an ACL

Before configuring the security channel, configure an ACL. For details about how to configure an ACL, see the earlier descriptions about ACL configuration.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, or expert extended ACL.

Configuring a Security Channel on an Interface

By default, no security channel is configured on an interface of a device.

Run the **security access-group** {*acl-id* | ***acl-name*** } command in interface configuration mode to configure the security channel on an interface.

Configuring a Global Security Channel

By default, no global security channel is configured on a device.

Run the **security global access-group** {*acl-id* | ***acl-name*** } command in global configuration mode to configure a global security channel.

Configuring an Exclusive Interface for the Global Security Channel

By default, no exclusive interface is configured for the global security channel on a device.

Run the **security uplink enable** command in interface configuration mode to configure a specified interface as the exclusive interface of the global security channel.

1.3.9 SVI Router ACL

By default, an ACL that is applied to an SVI also takes effect on L2 packets forwarded within a VLAN and L3 packets forwarded between VLANs. Consequently, users in the same VLAN may fail to communicate with each other. Therefore, a switchover method is provided so that the ACL that is applied to an SVI takes effect only on routing packets between VLANs.

Working Principle

By default, the SVI router ACL function is disabled, and an SVI ACL takes effect on L3 packets forwarded between VLANs and L2 packets forwarded within a VLAN. After the SVI router ACL function is enabled, the SVI ACL takes effect only on L3 packets forwarded between VLANs.

Only switches support the SVI router ACL.

Related Configuration

Configuring an ACL

Before configuring the SVI router ACL, configure and apply an ACL. For details about how to configure an ACL, see the earlier descriptions about ACL configuration.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Applying an ACL

For details about how to apply an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL. Apply the ACL in SVI configuration mode.

Configuring the SVI Router ACL

Run the **svi router-acls enable** command in global configuration mode to enable the SVI router ACL so that the ACL that is applied to an SVI takes effect only on packets forwarded at L3, and not on packets forwarded at L2 within a VLAN.

1.3.10 ACL Logging

ACL logging is used to monitor the running status of ACEs in an ACL and provide essential information for routine network maintenance and optimization.

Working Principle

To better learn the running status of ACLs on a device, you can determine whether to specify the ACL logging option as required when adding ACEs. If this option is specified, logs are output when packets matching ACEs are found. ACL logs are displayed based on ACEs. That is, the device periodically displays ACEs with matched packets and the number of matched packets. An example of the log is as follows:

```
*Sep    9 16:23:06: %ACL-6-MATCH: ACL 100 ACE 10 permit icmp any any, match 78 packets.
```

To control the amount of logs and output frequency, you can configure the log update interval.

An ACE containing the ACL logging option consumes more hardware resources. If all configured ACEs contain this option, the ACE capacity of a device will be reduced by half.

By default, the log update interval is 0, that is, no log is output. After the ACL logging option is specified in an ACE, you need to configure the log update interval to output related logs; otherwise, logs are not output.

For an ACE containing the ACL logging option, if no packet is matched in the specified interval, no packet matching log related to this ACE will be output. If matched packets are found in the specified interval, packet matching logs related to this ACE will be output when the interval expires. The number of matched packets is the total number of packets that match the ACE during the specified interval, that is, the period from the previous log output to the current log output.

Only switches support the ACL logging function.

You can configure the ACL logging option only for an IP ACL or an IPv6 ACL.

Related Configuration

Configuring an ACL

Configure an ACL before configuring ACEs containing the ACL logging option. For details about how to configure an ACL, see the earlier descriptions about ACL configuration.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL and IPv6 ACL. Note that the ACL logging option must be configured.

Configuring the Log Update Interval

Run the {**ip | ipv6**} **access-list log-update inerval** *time* command in the configuration mode to configure the interval at which the ACL logs are output.

Applying an ACL

For details about how to apply an ACL, see the earlier descriptions about the IP ACL and IPv6 ACL.

1.3.11 Packet Matching Counters

In addition to ACL logs, packet matching counters provide another choice for routine network maintenance and optimization.

Working Principle

To implement network management, users may want to know whether an ACE has any matched packets and how many packets are matched.ACLs provide the ACE-based packet matching counters. You can enable or disable packet matching counters for all ACEs in an ACL. When a packet matches the ACE, the corresponding counter increments by 1. You can run the **clear counters access-list** [*acl-id* | *acl-name*] command to reset counters of all ACEs in an ACL for a new round of statistics.

Enabling ACL counters requires more hardware entries. In an extreme case, this will reduce by half the number of ACEs that can be configured on a device.

You can enable packet matching counters on an IP ACL, MAC ACL, expert ACL, or IPv6 ACL.

Only switches support the ACL packet matching counters.

Related Configuration

Configuring an ACL

Configure an ACL before configuring ACEs containing the ACL logging option. For details about how to configure an ACL, see the earlier descriptions about ACL configuration.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL and IPv6 ACL. Note that the ACL logging option must be configured.

Enabling Packet Matching Counters

To enable packet matching counters on an IP ACL, MAC ACL, or expert ACL, run the {**mac | expert** | **ip**} **access-list counter** { *acl-id* | *acl-name* } command in global configuration mode.

To enable packet matching counters on an IPv6 ACL, run the **ipv6 access-list counter** *acl-name* command in global configuration mode.

Applying an ACL

For details about how to apply an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Clearing Packet Matching Counters

Run the **clear countersaccess-list** [*acl-id* | *acl-name*] command in privileged EXEC mode to reset packet matching counters.

1.3.12 Fragmented Packet Matching Mode

In fragmented packet matching mode, an ACL can implement more refined control on fragmented packets.

Working Principle

IP packets may be fragmented when transmitted on the network. When fragmentation occurs, only the first fragment of the packet contains the L4 information, such as the TCP/UDP port number, ICMP type, and ICMP code, and other fragmented packets do not contain the L4 information. By default, if an ACE contains the fragment flag, fragmented packets except the first fragments are filtered. If an ACE does not contain the fragment flag, all fragmented packets (including the first fragments) are filtered. In addition to this default fragmented packet matching mode, a new fragmented packet matching mode is provided. You can switch between the two fragmented packet matching modes as required on a specified ACL. In the new fragmented packet matching mode, if an ACE does not contain the fragment flag and packets are fragmented, the first fragments are compared with all the matching fields (including L3 and L4 information) defined in the ACE, and other fragmented packets are compared with only the non-L4 information defined in the ACE.

In the new fragmented packet matching mode, if an ACE does not contain the fragment flag and the action is Permit, this type of ACE occupies more hardware entries. In an extreme case, this will reduce by half the number of hardware entries. If Established is configured for filter the TCP flag in an ACE, more hardware entries will be occupied.

The ACL will be temporarily ineffective during switchover of the fragmented packet matching mode.

In the new fragmented packet matching mode, if an ACE does not contain the fragment flag, the L4 information of packets needs to be compared, and the action is Permit, the ACE checks the L3 and L4 information of the first fragments of packets, and checks only the L3 information of other fragmented packets. If the action is Deny, the ACE checks only the first fragments of packets, and ignores other fragmented packets.

In the new fragmented packet matching mode, if an ACE contains the fragment flag, the ACE checks only fragmented packets but not the first fragments of packets no matter whether the action in the ACE is Permit or Deny.

Only the IP extended ACL and the expert extended ACL support switching between the two fragmented packet matching modes.

Only switches support filtering of fragmented packets.

Related Configuration

Configuring an ACL

For details about how to configure an ACL, see the earlier descriptions about the IP ACL and expert extended ACL.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL and expert extended ACL. Note that the fragment option must be added.

Switching the Fragmented Packet Matching Mode

Run the [**no**] {**ip | expert**} **access-list new-fragment-mode** { *acl-id* | *acl-name* } command in global configuration mode to switch the fragmented packet matching mode.

Applying an ACL

For details about how to apply an ACL, see the earlier descriptions about the IP ACL and expert extended ACL.

1.4 Configuration

Configuration Item	Description and Command	
Configuring an IP ACL	(Optional) It is used to filter IPv4 packets.	
	ip access-list standard	Configures a standard IP ACL.
	ip access-list extended	Configures an extended IP ACL.
	permit host any time-range log	Adds a permit ACE to a standard IP ACL.

Configuration Item	Description and Command	
	deny host any time-range log	Adds a deny ACE to a standard IP ACL.
	permit host any host any tos dscp precedence fragment time-range log	Adds a permit ACE to an extended IP ACL.
	deny host any host any tos dscp precedence fragment time-range log	Adds a deny ACE to an extended IP ACL.
	ip access-group in out	Applies a standard or an extended IP ACL.
Configuring an MAC Extended ACL	(Optional) It is used to filter L2 packets.	
	mac access-list extended	Configures an MAC extended ACL.
	permit any host any host cos inner time-range	Adds a permit ACE to an MAC extended ACL.
	deny any host any host cos inner time-range	Adds a deny ACE to an MAC extended ACL.
	mac access-group in out	Applies an MAC extended ACL.
Configuring an Expert Extended ACL	(Optional) It is used to filter L2 and L3 packets.	
	expert access-list extended	Configures an expert extended ACL.
	permit cos inner VID inner host any host any host any host any precedence tos fragment range time-range	Adds a permit ACE to an expert extended ACL.
	deny cos inner VID inner host any host any host any host any precedence tos fragment range time-range	Adds a deny ACE to an expert extended ACL.
	expert access-group in out	Applies an expert extended ACL.
Configuring an IPv6 Extended ACL	(Optional) It is used to filter IPv6 packets.	
	ipv6 access-list	Configures an IPv6 ACL.
	permit host any host any range dscp flow-label fragment time-range log	Adds a permit ACE to an IPv6 ACL.
	deny host any host any range dscp flow-label fragment time-rangelog	Adds a deny ACE to an IPv6 ACL.
	ipv6 traffic-filter in out	Applies an IPv6 ACL.
Configuring an ACL80	(Optional) It is used to customize the fields for filter L2 and L3 packets.	
	expert access-list advanced	Configures an expert advanced ACL.
	permit	Adds a permit ACE to an expert advanced ACL.
	deny	Adds a deny ACE to an expert advanced ACL.
	expert access-group in out	Applies an expert advanced ACL
Configuring ACL Redirection	(Optional) It is used to redirect packets meeting the rules to a specified interface.	
	redirect destination interface acl in	Configures ACL redirection.
Configuring a Global Security ACL	(Optional) It is used to make an ACL take effect globally.	

Configuration Item	Description and Command	
	ip access-group in out	Applies a global security ACL in global configuration mode.
	no global ip access-group	Configures an interface as the exclusive interface of the global security ACL in interface configuration mode.
Configuring a Security Channel	(Optional) It is used to enable packets meeting some characteristics to bypass the checks of access control applications, such as the DOT1X and Web authentication.	
	security access-group	Enables the security channel in interface configuration mode.
	security global access-group	Enables the security channel in global configuration mode.
	security uplink enable	Configures an interface as the exclusive interface of the global security channel in interface configuration mode.
Configuring Comments for ACLs	(Optional) It is used to configure comments for an ACL or ACE so that users can easily identify the functions of the ACL or ACE.	
	list-remark	Configures a comment for an ACL in ACL configuration mode.
	access-list list-remark	Configures a comment for an ACL in global configuration mode.
	remark	Configures a comment for an ACE in ACL configuration mode.

1.4.1 Configuring an IP ACL

Configuration Effect

Configure and apply an IP ACL to an interface/VXLAN to control all incoming and outgoing IPv4 packets of this interface/VXLAN. You can permit or deny the entry of specific IPv4 packets to a network to control access of IP users to network resources.

Notes

N/A

Configuration Steps

Configuring an IP ACL

- (Mandatory) Configure an IP ACL if you want to control access of IPv4 users to network resources.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The IP ACL takes effect only on the local device, and does not affect other devices on the network.

Adding ACEs to an IP ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, all incoming IPv4 packets of the device are denied by default.

Applying an IP ACL

- (Mandatory) Apply an IP ACL to a specified interface/VXLAN if you want this ACL take effect.

- You can apply an IP ACL on a specified interface/VXLAN of an access, an aggregate, or a core device based on the distribution of users.

Verification

- Use the following methods to verify the configuration effects of the IP ACL:
- Run the **ping** command to verify that the IP ACL takes effect on the specified interface. For example, if an IP ACL is configured to prohibit a host with a specified IP address or hosts in a specified IP address range from accessing the network, run the **ping** command to verify that the host(s) cannot be successfully pinged.
- Access related network resources to verify that the IP ACL takes effect on the specified interface. For example, access the Internet or access the FTP resources on the network through FTP.

Related Commands

Configuring an IP ACL

Command	**ip access-list** { **standard | extended** } {*acl-name* | *acl-id* }
Parameter Description	**standard**: Indicates that a standard IP ACL is created. **extended**: Indicates that an extended IP ACL is created. *acl-name*: Indicates the name of a standard or an extended IP ACL. If this option is configured, a named ACL is created. The name is a string of 1 to 99 characters. The ACL name cannot start with numbers (0–9), "in", or "out". *acl-id*: Indicates the ID that uniquely identifies a standard or extended IP ACL. If this option is configured, a numbered ACL is created. If a standard IP ACL is created, the value range of *acl-id* is 1–99 and 1300–1999. If an extended IP ACL is created, the value range of *acl-id* is 100–199 and 2000–2699.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure a standard or an extended IP ACL and enter standard or extended IP ACL configuration mode. If you want to control access of users to network resources by checking the source IP address of each packet, configure a standard IP ACL. If you want to control access of users to network resources by checking the source or destination IP address, protocol number, and TCP/UDP source or destination port, configure an extended IP ACL.

Adding ACEs to an IP ACL

- Add ACEs to a standard IP ACL.

Use either of the following methods to add ACEs to a standard IP ACL:

Command	[*sn*] { **permit** | **deny** } {**host** *source* | **any** | *source source-wildcard* } [**time-range** *time-range-name*] [**log**]
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. **host** *source*: Indicates that IP packets sent from a host with the specified source IP address are filtered. **any**: Indicates that IP packets sent from any host are filtered. *source source-wildcard:* Indicates that IP packets sent from hosts in the specified IP network segment are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within

	this time range. For details about the time range, see the configuration manual of the time range. **log**: Indicates that logs will be periodically output if packets matching the ACEs are found. For details about logs, see "ACL Logging" in this document.
Command Mode	Standard IP ACL configuration mode
Usage Guide	Run this command to add ACEs in standard IP ACL configuration mode. The ACL can be a named or numbered ACL.

Command	**access-list** *acl-id* { **permit** | **deny** } {**host** *source* | **any** | *source source-wildcard* } [**time-range** *tm-rng-name*] [**log**]
Parameter Description	*acl-id*: Indicates the ID of a numbered ACL. It uniquely identifies an ACL. The value range of *acl-id* is 100–199 and 1300–1999. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. **host** *source*: Indicates that IP packets sent from a host with the specified source IP address are filtered. **any**: Indicates that IP packets sent from any host are filtered. *source source-wildcard:* Indicates that IP packets sent from hosts in the specified IP network segment are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range. **log**: Indicates that logs will be periodically output if packets matching the ACEs are found. For details about logs, see "ACL Logging" in this document.
Command Mode	Standard IP ACL configuration mode
Usage Guide	Run this command to add ACEs to a numbered IP ACL in global configuration mode.It cannot be used to add ACEs to a named IP ACL.

- Add ACEs to an extended IP ACL.

Use either of the following methods to add ACEs to an extended IP ACL:

Command	[*sn*] { **permit** | **deny** } *protocol* {**host** *source* | **any** | *source source-wildcard* } {**host** *destination* | **any** | *destination destination-wildcard* } [[**precedence** *precedence* [**tos** *tos*]] | **dscp** *dscp*] [**fragment**] [**time-range** *time-range-name*] [**log**]
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *protocol*: Indicates the IP protocol number. The value ranges from 0 to 255. To facilitate the use, the system provides frequently-used abbreviations to replace the specific IP protocol numbers, including eigrp, gre, icmp, igmp, ip, ipinip, nos, ospf, tcp, and udp. **host** *source*: Indicates that IP packets sent from a host with the specified source IP address are filtered. *source source-wildcard:* Indicates that IP packets sent from hosts in the specified IP network segment are filtered. **host** *destination:* Indicates that IP packets sent to a host with the specified destination IP address are filtered. If the

	any keyword is configured, IP packets sent to any host are filtered. *destination destination-wildcard:* Indicates that IP packets sent to hosts in a specified IP network segment are filtered. **any**: Indicates that IP packets sent to or from any host are filtered. **precedence** *precedence:* Indicates that IP packets with the specified precedence field in the header are filtered. **tos** *tos:* Indicates that IP packets with the specified the type of service (TOS) field in the header are filtered. **dscp** *dscp*: Indicates that IP packets with the specified the dcsp field in the header are filtered. **fragment:** Indicates that only fragmented IP packets except the first fragments are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range. **log**: Indicates that logs will be periodically output if packets matching the ACEs are found. For details about logs, see "ACL Logging" in this document.
Command Mode	Extended IP ACL configuration mode
Usage Guide	Run this command to add ACEs in extended IP ACL configuration mode. The ACL can be a named or numbered ACL.

Command	**access-list** *acl-id* { **permit** | **deny** } *protocol* {**host** *source* | **any** | *source source-wildcard* } {**host** *destination* | **any** | *destination destination-wildcard* } [[**precedence** *precedence* [**tos** *tos*]] | **dscp** *dscp*] [**fragment**] [**time-range** *time-range-name*] [**log**]
Parameter Description	*acl-id*: Indicates the ID of a numbered ACL. It uniquely identifies an ACL. The value range of *acl-id* is 100–199 and 2000–1999. *sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *protocol*: Indicates the IP protocol number. The value ranges from 0 to 255. To facilitate the use, the system provides frequently-used abbreviations to replace the specific IP protocol numbers, including eigrp, gre, icmp, igmp, ip, ipinip, nos, ospf, tcp, and udp. **host** *source*: Indicates that IP packets sent from a host with the specified source IP address are filtered. *source source-wildcard:* Indicates that IP packets sent from hosts in the specified IP network segment are filtered. **host** *destination:* Indicates that IP packets sent to a host with the specified destination IP address are filtered. If the **any** keyword is configured, IP packets sent to any host are filtered. *destination destination-wildcard:* Indicates that IP packets sent to hosts in a specified IP network segment are filtered. **any**: Indicates that IP packets sent to or from any host are filtered. **precedence** *precedence:* Indicates that IP packets with the specified precedence field in the header are filtered. **tos** *tos:* Indicates that IP packets with the specified the type of service (TOS) field in the header are filtered. **dscp** *dscp*: Indicates that IP packets with the specified the dcsp field in the header are filtered. **fragment: Indicates that only fragmented IP packets except the first fragments are filtered.** **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range.

	log: Indicates that logs will be periodically output if packets matching the ACEs are found. For details about logs, see "ACL Logging" in this document.
Command Mode	Extended IP ACL configuration mode
Usage Guide	Run this command to add ACEs to a numbered IP ACL in extended IP ACL configuration mode.It cannot be used to add ACEs to a named extended IP ACL.

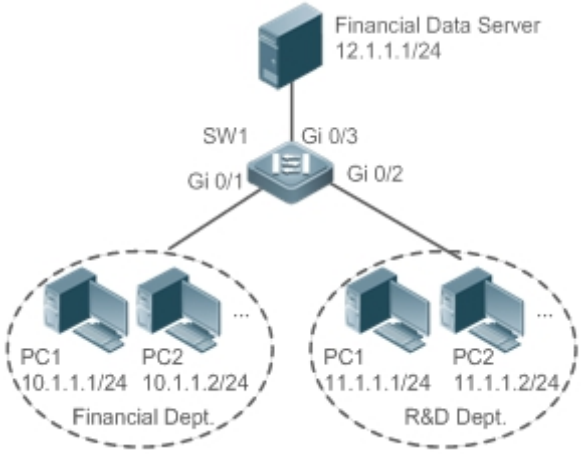
Applying an IP ACL

Command	**ip access-group** { *acl-id* **|** *acl-name* } { **in** | **out** }
Parameter Description	*acl-id*: Indicates that a numbered standard or extended IP ACL will be applied to the interface. *acl-name*: Indicates that a named standard or extended IP ACL will be applied to the interface. **in**: Indicates that this ACL controls incoming IP packets of the interface. **out**: Indicates that this ACL controls outgoing IP packets of the interface. **reflect:** Indicates that the reflexive ACL is enabled.
Command Mode	Interface/VXLAN configuration mode
Usage Guide	This command makes an IP ACL take effect on the incoming or outgoing packets of a specified interface/VXLAN.

Configuration Example

The following configuration example describes only ACL-related configurations.

Configuring an IP ACL to Prohibit Departments Except the Financial Department from Accessing the Financial Data Server

Scenario **Figure 1- 3**	
Configuration Steps	• Configure an IP ACL. • Add ACEs to the IP ACL. • Apply the IP ACL to the outgoing direction of the interface connecting the financial data server.
SW1	sw1(config)#ip access-list standard 1 sw1(config-std-nacl)#permit 10.1.1.0 0.0.0.255 sw1(config-std-nacl)#deny 11.1.1.1 0.0.0.255 sw1(config-std-nacl)#exit sw1(config)#int gigabitEthernet 0/3 sw1(config-if-GigabitEthernet 0/3)#ip access-group 1 out

Verification	● On a PC of the R&D department, ping the financial data server. Verify that the ping operation fails. ● On a PC of the financial department, ping the financial data server. Verify that the ping operation succeeds.
SW1	sw1(config)#show access-lists ip access-list standard 1 10 permit 10.1.1.0 0.0.0.255 20 deny 11.1.1.0 0.0.0.255 sw1(config)#show access-group ip access-group 1 out Applied On interface GigabitEthernet 0/3

1.4.2 Configuring an MAC Extended ACL

Configuration Effect

Configure and apply an MAC extended ACL to an interface/VXLAN to control all incoming and outgoing IPv4 packets of this interface/VXLAN. You can permit or deny the entry of specific L2 packets to a network to control access of users to network resources based on L2 packets.

Notes

N/A

Configuration Steps

Configuring an MAC Extended ACL

- (Mandatory) Configure an MAC extended ACL if you want to control users' access to network resources based on the L2 packet header, for example, the MAC address of each user's PC.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The MAC extended ACL takes effect only on the local device, and does not affect other devices on the network.

Adding ACEs to an MAC Extended ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, all incoming L2 Ethernet packets of the device are denied by default.

Applying an MAC extended ACL

- (Mandatory) Apply an MAC extended ACL to a specified interface if you want this ACL take effect.
- You can apply an MAC extended ACL on a specified interface of an access, an aggregate, or a core device based on the distribution of users.

Verification

- Use the following methods to verify the configuration effects of the MAC extended ACL:
- If an MAC extended ACL is configured to permit or deny some IP packets, run the **ping** command to check whether ACEs of this ACL takes effect on the specified interface. For example, an MAC extended ACL is configured to prevent a device interface from receiving IP packets (Ethernet type is 0x0800), run the **ping** command for verification.
- If an MAC extended ACL is configured to permit or deny some non-IP packets (e.g. ARP packets), also run the **ping** command to check whether ACEs of this ACL takes effect on the specified interface. For example, to filter out ARP packets, run the **ping** command for verification.

- You can also construct L2 packets meeting some specified characteristics to check whether the MAC extended ACL takes effect. Typically, prepare two PCs, construct and send L2 packets on one PC, enable packet capturing on another PC, and check whether packets are forwarded as expected (forwarded or blocked) according to the action specified in the ACEs.

Related Commands

Configuring an MAC Extended ACL

Command	**mac access-list extended** {*acl-name* | *acl-id* }
Parameter Description	*acl-name*: Indicates the name of an MAC extended ACL. If this option is configured, a named ACL is created. The name is a string of 1 to 99 characters. The ACL name cannot start with numbers (0–9), "in", or "out". *acl-id*: Indicates the ID that uniquely identifies an MAC extended ACL. If this option is configured, a numbered ACL is created. The value range of *acl-id* is 700–799.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure an MAC extended ACL and enter MAC extended ACL configuration mode. You can configure an MAC extended ACL to control users' access to network resources by checking the L2 information of Ethernet packets.

Adding ACEs to an MAC Extended ACL

Use either of the following methods to add ACEs to an MAC extended ACL:

- Add ACEs in MAC extended ACL configuration mode.

Command	[*sn*] { **permit** | **deny** } {**any** | **host** *src-mac-addr* | *src-mac-addr mask*} {**any** | **host** *dst-mac-addr* | *dst-mac-addr mask* } [*ethernet-type*] [**cos** *cos* [**inner** *cos*]] [**time-range** *tm-rng-name*]
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. **any**: Indicates that L2 packets sent from any host are filtered. **host** *src-mac-addr*: Indicates that IP packets sent from a host with the specified source MAC address are filtered. *src-mac-addr mask*: Indicates that the source MAC address is reversed. **any**: Indicates that L2 packets sent to any host are filtered. **host** *dst-mac-addr*: Indicates that IP packets sent to a host with the specified destination MAC address are filtered. *dst-mac-addr mask*: Indicates that the destination MAC address is reversed. *ethernet-type*: Indicates that L2 packets of the specified Ethernet type are filtered. **cos** *cos*: Indicates that L2 packets with the specified class of service (cos) field in the outer tag are filtered. **inner** *cos*: Indicates that L2 packets with the specified cos field in the inner tag are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range.
Command Mode	MAC extended ACL configuration mode
Usage Guide	Run this command to add ACEs in MAC extended ACL configuration mode. The ACL can be a named or numbered

	ACL.

- Add ACEs to an MAC extended ACL in global configuration mode.

Command	**access-list** *acl-id* { **permit** | **deny** } {**any** | **host** *src-mac-addr* | *src-mac-addr mask* } {**any** | **host** *dst-mac-addr* | *dst-mac-addr mask* } [*ethernet-type*] [**cos** *cos* [**inner** *cos*]] [**time-range** *tm-rng-name*]
Parameter Description	*acl-id*: Indicates the ID of a numbered ACL. It uniquely identifies an ACL. The value range of *acl-id* is 700–799. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. **host** *src-mac-addr*: Indicates that IP packets sent from a host with the specified source MAC address are filtered. *src-mac-addr mask*: Indicates that the source MAC address is reversed. **any**: Indicates that L2 packets sent to any host are filtered. **host** *dst-mac-addr*: Indicates that IP packets sent to a host with the specified destination MAC address are filtered. *dst-mac-addr mask*: Indicates that the destination MAC address is reversed. *ethernet-type*: Indicates that L2 packets of the specified Ethernet type are filtered. **cos** *cos*: Indicates that L2 packets with the specified cos field in the outer tag are filtered. **inner** *cos*: Indicates that L2 packets with the specified cos field in the inner tag are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range.
Command Mode	Global configuration mode
Usage Guide	Run this command to add ACEs to a numbered MAC extended ACL in global configuration mode. It cannot be used to add ACEs to a named MAC extended ACL.

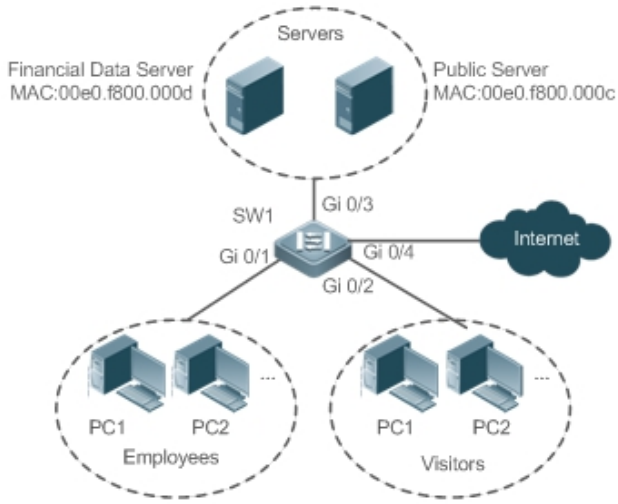
Applying an MAC Extended ACL

Command	**mac access-group** { *acl-id* | *acl-name* } { **in** | **out** }
Parameter Description	*acl-id*: Indicates that a numbered MAC extended IP ACL will be applied to the interface. *acl-name*: Indicates that a named MAC extended IP ACL will be applied to the interface. **in**: Indicates that this ACL controls incoming L2 packets of the interface. **out**: Indicates that this ACL controls outgoing L2 packets of the interface.
Command Mode	Interface configuration mode
Usage Guide	This command makes an MAC extended ACL take effect on the incoming or outgoing packets of a specified interface.

Configuration Example

The following configuration example describes only ACL-related configurations.

Configuring an MAC Extended ACL to Restrict Resources Accessible by Visitors

Scenario **Figure 1- 4**	
Configuration Steps	● Configure an MAC extended ACL. ● Add ACEs to the MAC extended ACL. ● Apply the MAC extended ACL to the outgoing direction of the interface connected to the visitor area so that visitors are allowed to access Internet and the public server of the company, but prohibited from accessing the financial data server of the company. That is, visitors cannot access the server with the MAC address 00e0.f800.000d.
SW1	sw1(config)#mac access-list extended 700 sw1(config-mac-nacl)#deny any host 00e0.f800.000d sw1(config-mac-nacl)#pemit any any sw1(config-mac-nacl)#exit sw1(config)#int gigabitEthernet 0/2 sw1(config-if-GigabitEthernet 0/2)#mac access-group 700 in
Verification	● On a visitor's PC, ping the financial data server. Verify that the ping operation fails. ● On a visitor's PC, ping the public resource server. Verify that the ping operation succeeds. ● On a visitor's PC, access the Internet, for example, visit the Baidu website. Verify that the webpage can be opened.
SW1	sw1(config)#show access-lists mac access-list extended 700 10 deny any host 00e0.f800.000d etype-any 20 permit any any etype-any sw1(config)#show access-group mac access-group 700 in Applied On interface GigabitEthernet 0/2

1.4.3 Configuring an Expert Extended ACL

Configuration Effect

Configure and apply an expert extended ACL to an interface/VXLAN to control incoming and outgoing packets of the interface/VXLAN based on the L2 and L3 information, and allow or prohibit the entry of specific packets to the network. In addition, you can configure an expert extended ACL to control all L2 packets based on the VLAN to permit or deny the access of users in some network segments to

network resources. Generally, you can use an expert extended ACL if you want to incorporate ACEs of the IP ACL and MAC extended ACL into one ACL.

Configuration Steps

Configuring an Expert Extended ACL

- (Mandatory) Configure an expert extended ACL if you want to control users' access to network resources based on the L2 packet header, for example, the VLAN ID.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The expert extended ACL takes effect only on the local device, and does not affect other devices on the network.

Adding ACEs to an Expert Extended ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, all incoming packets of the device are denied by default.

Applying an Expert Extended ACL

- (Mandatory) Apply an expert extended ACL to a specified interface if you want this ACL take effect.
- You can apply an expert extended ACL in the incoming or outgoing direction of a specified interface of an access, an aggregate, or a core device based on the distribution of users.

Verification

- Use the following methods to verify the configuration effects of the expert extended ACL:
- If IP-based access rules are configured in an expert extended ACL to permit or deny some IP packets, run the **ping** command to verify whether these rules take effect.
- If MAC-based access rules are configured in an expert extended ACL to permit or deny some L2 packets (e.g. ARP packets), also run the **ping** command to check whether ACEs of this ACL takes effect on the specified interface. For example, to filter out ARP packets, run the **ping** command for verification.
- If VLAN ID-based access rules are configured in an expert extended ACL to permit or deny some L2 packets in some network segments (e.g., to prevent communication between VLAN 1 users and VLAN 2 users), ping PCs of VLAN 2 on a PC of VLAN 1. If the ping operation fails, the rules take effect.

Related Commands

Configuring an Expert Extended ACL

Command	**expert access-list extended** {*acl-name* | *acl-id* }
Parameter Description	*acl-name*: Indicates the name of an expert extended ACL. If this option is configured, a named ACL is created. The name is a string of 1 to 99 characters. The ACL name cannot start with numbers (0–9), "in", or "out". *acl-id*: Indicates the ID of an expert extended ACL. If this option is configured, a numbered ACL is created. The value range of *acl-id* is 2700-2899.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure an expert extended ACL and enter expert extended ACL configuration mode.

Adding ACEs to an Expert Extended ACL

Use either of the following methods to add ACEs to an expert extended ACL:

- Add ACEs in expert extended ACL configuration mode.

Command	[*sn*] { **permit** | **deny** } [*protocol* | [*ethernet-type*] [**cos** [*out*] [**inner** *in*]]] [[**VID** [*out*] [**inner** *in*]]] { *source source-wildcard* | **host** *source* | **any** } { **host** *source-mac-address* | **any** } { *destination destination-wildcard* | **host** *destination* | **any** } { **host** *destination-mac-address* | **any** } [**precedence** *precedence*] [**tos** *tos*] [**fragment**] [**range**

	lower upper] [**time-range** *time-range-name*]]
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *protocol*: Indicates the IP protocol number. The value ranges from 0 to 255. To facilitate the use, the system provides frequently-used abbreviations to replace the specific IP protocol numbers, including eigrp, gre, icmp, igmp, ip, ipinip, nos, ospf, tcp, and udp. *ethernet-type*: Indicates that L2 packets of the specified Ethernet type are filtered. **cos** *out*: Indicates that L2 packets with the specified cos field in the outer tag are filtered. **cos inner** *in*: Indicates that L2 packets with the specified cos field in the inner tag are filtered. **VID** *out*: Indicates that L2 packets with the specified VLAN ID field in the outer tag are filtered. **VID inner** *in*: Indicates that L2 packets with the specified VLAN ID field in the inner tag are filtered. *source source-wildcard:* Indicates that IP packets sent from hosts in the specified IP network segment are filtered. **host** *source*: Indicates that IP packets sent from a host with the specified source IP address are filtered. **any**: Indicates that IP packets sent from any host are filtered. **host** *source-mac-address*: Indicates that IP packets sent from a host with the specified source MAC address are filtered. **any**: Indicates that L2 packets sent to any host are filtered. *destination destination-wildcard:* Indicates that IP packets sent to hosts in a specified IP network segment are filtered. **host** *destination*: Indicates that IP packets sent to a host with the specified destination IP address are filtered. **any**: Indicates that IP packets sent to any host are filtered. **host** *destination-mac-address*: Indicates that IP packets sent to a host with the specified destination MAC address are filtered. **any**: Indicates that L2 packets sent to any host are filtered. **precedence** *precedence:* Indicates that IP packets with the specified precedence field in the header are filtered. **tos** *tos:* Indicates that IP packets with the specified the TOS field in the header are filtered. **dscp** *dscp*: Indicates that IP packets with the specified the dcsp field in the header are filtered. **fragment:** Indicates that only fragmented IP packets except the first fragments are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range.
Command Mode	Expert extended ACL configuration mode
Usage Guide	Run this command to add ACEs in expert extended ACL configuration mode. The ACL can be a named or numbered ACL.

- Add ACEs to an expert extended ACL in global configuration mode.

Command	**access-list** *acl-id* { **permit** | **deny** } [*protocol* | [*ethernet-type*] [**cos** [*out*] [**inner** *in*]]] [[**VID** [*out*] [**inner** *in*]]] { *source source-wildcard* | **host** *source* | **any** } { **host** *source-mac-address* | **any** } { *destination destination-wildcard* | **host** *destination* | **any** } { **host** *destination-mac-address* | **any** } [**precedence** *precedence*] [**tos** *tos*] [**fragment**] [**range** *lower upper*] [**time-range** *time-range-name*]]

Parameter Description	*acl-id*: Indicates the ID of a numbered ACL. It uniquely identifies an ACL. The value range of *acl-id* is 2700-2899. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *protocol*: Indicates the IP protocol number. The value ranges from 0 to 255. To facilitate the use, the system provides frequently-used abbreviations to replace the specific IP protocol numbers, including eigrp, gre, icmp, igmp, ip, ipinip, nos, ospf, tcp, and udp. *ethernet-type*: Indicates that L2 packets of the specified Ethernet type are filtered. **cos** *out*: Indicates that L2 packets with the specified cos field in the outer tag are filtered. **cos inner** *in*: Indicates that L2 packets with the specified cos field in the inner tag are filtered. **VID** *out*: Indicates that L2 packets with the specified VLAN ID field in the outer tag are filtered. **VID inner** *in*: Indicates that L2 packets with the specified VLAN ID field in the inner tag are filtered. *source source-wildcard:* Indicates that IP packets sent from hosts in the specified IP network segment are filtered. **host** *source*: Indicates that IP packets sent from a host with the specified source IP address are filtered. **any**: Indicates that IP packets sent from any host are filtered. **host** *source-mac-address*: Indicates that IP packets sent from a host with the specified source MAC address are filtered. **any**: Indicates that L2 packets sent to any host are filtered. *destination destination-wildcard:* Indicates that IP packets sent to hosts in a specified IP network segment are filtered. **host** *destination*: Indicates that IP packets sent to a host with the specified destination IP address are filtered. **any**: Indicates that IP packets sent to any host are filtered. **host** *destination-mac-address*: Indicates that IP packets sent to a host with the specified destination MAC address are filtered. **any**: Indicates that L2 packets sent to any host are filtered. **precedence** *precedence:* Indicates that IP packets with the specified precedence field in the header are filtered. **tos** *tos:* Indicates that IP packets with the specified the TOS field in the header are filtered. **dscp** *dscp*: Indicates that IP packets with the specified the dcsp field in the header are filtered. **fragment:** Indicates that only fragmented IP packets except the first fragments are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range.
Command Mode	Global configuration mode
Usage Guide	Run this command to add ACEs to a numbered expert extended ACL in global configuration mode. It cannot be used to add ACEs to a named expert extended ACL.

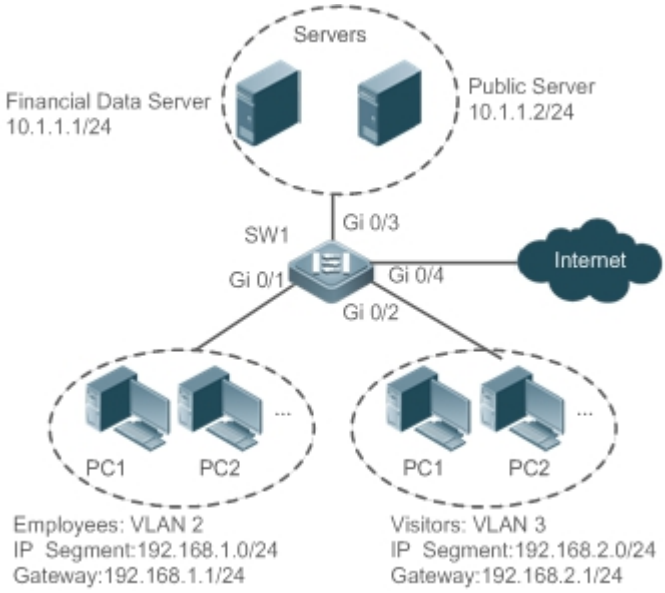
Applying an Expert Extended ACL

Command	**expert access-group** { *acl-id* **|** *acl-name* } { **in** | **out** }
Parameter Description	● *acl-id*: Indicates that a numbered expert extended ACL will be applied to the interface. ● *acl-name*: Indicates that a named expert extended ACL will be applied to the interface. ● **in**: Indicates that this ACL controls incoming L2 packets of the interface. ● **out**: Indicates that this ACL controls outgoing L2 packets of the interface.
Command Mode	Interface configuration mode
Usage Guide	This command makes an expert extended ACL take effect on the incoming or outgoing packets of a specified interface.

Configuration Example

The following configuration example describes only ACL-related configurations.

Configuring an Expert Extended ACL to Restrict Resources Accessible by Visitors (It is required that visitors and employees cannot communicate with each other, visitors can access the public resource server but not the financial data server of the company.)

Scenario **Figure 1- 5**	
Configuration Steps	● Configure an expert extended ACL. ● Add an ACE to deny packets sent from PCs in the visitor area (VLAN 3) to employee PCs in VLAN 2. ● Add an ACE to prevent visitors from accessing the financial data server of the company. ● Add an ACE to permit all packets. ● Apply the ACL to the incoming direction of the interface of the switch that connects to the visitor area.
SW1	sw1(config)#expert access-list extended 2700 sw1(config-exp-nacl)#deny ip any any 192.168.1.0 0.0.0.255 any sw1(config-exp-nacl)#deny ip any any host 10.1.1.1 any sw1(config-exp-nacl)#pemit any any any any sw1(config-exp-nacl)#exit sw1(config)#int gigabitEthernet 0/2 sw1(config-if-GigabitEthernet 0/2)#expert access-group 2700 in
Verification	● On a visitor's PC, ping the financial data server. Verify that the ping operation fails. ● On a visitor's PC, ping the public resource server. Verify that the ping operation succeeds. ● On a visitor's PC, ping the gateway address 192.168.1.1 of an employee. Verify that the ping operation fails. ● On a visitor's PC, access the Internet, for example, visit the Baidu website. Verify that the webpage can be opened.

SW1	sw1(config)#show access-lists expert access-list extended 2700 10 deny ip any any 192.168.1.0 0.0.0.255 any 20 deny ip any any host 10.1.1.1 any 30 permit ip any any any any sw1(config)#show access-group expert access-group 2700 in Applied On interface GigabitEthernet 0/2

1.4.4 Configuring an IPv6 Extended ACL

Configuration Effect

Configure and apply an IPv6 ACL to an interface/VXLAN to control all incoming and outgoing IPv5 packets of this interface/VXLAN. You can permit or deny the entry of specific IPv6 packets to a network to control access of IPv6 users to network resources.

Configuration Steps

Configuring an IPv6 ACL

- (Mandatory) Configure an IP ACL if you want to access of IPv4 users to network resources.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The IPv6 ACL takes effect only on the local device, and does not affect other devices on the network.

Adding ACEs to an IPv6 ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, all incoming IPv6 packets of the device are denied by default.

Applying an IPv6 ACL

- (Mandatory) Apply an IPv6 ACL to a specified interface on a device if you want this ACL take effect.
- You can apply an IPv6 ACL on a specified interface/VXLAN of an access, an aggregate, or a core device based on the distribution of users.

Verification

- Use the following methods to verify the configuration effects of the IPv6 ACL:
- Run the **ping** command to verify that the IPv6 ACL takes effect on the specified interface. For example, if an IPv6 ACL is configured to prohibit a host with a specified IP address or hosts in a specified IPv6 address range from accessing the network, run the **ping** command to verify that the host(s) cannot be successfully pinged.
- Access network resources, for example, visit an IPv6 website, to check whether the IPv6 ACL takes effect on the specified interface.

Related Commands

Configuring an IPv6 ACL

Command	**ipv6 access-list** *acl-name*
Parameter Description	*acl-name*: Indicates the name of a standard or an extended IP ACL. The name is a string of 1 to 99 characters. The ACL name cannot start with numbers (0–9), "in", or "out".
Command Mode	Global configuration mode
Usage Guide	Run this command to configure an IPv6 ACL and enter IPv6 configuration mode.

Adding ACEs to an IPv6 ACL

- To filter TCP or UDP packets, add ACEs to an IPv6 ACL as follows:

Command	[*sn*] {**permit** | **deny** } *protocol* {*src-ipv6-prefix/prefix-len* | **host** *src-ipv6-addr* | **any**} {*dst-ipv6-pfix/pfix-len* | **host** *dst-ipv6-addr* | **any**} [*op dstport* | **range** *lower upper*] [**dscp** *dscp*] [**flow-label** *flow-label*] [**fragment]** [**time-range***tm-rng-name*][**log**]
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *protocol*: Indicates the IPv6 protocol number. The value ranges from 0 to 255. To facilitate the use, the system provides frequently-used abbreviations of IPv6 protocol numbers to replace the specific IP protocol numbers, including icmp, ipv6, tcp, and udp. *src-ipv6-prefix/prefix-len:* Indicates that IP packets sent from hosts in the specified IPv6 network segment are filtered. **host** *src-ipv6-addr*: Indicates that IPv6 packets sent from a host with the specified source IP address are filtered. **any**: Indicates that IPv6 packets sent from any host are filtered. *dst-ipv6-pfix/pfix-len:* Indicates that IPv6 packets sent from hosts in the specified IPv6 network segment are filtered. **host** *dst-ipv6-addr*: Indicates that IPv6 packets sent to a host with the specified destination IP address are filtered. **any**: Indicates that IPv6 packets sent to any host are filtered. *op dstport*: Indicates that TCP or UDP packets are filtered based on the L4 destination port number. The value of the **op** parameter can be **eq** (equal to), **neq** (not equal to), **gt** (greater than), or **lt** (smaller than). **range** *lower upper*: Indicates that TCP or UDP packets with the L4 destination port number in the specified range are filtered. **dscp** *dscp*: Indicates that IPv6 packets with the specified the dcsp field in the header are filtered. **flow-label** *flow-label*: Indicates that IPv6 packets with the specified the flow label field in the header are filtered. **fragment:** Indicates that only fragmented IPv6 packets except the first fragments are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range. **log**: Indicates that logs will be periodically output if packets matching the ACEs are found. For details about logs, see "ACL Logging" in this document.
Command Mode	IPv6 ACL configuration mode
Usage Guide	Run this command to add ACEs in IPv6 ACL configuration mode.

- To filter IPv6 packets except for the TCP or UDP packets, add ACEs to an IPv6 ACL as follows:

Command	[*sn*] { **permit** | **deny** } *protocol* { *src-ipv6-prefix/prefix-len* | **host** *src-ipv6-addr* | **any** } { *dst-ipv6-pfix/pfix-len* | **host** *dst-ipv6-addr* | **any** } [**dscp** *dscp*] [**flow-label** *flow-label*] [**fragment**] [**time-range***tm-rng-name*] [**log**]
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command.

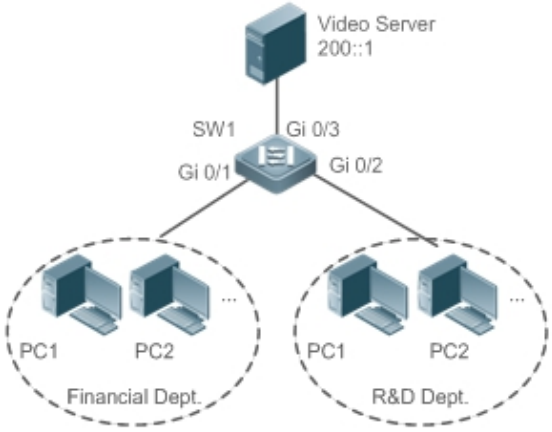
	permit: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *protocol*: Indicates the IPv6 protocol number. The value ranges from 0 to 255. To facilitate the use, the system provides frequently-used abbreviations of IPv6 protocol numbers to replace the specific IP protocol numbers, including icmp, ipv6, tcp, and udp. *src-ipv6-prefix/prefix-len:* Indicates that IP packets sent from hosts in the specified IPv6 network segment are filtered. **host** *src-ipv6-addr*: Indicates that IPv6 packets sent from a host with the specified source IP address are filtered. **any**: Indicates that IPv6 packets sent from any host are filtered. *dst-ipv6-pfix/pfix-len:* Indicates that IPv6 packets sent from hosts in the specified IPv6 network segment are filtered. **host** *dst-ipv6-addr*: Indicates that IPv6 packets sent to a host with the specified destination IP address are filtered. **any**: Indicates that IPv6 packets sent to any host are filtered. **dscp** *dscp*: Indicates that IPv6 packets with the specified the dcsp field in the header are filtered. **flow-label** *flow-label*: Indicates that IPv6 packets with the specified the flow label field in the header are filtered. **fragment:** Indicates that only fragmented IPv6 packets except the first fragments are filtered. **time-range** *time-range-name:* Indicates that this ACE is associated with a time range. The ACE takes effect only within this time range. For details about the time range, see the configuration manual of the time range. **log**: Indicates that logs will be periodically output if packets matching the ACEs are found. For details about logs, see "ACL Logging" in this document.
Command Mode	IPv6 ACL configuration mode
Usage Guide	Run this command to add ACEs in IPv6 ACL configuration mode.

Applying an IPv6 ACL

Command	**ipv6 traffic-filter** *acl-name* { **in** | **out** }
Parameter Description	*acl-name*: Indicates the name of an IPv6 ACL. **in**: Indicates that this ACL controls incoming IPv6 packets of the interface. **out**: Indicates that this ACL controls outgoing IPv6 packets of the interface.
Command Mode	Interface configuration mode
Usage Guide	This command makes an IPv6 ACL take effect on the incoming or outgoing packets of the specified interface.

Configuration Example

Configuring an IPv6 ACL to Prohibit the R&D Department from Accessing the Video Server

Scenario **Figure 1- 6**	

Configuration Steps	● Configure an IPv6 ACL. ● Add an ACE to the IPv6 ACL to prevent access to the video server. ● Add an ACE to the IPv6 ACL to permit all IPv6 packets. ● Apply the IPv6 ACL to the incoming direction of the interface connected to the R&D department.
SW1	sw1(config)#ipv6 access-list dev_deny_ipv6video sw1(config-ipv6-nacl)#deny ipv6 any host 200::1 sw1(config-ipv6-nacl)#permit ipv6 any any sw1(config-ipv6-nacl)#exit sw1(config)#int gigabitEthernet 0/2 sw1(config-if-GigabitEthernet 0/2)# ipv6 traffic-filter dev_deny_ipv6video in
Verification	● On a PC of the R&D department, ping the video server. Verify that the ping operation fails.
SW1	sw1(config)#show access-lists ipv6 access-list dev_deny_ipv6video 10 deny ipv6 any host 200::1 20 permit ipv6 any any sw1(config)#show access-group ipv6 traffic-filter dev_deny_ipv6video in Applied On interface GigabitEthernet 0/2

1.4.5 Configuring an ACL80

Configuration Effect

When the IP ACL, MAC extended ACL, expert extended ACL, and IPv6 ACL with fixed matching fields cannot meet requirements, configure the ACL80 to customize the packet fields that need to be matched.

Configuration Steps

Configuring an Expert Advanced ACL

- (Mandatory) Configure an expert advanced ACL if you want to implement the ACL80 function. For details about how to configure the expert advanced ACL, see the related descriptions.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The expert advanced ACL takes effect only on the local device, and does not affect other devices on the network.

Adding ACEs to an Expert Advanced ACL

- (Mandatory) Add ACEs to an expert advanced ACL to customize matching fields. If no ACE is added to the expert advanced ACL, the deny ACEs will drop all packets by default. For details about how to add an ACE to an expert advanced ACL, see the related descriptions.

Applying an Expert Advanced ACL

- (Mandatory) Apply an expert advanced ACL to a specified interface if you want this ACL take effect.
- You can apply an expert advanced ACL on a specified interface of an access, an aggregate, or a core device based on the distribution of users.

Verification

- Use the following methods to verify the configuration effects of the expert advanced ACL:
- Run the **ping** command to check whether the configurations take effect.
- Construct packets matching the ACEs to check whether ACEs take effect.

Related Commands

Configuring an Expert Advanced ACL

Command	**expert access-list advanced** *acl-name*
Parameter Description	*acl-name*: Indicates the name of an expert advanced ACL. The name is a string of 1 to 99 characters. The ACL name cannot start with numbers (0–9), "in", or "out".
Command Mode	Global configuration mode
Usage Guide	Run this command to configure an expert advanced ACL and enter expert advanced ACL configuration mode.

Adding ACEs to an Expert Advanced ACL

Command	[*sn*] { **permit** | **deny** } *hex hex-mask offset*
Parameter Description	*sn*: Indicates the sequence number of an ACE. The value ranges from 1 to 2,147,483,647. This sequence number determines the priority of this ACE in the ACL. A smaller sequence number indicates a higher priority. An ACE with a higher priority will be preferentially used to match packets. If you do not specify the sequence number when adding an ACE, the system automatically allocates a sequence number, which is equal to an increment (10 by default) plus the sequence number of the last ACE in the current ACL. For example, if the sequence number of the last ACE is 100, the sequence number of a newly-added ACE will be 110 by default. You can adjust the increment using a command. **permit**: Indicates that the ACE is a permit ACE. **deny**: Indicates that the ACE is a deny ACE. *hex*: Indicates the customized matching rule expressed in hexadecimal format, for example, 00d0f800. *hex-mask*: Indicates the matching mask. *offset*: Indicates the start position of matching. For example, if the matching content is 00d0f800, the matching mask is 00ff0000, and start position is 6, the destination MAC address of each packet is compared. All packets whose second byte of the destination MAC address is d0 match this ACE.
Command Mode	Expert advanced ACL configuration mode
Usage Guide	Run this command to add ACEs in expert advanced ACL configuration mode.

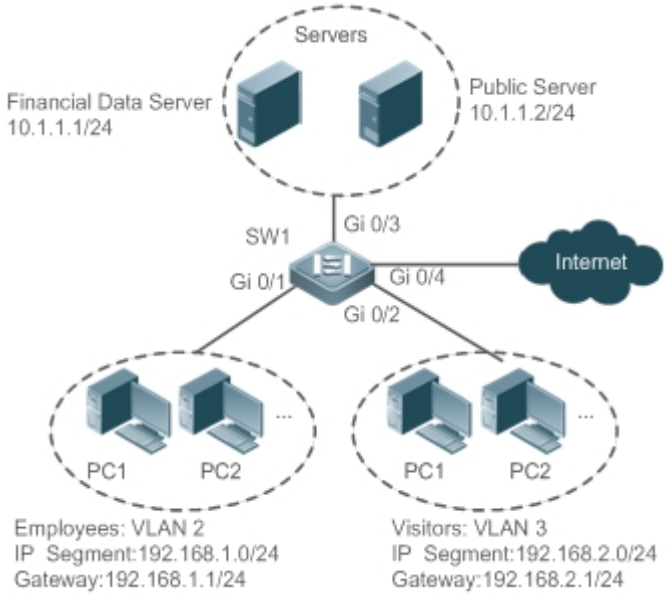
Applying an Expert Advanced ACL

Command	**expert access-group** *acl-n* { **in** | **out** }
Parameter Description	*acl-id*: Indicates that a numbered expert advanced ACL will be applied to the interface. *acl-name*: Indicates that a named expert advanced ACL will be applied to the interface. **in**: Indicates that this ACL controls incoming L2 packets of the interface. **out**: Indicates that this ACL controls outgoing L2 packets of the interface.
Command Mode	Interface configuration mode
Usage Guide	This command makes an expert advanced ACL take effect on the incoming or outgoing packets of a specified interface.

Configuration Example

The following configuration example describes only ACL-related configurations.

Configuring an ACL80 to Restrict Resources Accessible by Visitors (It is required that visitors and employees cannot communicate with each other, visitors can access the public resource server but not the financial data server of the company.)

Scenario **Figure 1- 7**	
Configuration Steps	● Configure an expert advanced ACL. ● Add an ACE to deny packets sent from PCs in the visitor area (VLAN 3) to employee PCs in VLAN 2. ● Add an ACE to prevent visitors from accessing the financial data server of the company. ● Add an ACE to permit all packets. ● Apply the ACL to the incoming direction of the interface of the switch that connects to the visitor area.
SW1	sw1(config)#expert access-list advanced acl80-guest sw1(config-exp-dacl)#deny C0A801 FFFFFF 42 sw1(config-exp-dacl)#deny 0A010101 FFFFFFFF 42 sw1(config-exp-dacl)#permit 0806 FFFF 24 sw1(config-exp-dacl)#permit 0800 FFFF 24 sw1(config-exp-dacl)#exit sw1(config)#int gigabitEthernet 0/2 sw1(config-if-GigabitEthernet 0/2)#expert access-group acl80-guest in
Verification	● On a visitor's PC, ping the financial data server. Verify that the ping operation fails. ● On a visitor's PC, ping the public resource server. Verify that the ping operation succeeds. ● On a visitor's PC, ping the gateway address 192.168.1.1 of an employee. Verify that the ping operation fails. ● On a visitor's PC, access the Internet, for example, visit the Baidu website. Verify that the webpage can be opened.

SW1	sw1(config)#show access-lists expert access-list advanced sss 10 deny C0A801 FFFFFF 42 20 deny 0A010101 FFFFFFFF 42 30 permit 0806 FFFF 24 40 permit 0800 FFFF 24 expert access-group acl80-guest in Applied On interface GigabitEthernet 0/2

1.4.6 Configuring ACL Redirection

Configuration Effect

Configure the ACL redirection function on a specified interface to directly redirect specified packets on the interface to a specified port for further forwarding.

Configuration Steps

Configuring an ACL

- (Mandatory) To implement ACL redirection, you must first configure an ACL, for example, an IP, MAC extended, or expert extended ACL. For details about how to configure an ACL, see the related descriptions.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The IPv6 ACL takes effect only on the local device, and does not affect other devices on the network.

Adding ACEs to an ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, the ACL redirection function is not available. For details about how to add an ACE to an ACL, see the related descriptions.

Configuring ACL Redirection

- (Mandatory) Enable ACL redirection on a specified interface if you want to implement ACL redirection.
- You can configure the ACL redirection function on a specified interface of an access, an aggregate, or a core device based on the distribution of users.

Verification

Send packets matching ACEs on the port where ACL redirection is enabled, and then use the packet capturing software on the destination port to check whether the ACL redirection function takes effect.

Related Commands

Configuring an ACL

For details about how to configure an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Configuring ACL Redirection

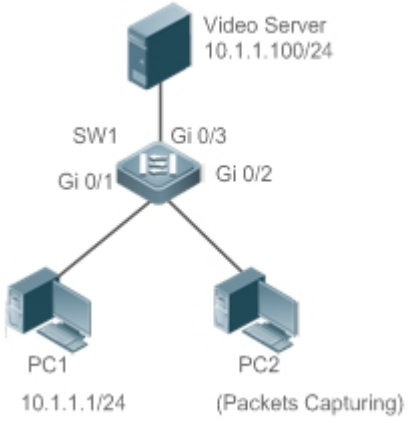
Command	**redirect destination interface** *interface-name* **acl** {*acl-id* | *acl-name* } **in**

Parameter Description	**interface** *interface-name*: Indicates the name of the destination port for redirection. *acl-id*: Indicates the ID of an ACL. *acl-name*: Indicates the name of an ACL. **in**: Indicates that incoming packets of the interface are redirected.
Command Mode	Interface configuration mode
Usage Guide	Run this command to redirect incoming packets of the interface that match ACEs to the destination port for further forwarding.

Configuration Example

The following configuration example describes only ACL-related configurations.

Enabling ACL Redirection to Redirect Packets Sent from the Host 10.1.1.1 to the Packet Capturing Device for Analysis

Scenario Figure 1- 8	
Configuration Steps	● Configures an IP ACL. ● Add an ACE to the IP ACL to permit packets sent from the host 10.1.1.1. ● Enable ACL redirection on the port GI 0/1, and set the destination port to Gi 0/2.
SW1	sw1(config)#ip access-list standard 1 sw1 (config-std-nacl)#permit host 10.1.1.1 sw1(config-std-nacl)#exit sw1(config)#int gigabitEthernet 0/1 sw1(config-if-GigabitEthernet 0/1)# redirect destination interface gigabitEthernet 0/2 acl 1
Verification	● Capture packets on PC 2. Ping the video server on PC 1. Verify that ICMP requests sent from PC 1 are captured on PC 2.
SW1	sw1#show access-lists ip access-list standard 1 10 permit host 10.1.1.1 sw1#show redirect interface gigabitEthernet 0/1 acl redirect configuration on interface gigabitEthernet 0/1 redirect destination interface gigabitEthernet 0/2 acl 1 in

1.4.7 Configuring a Global Security ACL

Configuration Effect

Configure a global security ACL to prevent internal PCs of a company from accessing illegal websites or prevent virus from attacking the company's internal network. You can also configure exclusive interfaces to allow specified departments of the company to access external websites.

Configuration Steps

Configuring an ACL

- (Mandatory) Configure an ACL if you want to protect the internal network globally. For details about the configuration method, see the earlier descriptions about the ACL.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The configurations take effect only on the local device, and do not affect other devices on the network.

Adding ACEs to an ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, it is equivalent that the global security ACL does not exist. For details about how to add an ACE to an ACL, see the related descriptions.

Configuring a Global Security ACL

- (Mandatory) Enable the global security function if you want to make the global security ACL take effect.
- You can configure a global security ACL on an access, an aggregate, or a core device based on the distribution of users.

Verification

On the internal network protected by the global security ACL, ping the website or device that are denied by ACEs to check whether the global security ACL takes effect.

Related Commands

Configuring an ACL

For details about the configuration method, see the earlier descriptions about the ACL.

Adding ACEs to an ACL

For details about the configuration method, see the earlier descriptions about the ACL.

Configuring a Global Security ACL

Command	{ **ip** | **mac** | expert } **access-group** *acl-id* { **in** | **out** }
Parameter Description	*acl-id*: Indicates the ID of an ACL. **in**: Filters the incoming packets of the device. **out**: Filters the outgoing packets of the device.
Command Mode	Global configuration mode
Usage Guide	Run this command to enable the global security ACL so that the ACL takes effect on all L2 interfaces of the device.

Configuring an Exclusive Interface of the Global Security ACL

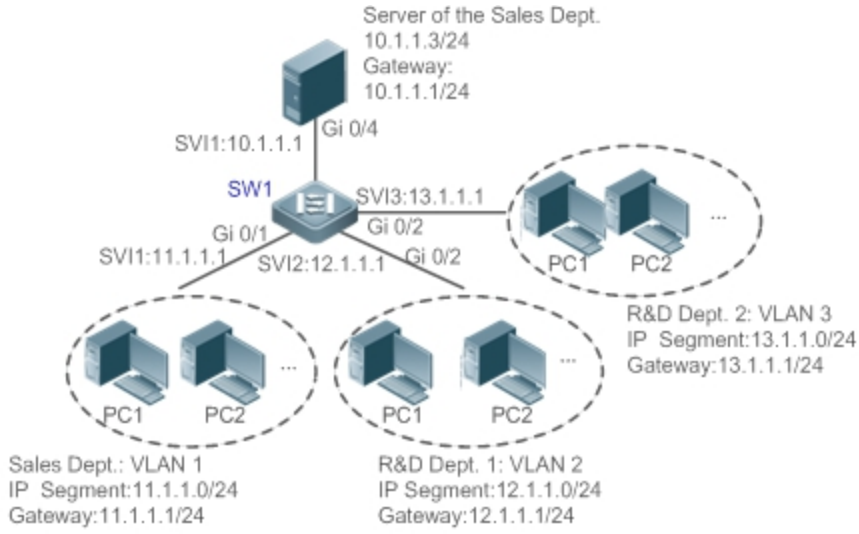
Command	**no global ip access-group**
Parameter Description	N/A
Command Mode	Interface configuration mode

Usage Guide	Run this command to invalidate a global security ACL on a specified interface.

Configuration Example

The following configuration example describes only ACL-related configurations.

Configuring a Global Security ACL to Prevent the R&D Department From Accessing the Server of the Sales Department but Allow the Sales Department to Access This Server

Scenario **Figure 1- 9**	
Configuration Steps	● Configure an extended IP ACL "ip_ext_deny_dst_sale_server". ● Add the ACE that prevents the device to forward packets to the destination host 10.1.1.3/24. ● Configure the ACL "ip_ext_deny_dst_sale_server" as a global security ACL. ● Configure the interface directly connected to the sales department as the exclusive interface of the global security ACL.
SW1	sw1(config)#ip access-list extended ip_ext_deny_dst_sale_server sw1(config-ext-nacl)# deny ip any host 10.1.1.3 sw1(config-ext-nacl)#exit sw1(config)#ip access-group ip_ext_deny_dst_sale_server in sw1(config)#int gigabitEthernet 0/1 sw1(config-if-GigabitEthernet 0/1)# no global ip access-group
Verification	● On a PC of the sales department, ping the server of the sales department. Verify that the ping operation succeeds. ● On the PCs of R&D department 1 and R&D department 2, ping the server of the sales department. Verify that the ping operations fail.
	sw1#show access-lists ip access-list extended ip_ext_deny_dst_sale_server 10 deny ip any host 10.1.1.3 sw1#show running …… ! ip access-group ip_ext_deny_dst_sale_server in

```
!
!
!
!
!
!
!
!
interface GigabitEthernet 0/1
 no global ip access-group
!
……
```

1.4.8 Configuring a Security Channel

Configuration Effect

Configure a security channel to enable packets meeting the security channel rules to bypass the checks of access control applications.

Configuration Steps

Configuring an ACL

- (Mandatory) Configure an ACL before configuring the security channel. For details about the configuration method, see the earlier descriptions.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The configurations take effect only on the local device, and do not affect other devices on the network.

Adding ACEs to an ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured for an ACL, it is equivalent that the security channel does not take effect. For details about how to add an ACE to an ACL, see the related descriptions.

Configuring a Security Channel on a Specified Interface, VXLAN or Globally

- Configure a security channel on an interface if you want this security channel to take effect on the interface. Configure a VXLAN security channel if you want this security channel to take effect on VNI. Configure a global security channel if you want this security channel to take effect globally. You must configure either the interface-based security channel or the global security channel.
- You can configure a security channel on an access, an aggregate, or a core device based on the distribution of users.

Configuring an Exclusive Interface for the Global Security Channel

- (Optional) Configure an interface as the exclusive interface for the global security channel if you do not want the global security channel to take effect on this interface.

Configuring an Access Control Application

- (Optional) You can enable the DOT1X or Web authentication function to verify the security channel function.
- You can configure the access control function on an access, an aggregate, or a core device based on the distribution of users.

Verification

On a PC that is subject to the control of an access control application, ping the resources (devices or servers) that are allowed to bypass the check of the access control application to verify the configuration of the security channel.

Related Commands

Configuring an ACL

For details about how to configure an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Configuring a Security Channel on an Interface

Command	**security access-group** {*acl-id* | *acl-name* }
Parameter Description	*acl-id*: Indicates that ID of the ACL that is configured as the security channel. *acl-name*: Indicates that name of the ACL that is configured as the security channel.
Command Mode	Interface configuration mode
Usage Guide	Run this command to configure a specified ACL as the security channel on the specified interface.

Configuring a VXLAN Security Channel

Command	**security access-group** {*acl-id* | *acl-name* }
Parameter Description	*acl-id*: Indicates that ID of the ACL that is configured as the security channel. *acl-name*: Indicates that name of the ACL that is configured as the security channel.
Command Mode	VXLAN configuration mode
Usage Guide	Run this command to configure a specified ACL as the security channel on the specified VXLAN.

Configuring a Global Security Channel

Command	**security global access-group** {*acl-id* | *acl-name* }
Parameter Description	*acl-id*: Indicates that ID of the ACL that is configured as the security channel. *acl-name*: Indicates that name of the ACL that is configured as the security channel.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure the specified ACL as the global security channel.

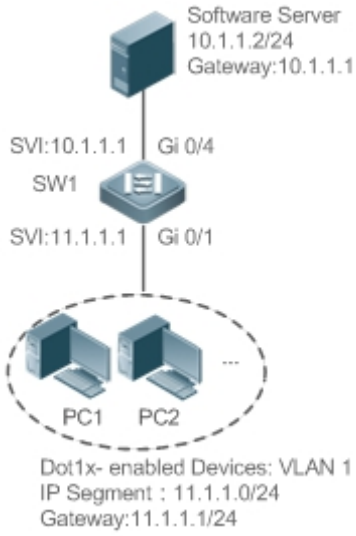
Configuring an Exclusive Interface for the Global Security Channel

Command	**security uplink enable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	Run this command to configure the specified interface as the exclusive interface of the global security channel.

Configuration Example

The following configuration example describes only ACL-related configurations.

Enabling DOT1X Authentication and Configuring a Security Channel to Allow Users to Download the SU Software From the Server Before Authentication

Scenario **Figure 1- 10**	
Configuration Steps	● Configure an expert extended ACL "exp_ext_esc". ● Add an ACE to allow forwarding packets to the destination host 10.1.1.2. ● Add an ACE to permit the DHCP packets. ● Add an ACE to permit the ARP packets. ● On the interface where DOT1X authentication is enabled, configure the ACL "exp_ext_esc" as the security channel.
SW1	sw1(config)#expert access-list extended exp_ext_esc sw1(config-exp-nacl)# permit ip any any host 10.1.1.2 any sw1(config-exp-nacl)# permit 0x0806 any any any any any sw1(config-exp-nacl)# permit tcp any any any any eq 67 sw1(config-exp-nacl)# permit tcp any any any any eq 68 sw1(config)#int gigabitEthernet 0/1 sw1(config-if-GigabitEthernet 0/1)# security access-group exp_ext_esc
Verification	● On a PC of the sales department, ping the server of the sales department. Verify that the ping operation succeeds. ● On the PCs of R&D department 1 and R&D department 2, ping the server of the sales department. Verify that the ping operations fail.
	sw1#show access-lists expert access-list extended exp_ext_esc 10 permit ip any any host 10.1.1.2 any 20 permit arp any any any any any 30 permit tcp any any any any eq 67 40 permit tcp any any any any eq 68…… sw1#show running-config interface gigabitEthernet 0/1 Building configuration... Current configuration : 59 bytes

	interface GigabitEthernet 0/1 security access-group exp_ext_esc

1.4.9 Configuring the Time Range-Based ACEs

Configuration Effect

Configure the time range-based ACEs if you want some ACEs to take effect or to become invalid in a specified period of time, for example, in some time ranges during a week.

Configuration Steps

Configuring an ACL

- (Mandatory) Configure an ACL if you want ACEs to take effect in the specified time range. For details about the configuration method, see the earlier descriptions.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The configurations take effect only on the local device, and do not affect other devices on the network.

Adding an ACE with the Time Range Specified

- (Mandatory) Specify the time range when adding an ACE. For details about how to configure the time range, see the configuration manual related to the time range.

Applying an ACL

- (Mandatory) Apply the ACL to a specified interface if you want to make ACEs take effect in the specified time range.
- You can apply an IP ACL on a specified interface of an access, an aggregate, or a core device based on the distribution of users.

Verification

In the time range that the configured ACE takes effect or becomes invalid, run the **ping** command or construct packets matching the ACE to check whether the ACE takes effect or becomes invalid.

Related Commands

Configuring an ACL

For details about the ACL configuration commands, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Adding an ACE with the Time Range Specified

For details about the ACE configuration commands, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

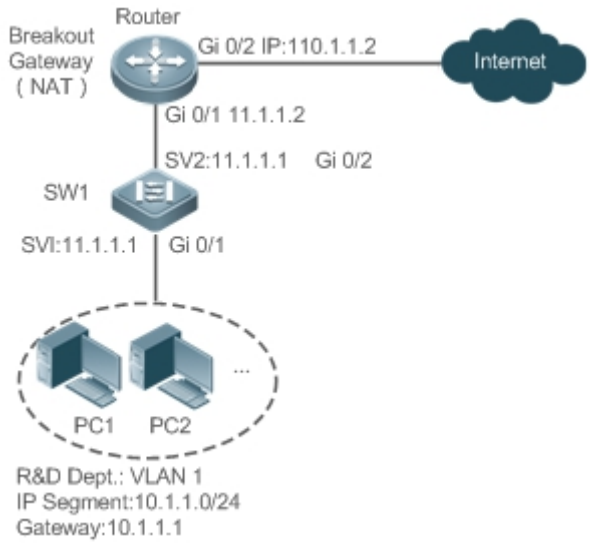
Applying an ACL

For details about the command for applying an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Configuration Example

The following configuration example describes only ACL-related configurations.

Adding an ACE With the Time Range Specified to Allow the R&D Department to Access the Internet Between 12:00 and 13:30 Every Day

Scenario **Figure 1- 11**	
Configuration Steps	● Configure a time range named "access-internet", and add an entry of the time range between 12:00 and 13:30 every day. ● Configure an IP ACL "ip_std_internet_acl". ● Add an ACE to allow packets with the source IP address in the network segment 10.1.1.0/24, and associate this ACE with the time zone "access-internet". ● Add an ACE to deny packets with the source IP address the network segment 10.1.1.0/24. Access to the Internet is not allowed except in the specified time range. ● Add an ACE to permit all packets. ● Apply the ACL to the outgoing direction of the interface connected to the breakout gateway.
SW1	FS(config)# time-range access-internet FS(config-time-range)# periodic daily 12:00 to 13:30 FS(config-time-range)# exit sw1(config)# ip access-list standard ip_std_internet_acl sw1(config-std-nacl)# permit 10.1.1.0 0.0.0.255 time-range access-internet sw1(config-std-nacl)# deny 10.1.1.0 0.0.0.255 sw1(config-std-nacl)# permit any sw1(config-std-nacl)# exit sw1(config)#int gigabitEthernet 0/2 sw1(config-if-GigabitEthernet 0/2)# ip access-group ip_std_internet_acl out
Verification	● Within the time range between 12:00 and 13:30, visit the Baidu website on a PC of the R&D department. Verify that the website can be opened normally. ● Beyond the time range between 12:00 and 13:30, visit the Baidu website on a PC of the R&D department. Verify that the website cannot be opened.
SW1	sw1#show time-range time-range entry: access-internet (inactive) periodic Daily 12:00 to 13:30

```
sw1#show access-lists

ip access-list standard ip_std_internet_acl
  10 permit 10.1.1.0 0.0.0.255 time-range access-internet (inactive)
  20 deny 10.1.1.0 0.0.0.255
  30 permit any

sw1#show access-group
ip access-group ip_std_internet_acl out
Applied On interface GigabitEthernet 0/2
```

1.4.10 Configuring Comments for ACLs

Configuration Effect

During network maintenance, if a lot of ACLs are configured without any comments, it is difficult to distinguish these ACLs later on. You can configure comments for ACLs to better understand the intended use of ACLs.

Configuration Steps

Configuring an ACL

- (Mandatory) Configure an ACL before configuring the security channel. For details about the configuration method, see the earlier descriptions.
- You can configure this ACL on an access, an aggregate, or a core device based on the distribution of users. The configurations take effect only on the local device, and do not affect other devices on the network.

Configuring Comments for ACLs

- (Optional) Configure comments for ACLs so that it is easy to manage and understand the configured ACLs.

Adding ACEs to an ACL

- (Optional) An ACL may contain zero or multiple ACEs. If no ACE is configured, it is equivalent that the security channel does not take effect. For details about how to add an ACE to an ACL, see the related descriptions.

Configuring Comments for ACEs

- (Optional) To facilitate understanding of a configured ACL, you can configure comments for ACEs in addition to comments for the ACL.

Verification

Run the **show access-lists** command on the device to display the comments configured for ACLs.

Related Commands

Configuring an ACL

For details about how to configure an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Configuring a Comment for an ACL

Use either of the following two methods to configure a comment for an ACL:

Command	**list-remark** *comment*
Parameter Description	***comment***: Indicates the comment. The value is a string of 1 to 100 characters. A comment longer than 100 characters will be truncated to 100 characters.
Command	ACL configuration mode

Mode	
Usage Guide	Run this command to configure the comment for a specified ACL.

Command	**access-list** *acl-id* **list-remark** *comment*
Parameter Description	***acl-id***: Indicates the ID of an ACL. ***comment***: Indicates the comment. The value is a string of 1 to 100 characters. A comment longer than 100 characters will be truncated to 100 characters.
Command Mode	Configuration mode
Usage Guide	Run this command to configure the comment for a specified ACL.

Adding ACEs to an ACL

For details about how to add ACEs to an ACL, see the earlier descriptions about the IP ACL, MAC extended ACL, expert extended ACL, or IPv6 ACL.

Configuring Comments for ACEs

Use either of the following two methods to configure a comment for an ACE:

Command	[sn] **remark** *comment*
Parameter Description	***comment***: Indicates the comment. The value is a string of 1 to 100 characters. A comment longer than 100 characters will be truncated to 100 characters. sn: Indicates the sequence number of ACE.
Command Mode	ACL configuration mode
Usage Guide	Run this command to configure the comment for a specified ACE. If sn is not specified, the remark is applied to the last ACE.

Command	**access-list** *acl-id* **sn remark** *comment*
Parameter Description	***acl-id***: Indicates the ID of an ACL. ***comment***: Indicates the comment. The value is a string of 1 to 100 characters. A comment longer than 100 characters will be truncated to 100 characters. sn: Indicates the sequence number of ACE.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure the comment for a specified ACE. If sn is not specified, the remark is applied to the last ACE.

1.5 Monitoring

Clearing

Description	Command
Clears the ACL packet matching counters.	**clear counters access-list** [*acl-id* | *acl-name*]
Clears the counters of packets matching the deny ACEs.	**clear access-list counters [***acl-id* |*acl-name* **]**

Displaying

Description	Command
Displays the basic ACLs.	**show access-lists** [*acl-id* | *acl-ame*] [**summary**]
Displays the redirection ACEs bound to a specified interface. If the interface is not specified, redirection ACEs bound to all interfaces are displayed.	**show redirect** [**interface** *interface-name*]
Displays the ACL configurations applied to an interface.	**show access-group** [**interface** *interface-name*]
Displays the IP ACL configurations applied to an interface.	**show ip access-group** [**interface** *interface-name*]
Displays the MAC extended ACL configurations applied to an interface.	**show mac access-group** [**interface** *interface-name*]
Displays the expert extended ACL configurations applied to an interface.	**show expert access-group** [**interface** *interface-name*]
Displays the IPv6 ACL configurations applied to an interface.	**show ipv6 traffic-filter** [**interface** *interface-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the ACL running process.	**debug acl acld event**
Debugs the ACL clients.	**debug acl acld client-show**
Debugs the ACLs created by all ACL clients.	**debug acl acld acl-show**

2 Configuring QoS

2.1 Overview

Quality of Service (QoS) indicates that a network can provide a good service capability for specified network communication by using various infrastructure technologies.

When the network bandwidth is sufficient, all data streams can be properly processed; when network congestion occurs, all data streams may be discarded. To meet users' requirements for different applications and different levels of service quality, a network must be able to allocate and schedule resources based on users' requirements and provide different levels of service quality for different data streams.

To be specific, the network can process real-time and important data packets in higher priorities, and process non-real-time and common data packets in lower priorities and even discard the data packets upon network congestion.

The "doing the best" forwarding mechanism used by traditional networks cannot meet the requirements any longer and then QoS comes into being. QoS-enabled devices provide transmission QoS quality service. A transmission priority can be assigned to data streams of a type to identify the importance of the data streams. Then, the devices provide forwarding policies for different priorities, congestion mitigation and other mechanisms to provide special transmission services for these data streams. A network environment configured with QoS can provide predictability for network performance, effectively allocate network bandwidth, and reasonably utilize network resources.

2.2 Applications

Application	Description
Interface Rate Limit + Priority Relabeling	Based on different service requirements for a campus network, provide rate control and priority-based processing for outgoing traffic of the teaching building, laboratories and dormitory building.
Priority Relabeling + Queue Scheduling	Provide priority-based processing and bandwidth control for traffic of internal access to servers of an enterprise.

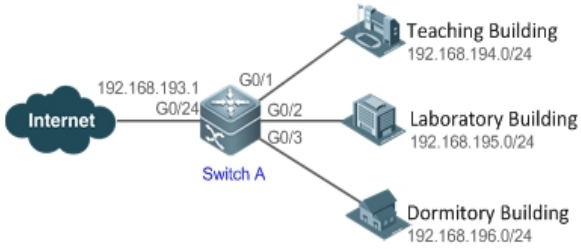
2.2.1 Interface Rate Limit + Priority Relabeling

Scenario

To meet the service requirements of normal teaching, a school puts forwards the following requirements:

- Control the Internet access traffic under 100M and discard packets out of control.
- Control the outgoing traffic of the dormitory building under 50M and discard packets out of control.
- Control the rate of packets with DSCP priority 7 sent from laboratories under 20M, and change the DSCP priorities of these packets whose rates exceed 20M to 16.
- Control the outgoing traffic of the teaching building under 30M and discard packets out of control.

Figure 2- 1



Remarks	A school connects GigabitEthernet 0/24 of Switch A to the Internet in the uplink and connects GigabitEthernet 0/1, GigabitEthernet 0/2 and GigabitEthernet 0/3 of Switch A to the teaching building, laboratory and dormitory building in the downlink respectively.

Deployment

- Configure the QoS interface rate limit for the interface G0/24 of Switch A for connecting the Internet.
- Configure the QoS rate limit for packets sent from the dormitory building on Switch A.
- Set the rate limit for packets with the DSCP priority 7 sent from the laboratory to 20M and relabel the DSCP priority of packets out of the rate limit to 16.
- Configure the QoS rate limit for packets sent from the teaching building on Switch A.

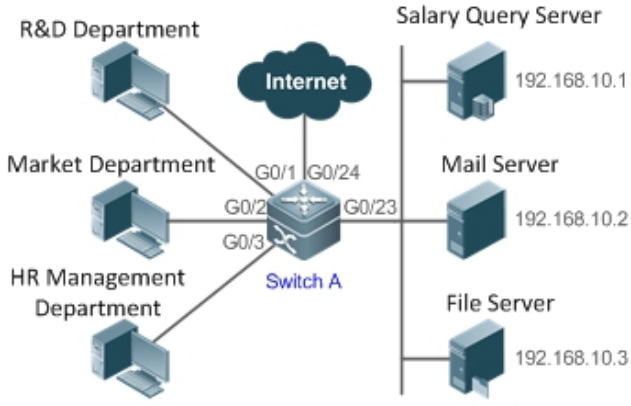
2.2.2 Priority Relabeling + Queue Scheduling

Scenario

Configure priority relabeling and queue scheduling to meet the following requirements:

- When the R&D department and market department access servers, the priorities of the server packets are as follows: mail server > file server > salary query server.
- No matter when the HR management department accesses the Internet or servers, the switch processes the corresponding packets in the highest priority.
- Since network congestion often occurs in switch running, in order to ensure smooth business operation, WRR queue scheduling must be used to schedule IP packets for the R&D and market departments to access the mail database, file database, and salary query database based on the ratio of 6:2:1.

Figure 2- 2



Remarks	The R&D, market and HR management departments access the interfaces GigabitEthernet 0/1, GigabitEthernet 0/2 and GigabitEthernet 0/3 of Switch A respectively. The salary query server, mail server and file server are connected to GigabitEthernet 0/23 of Switch A.

Deployment

- Configure the CoS values of data streams for accessing different servers to ensure that the switch processes packets for different servers in different priorities.
- Set the default CoS value of the interface to a specific value to ensure that the switch processes packets sent by the HR management department in the highest priority.
- Configure WRR queue scheduling to ensure that data packets are transmitted in a specific quantity ratio.

2.3 Features

Basic Concept

DiffServ

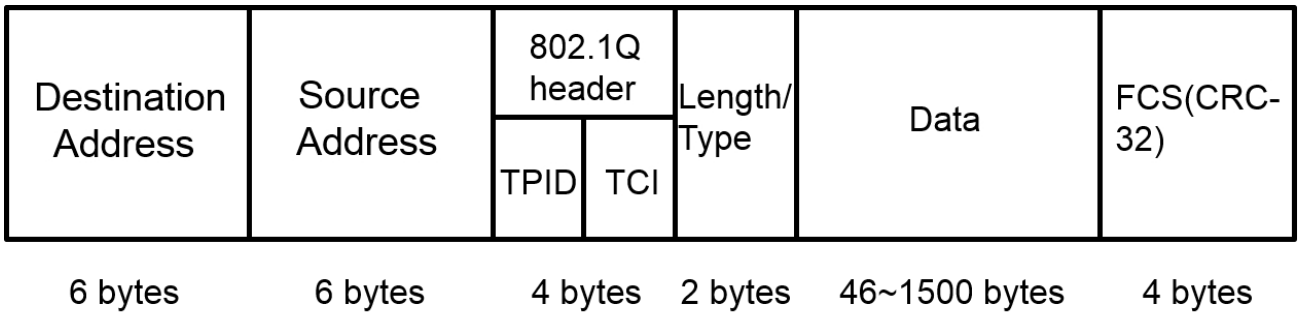
The Differentiated Services (DiffServ) Mode is an IETF system based on which QoS is implemented in FS products. The DiffServ system classifies all packets transmitted in a network into different types. The classification information is included in layer-2/3 packet headers, including 802.1P, IP and IP DSCP priorities.

In a DiffServ-compliant network, all devices apply the same transmission service policy to packets containing the same classification information and apply different transmission service policies to packets containing different classification information. Classification information of packets is either assigned by hosts or other devices in the network or assigned based on different application policies or different packet contents. Based on the classification information carried by packets, a device may provide different transmission priorities for different packet streams, reserve bandwidth for a kind of packet streams, discard certain packets with lower priorities, or take some other actions.

802.1P(PRI) priority

The 802.1 P priority is located at the header of a layer-2 packet with the 802.1Q header, and is used in scenarios where layer-3 headers do not need to be analyzed and QoS needs to be implemented at layer 2. Figure 2-3 shows the structure of a layer-2 packet.

Figure 2- 3

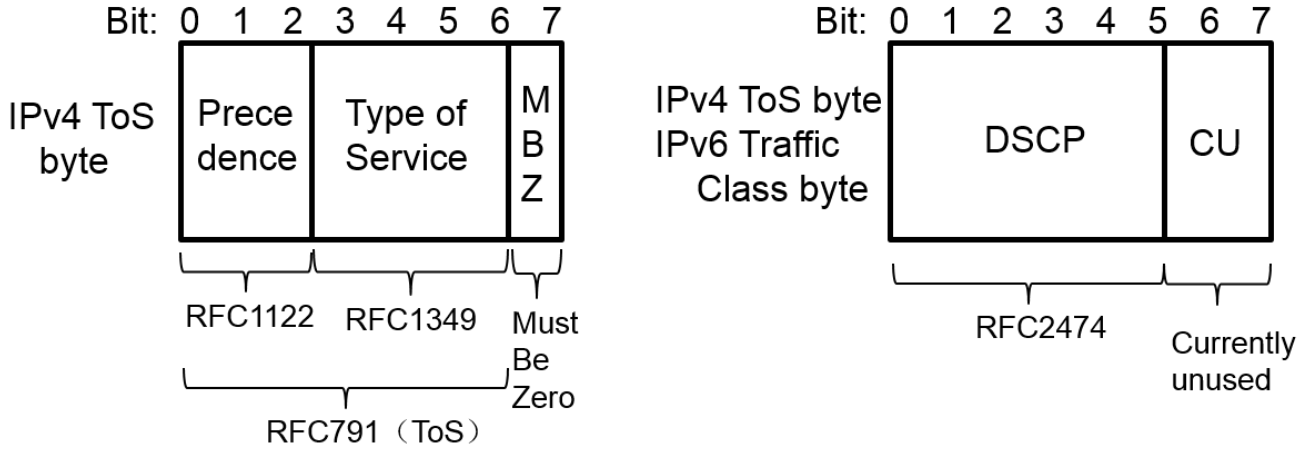


As shown in Figure 2-3, the 4-byte 802.1Q header contains 2-byte Tag ProtocolIdentifier (TPID) whose value is 0x8100 and 2-byte Tag ControlInformation (TCI). The first three bits of the TCI indicate the 802.1P priority.

IP priority (IP PRE) and DSCP priority

The priorities of IP packets are identified by the IP PRE and DSCP priority. The Type Of Service (ToS) field of the IPv4 header comprises 8 bits; where the first three bits indicate the IP precedence (IP PRE), ranging from 0 to 7. RFC 2474 redefines the ToS field of the IPv4 header, which is called the Differentiated Services (DS) field. The Differentiated Services Code Point (DSCP) priority is identified by the first 6 bits (bits 0 to 5) of the DS field, and by the first 6 bits of the Traffic Class field in the IPv6 header. Figure 2-4 shows the locations of the IP PRE and DSCP priorities in IPv4/IPv6 packets.

Figure 2- 4



CoS

Class of Service (COS). FS products convert packet priorities into CoS values to identity the local priorities of the packets and determine the input queue ID when packets are sent from the output interface.

Overview

Feature	Description
Stream Classification	Stream classification uses certain rules to identify packets with same characteristics and is the prerequisite and basis for distinguishing network services.
Priority Labeling and Mapping	Label packet priorities with specified values and map the values to corresponding CoS values.
Traffic Supervision	Supervise the specification of traffic flowing into a network, limit the traffic within a reasonable range, and discard the traffic out of the limit or modify the priority of the traffic.
Congestion Management	Determine the sequence of data packets sent from an interface based on the priorities of the data packets and ensure that key services can be processed in time when congestion occurs.

Congestion Mitigation	Monitor the usage of the output interface queue and reduce the network load by actively discarding packets and adjusting the network traffic when network congestion occurs.

2.3.1 Stream Classification

Stream classification uses certain rules to identify packets with same characteristics and is the prerequisite and basis for distinguishing network services. Stream classification rules are used to distinguish different packets in the network and specify different QoS parameters for packets at different service levels.

Working Principle

Stream classification rules can be matching the PRE or DSCP priorities of IP packets or classifying packets by identifying packet content through an ACL. You can define the binding between multiple streams and stream behaviors by using commands to form policies which can be applied to interfaces for stream classification and processing.

QoS policy

A QoS policy comprises three elements: class, stream behavior and policy.

- Class

A class identifies streams and comprises the class name and class rules. You can define the class rules by using commands to classify packets.

- Stream behavior

Stream behaviors define the QoS actions taken for packets, including priority labeling and traffic supervision for packets.

- Policy

A policy binds a specific class and specific stream behaviors and comprises the policy name, names of the classes bound, and stream behaviors. You can bind a specified class and stream behaviors by using a QoS policy and apply the policy to one or more interfaces.

QoS logical interface group

You can specify a series of interfaces as a QoS logical interface group (including both APs and Ethernet interfaces) and associate polices with the logical interface group for QoS processing. Take rate limit for stream behaviors for example. For packets that meet the rate limit conditions, all interfaces in the same logical interface group share the bandwidth specified by the policy.

Related Configuration

Creating a class

No class is defined by default.

You can run the **class-map** command to create a class and enter the class configuration mode.

Matching an ACL

No rules are defined for a class by default.

In the class configuration mode, you can run the **match acess-group** command to define a class rule as matching an ACL. You need to create ACL rules first.

Matching PRE priorities of IP packets

No rules are defined for a class by default.

In the class configuration mode, you can run the **match ip precedence** command to define a class rule as matching PRE priorities of IP packets. The value range of IP PRE is 0 to 7.

Matching DSCP priorities of IP packets

No rules are defined for a class by default.

In the class configuration mode, you can run the **match ip dscp** command to define a class rule as matching DSCP priorities of IP packets. The value range of DHCP priorities is 0 to 63.

Creating a policy

No policy is defined by default.

You can run the **policy-map** command to create a policy and enter the policy configuration mode.

Associating a class

A policy is not associated with any class by default.

In the policy configuration mode, you can run the **class** command to associate a class and enter the policy-class configuration mode.

Binding a stream behavior

A class is not bound to any stream behavior by default.

In the policy-class configuration mode, you can run the **set** command to modify the CoS, DSCP or VID values of a specified stream; where, the CoS value ranges from 0 to 7, the DSCP value ranges from 0 to 63 and the VID value ranges from 1 to 4094. You can run the **police** command to limit the bandwidth and process streams out of the limit for specified streams. The bandwidth limit ranges are determined by products.

Configuring a logical interface group

No logical interface group is defined and an interface is not added to any logical interface group by default.

In the global configuration mode, you can run the **virtual-group** command to create a logical interface group. In the interface configuration mode, you can run the **virtual-group** command to add an interface to a logical interface group. If this logical interface group is not created, you can create the logical interface group and add the interface to the group. You can create 128 logical interface groups, ranging from 1 to 128.

Applying a policy to an interface

No policy is applied to an interface by default.

In the interface configuration mode, you can run the **service-policy** command to apply a policy in the input/output directions of the interface. In the global configuration mode, you can run the **service-policy** command to apply a policy in the input/output directions of all interfaces.

2.3.2 Priority Labeling and Mapping

Priorities are used to label the scheduling weights of packets or the priorities of the packets in forwarding. Different packet types have different priority types including 802.1P(PRI), IP PRE and DSCP priorities. Priority labeling and mapping refer to labeling packet priorities with specified values and mapping the values to corresponding CoS values.

Working Principle

After data streams of packets enter a device interface, the device assigns priorities to the packets based on the trust mode configured for the interface. The following describes several trust modes:

- When the interface trust mode is untrust, which means not trusting the priority information carried in packets:

Modify the CoS value according to the default CoS value (0, which is configurable), COS-DSCP mapping table and DSCP-COS mapping table of the interface and put the packets into queues based on the final CoS value. For output packets carrying the 802.1Q tag, the packet priority will be modified to the corresponding CoS value.

- When the interface trust mode is trusting CoS:

For packets carrying the 802.1Q tag, modify the CoS value according to the PRI value, CoS-DSCP mapping table, and DSCP-CO mapping table, and put the packets into queues based on the final CoS value. For output packets carrying the 802.1Q tag, the packet priority will be modified to the corresponding CoS value.

For packets not carrying the 802.1Q tag, modify the CoS value according to the default CoS value (0, which is configurable), COS-DSCP mapping table and DSCP-COS mapping table of the interface, and put the packets into queues based on the final CoS value. For output packets carrying the 802.1Q tag, the packet priority will be modified to the corresponding CoS value.

- When the interface trust mode is trusting DSCP:

For non-IP packets, the processing is the same as that for trusting CoS.

For IP packets, modify the CoS value according to the DSCP value of the packets and the DSCP-CoS mapping table and put the packets into queues based on the final CoS value.

- When the interface trust mode is trusting IP PRE:

For non-IPv4 packets, the processing is the same as that for trusting CoS.

For IPv4 packets, obtain and modify the DSCP priority of the packets according to the IP PRE value of the packets and the IP-PRE-DSCP mapping table, obtain the CoS value according to the DSCP-CoS mapping table, and then put the packets into queues based on the final CoS value.

- When the trust mode and the applied policy of an interface work together:

When the trust mode and the applied policy of an interface work together, the trust mode has a lower priority than the policy and the CoS priority can be obtained according to the DSCP-CoS mapping table.

If a policy is applied to the interface but the policy does not has a configuration for modifying the DSCP and CoS values, the processing will be performed based on the trust mode of the interface.

Related Configuration

Configuring the trust mode of an interface

The default trust mode of an interface is untrust.

In the interface configuration mode, run the **mls qos trust** command to modify the trust mode. The trust mode can be trusting CoS, trusting DSCP or trusting IP PRE.

Configuring the default CoS value of an interface

The default CoS value of an interface is 0.

In the interface configuration mode, run the **mls qos cos** command to modify the default CoS value of the interface, which ranges from 0 to 7.

Labeling the priority of streams

The priorities of streams are not relabeled by default.

In the policy-class configuration mode, run the **set** command to modify the CoS, DSCP and VID values of streams. The CoS value ranges from 0 to 7; the DSCP value ranges from 0 to 63; the VID value ranges from 1 to 4094.

Configuring CoS-to-DSCP Map

By default, the CoS values 0, 1, 2, 3, 4, 5, 6 and 7 are mapped to the DSCP values 0, 8, 16, 24, 32, 40, 48 and 56 respectively.

Run the **mls qos map cos-dscp** command to configure the COS-DSCP mapping. The DSCP value ranges from 0 to 63.

Configuring DSCP-to-CoS Map

By default, DSCP 0 to 7 are mapped to CoS 0, DSCP 8 to 15 mapped to CoS 1, DSCP 16 to 23 mapped to CoS2, DSCP 24 to 31 mapped to CoS 3, DSCP 32 to 39 mapped to CoS 4, DSCP 40 to 47 mapped to CoS 5, DSCP 48 to 55 mapped to CoS 6, and DSCP 56 to 63 mapped to CoS 7.

Run the **mls qos map dscp-cos** command to configure the DSCP-CoS mapping. The CoS value ranges from 0 to 7 and the DSCP value ranges from 0 to 63.

Configuring IP-PRE-to-DSCP Map

By default, the IP PRE values 0, 1, 2, 3, 4, 5, 6 and 7 are mapped to the DSCP values 0, 8, 16, 24, 32, 40, 48 and 56 respectively.

Run the **mls qos map ip-prec-dscp** command to configure the IP PRE-DSCP mapping. The DSCP value ranges from 0 to 63.

2.3.3 Traffic Supervision

Supervise the specification of traffic flowing into a network, limit the traffic within a reasonable range, and discard the traffic out of the limit or modify the priority of packets. In addition, the total traffic of an interface can be monitored and the traffic out of the limit will be discarded.

Working Principle

Traffic supervision is used to monitor the specification of traffic flowing into a network and conduct preset supervision actions based on different assessment results. These actions can be:

- Forwarding: Normally forward packets within the traffic limit.
- Discarding: discard packets out of the traffic limit.
- Changing the priority and forwarding: modify the priorities of packets out of the traffic limit and then forward the packets.

Directly discard packets out of the total traffic limit of an interface.

Related Configuration

Configuring the action to be conducted for traffic out of limit

No action to be conducted for traffic out of limit is configured by default.

In the policy-class configuration mode, run the **police** command to configure the action to be conducted for traffic out of limit to discarding traffic out of limit, or modifying the CoS value or DSCP value. The traffic limit range is determined by products. When the traffic is out of the limit, you can modify the CoS value in the range of 0 to 7 and the DSCP value in the range of 0 to 63.

Configuring the total traffic limit for an interface

The total traffic limit for an interface is not configured by default.

In the interface configuration mode, run the **rate-limit** command to configure the total traffic limit for an interface in the input and output directions. The traffic limit range is determined by products.

2.3.4 Congestion Management

When the receiving rate of packets exceeds the sending rate of packets, congestion will occur on the sending interface. If no sufficient buffer is provided to store these packets, the packets may be lost. The congestion management mechanism determines the sequence of data packets to be sent from an interface based on the priorities of the data packets. The congestion management function allows for congestion control by increasing the priorities of important data packets. When congestion occurs, the important data packets are sent in higher priorities to ensure that key services are implemented in time.

Working Principle

A queue scheduling mechanism is used for congestion management and the process is as follows:

- After each packet passes all QoS processing in a switch, the packet will obtain a CoS value finally.
- At the output interface, the device classifies the packets into corresponding sending queues based on the CoS values.
- The output interface selects packets in a queue for sending based on various scheduling policies (SP, WRR, DRR, WFQ, SP+WRR, SP+DRR and SP+WFQ).

Scheduling policy

The queue scheduling policies include SP, WRR, DRR, WFQ, SP+WRR, SP+DRR, and SP+WFQ.

- Strict-Priority (SP) scheduling means scheduling packets strictly following queue IDs. Before sending packets each time, check whether a queue with the first priority has packets to be sent. If yes, the packets in this queue are sent first. If not, check whether a queue with the second priority has packets. Follow the same rules for packets in other queues.
- Weighted Round Robin (WRR) scheduling means scheduling queues in turn to ensure that all queues have certain service time. For example, a 1000 Mbps interface has 8 output queues. The WRR configures a weighted value (5, 5, 10, 20, 20, 10, 20 and 10, which indicate the proportions of obtained resources) for each queue. This scheduling method ensures that a queue with the lowest priority is assigned

with at least 50 Mbps bandwidth, which avoids that packets in the queue with the lowest priority are not served for long time when the SP scheduling method is used.

- Deficit Round Robin (DRR) scheduling is similar to the WRR, but applies weight values based on bytes, but not based on time slices.
- Weighted Fair Queueing (WFQ) scheduling provides dynamic and fair queuing and applies weighted values based on bytes, similar to the DRR. When encountering an empty queue, the DRR will shift to the next queue for transmission immediately. If a queue misses its transmission time, the queue must wait for the next time, which is the difference between the WFQ and DRR; therefore, the WFQ is more suitable for processing data packets with variable lengths than the DRR.
- SP+WRR scheduling means configuring the SP scheduling for one or more sending queues and configuring the WRR scheduling for the other queues. Among SP queues, only after all packets in the SP queue with the first priority are sent, the packets in the SP queue with the second priority can be sent. Among SP and WRR queues, only after the packets in all SP queues are sent, the packets in WRR queues can be sent.
- SP+DRR scheduling means configuring the SP scheduling for one or more sending queues and configuring the DRR scheduling for the other queues. Among SP queues, only after all packets in the SP queue with the first priority are sent, the packets in the SP queue with the second priority can be sent. Among SP and DRR queues, only after the packets in all SP queues are sent, the packets in DRR queues are sent.
- SP+WFQ scheduling means configuring the SP scheduling for one or more sending queues and configuring the WFQ scheduling for the other queues. Among SP queues, only after all packets in the SP queue with the first priority are sent, the packets in the SP queue with the second priority can be sent. Among SP and WFQ queues, only after the packets in all SP queues are sent, the packets in WFQ queues can be sent.

QoS multicast queue

On some products, interface queues are classified into unicast queues and multicast queues. There are 8 unicast queues. All known unicast packets enter corresponding unicast queues for forwarding based on their priorities. There are 1 to 8 multicast queues (depending on products. Certain products do not support multicast queues). Except for known unicast packets, all packets (such as broadcast packets, multicast packets, unknown unicast packets, and mirroring packets) enter corresponding multicast queues for forwarding based on their priorities. Similar to unicast queues, you can configure priority mappings and scheduling algorithms for multicast queues. The **Cos-to-Mc-Queue** command can be used to configure mapping from priorities to multicast queues. At present, multicast queues support the SP, WRR and SP+WRR scheduling algorithms.

Scheduling policy and round robin weight for output queues on an interface

The scheduling policies and round robin weight for output queues are based on global configurations. Some products support both global configurations and interface-based configurations. Interface-based configurations have higher priorities than global configurations. The global scheduling policy works with the corresponding global round robin weight whereas the interface scheduling policy works with the interface round robin weight. If only the global scheduling policy or interface scheduling policy is configured but no corresponding round robin weights are configured, the default round robin weights will work with the scheduling policy.

Queue bandwidth

Some products allow for configuring the guaranteed minimum bandwidth and the limited maximum bandwidth for a queue. A queue configured with the guaranteed minimum bandwidth ensures that the bandwidth for this queue is not smaller than the configured value. A queue configured with the limited maximum bandwidth ensures that the bandwidth for this queue is not greater than the configured value and packets out of the bandwidth limit will be discarded. The bandwidth limits for unicast and multicast queues are configured together on some products whereas configured separately on some other products. In addition, some products allow for configuring bandwidth only for unicast queues.

Related Configuration

Configuring CoS-to-Queue Map

By default, the CoS values 0, 1, 2, 3, 4, 5, 6 and 7 are mapped to the queues 1, 2, 3, 4, 5, 6, 7 and 8 respectively.

Run the **priority-queue cos-map** command to configure the CoS-to-queue mapping. The CoS value ranges from 0 to 7 and the queue value ranges from 1 to 8.

Configuring the scheduling policy for an output queue

By default, the scheduling policy for a global output queue is WRR.

Run the **mls qos scheduler** command to configure the output scheduling policy for a queue. Configurable scheduling policies include SP, WRR, DRR and WFQ. You can also run the **priority-queue** command to configure the scheduling policy as SP.

Configuring the round robin weight corresponding to the WRR scheduling policy for an output queue

By default, the weight of a global queue is 1:1:1:1:1:1:1:1.

Run the **wrr-queue bandwidth** command to configure the round robin weight corresponding to the WRR scheduling policy for an output queue. The configurable weight range is determined by products.

A higher weight means longer output time.

Configuring the round robin weight corresponding to the DRR scheduling policy for an output queue

By default, the weight of a global queue is 1:1:1:1:1:1:1:1.

Run the **drr-queue bandwidth** command to configure the round robin weight corresponding to the DRR scheduling policy for an output queue. The configurable weight range is determined by products.

A higher weight means more packet bytes that can be sent.

Configuring the round robin weight corresponding to the WFQ scheduling policy for an output queue

By default, the weight of a global queue is 1:1:1:1:1:1:1:1.

Run the **wfq-queue bandwidth** command to configure the round robin weight corresponding to the WFQ scheduling policy for an output queue. The configurable weight range is determined by products.

Configuring the bandwidth for a queue

Run the **qos queue** command to configure the guaranteed minimum bandwidth and the limited maximum bandwidth for each queue. The queue value ranges from 1 to 8 and the guaranteed minimum bandwidth and limited maximum bandwidth value ranges are determined by products. Supported queue types are determined by products.

2.3.5 Congestion Mitigation

Monitor the usage of the output interface queue and reduce the network load by actively discarding packets and adjusting the network traffic when network congestion occurs.

Working Principle

Mitigate congestion by effectively monitoring the network traffic and forecasting occurrence of congestion. Packets need to be discarded to mitigate congestion. Discarding policies include Tail-Drop, Random Early Detection (RED), and Weighted Random Early Detection (WRED).

Tail-Drop

Traditional packet loss policies include Tail-Drop. Tail-Drop is effective for all traffic and cannot distinguish service levels. When congestion occurs, data packets at the tail of a queue will be discarded until the congestion is removed.

RED and WRED

Hosts running TCP will decrease the rate of sending packets to respond to massive packet loss. After congestion is removed, the hosts increase the rate of sending packets. In this way, Tail-Drop may cause TCP Global Synchronization. When a queue discards multiple TCP packets simultaneously, multiple TCP connections enter the congestion mitigation and slow startup state simultaneously, and the traffic

is reduced and adjusted. When congestion is removed, traffic peaks may appear. The process repeats constantly, the network traffic goes up and down suddenly, and the line traffic always fluctuates between the lowest quantity and the highest quantity. When TCP global synchronization occurs, the connection bandwidth cannot be adequately used, which causes bandwidth waste.

To avoid this circumstance, you can use the RED/WRED packet discarding policy. This policy provides a mechanism for discarding packets in random, which avoids TCP global synchronization. When packets of a TCP connection are discarded and sent at a lower rate, packets of other TCP connections are still sent at higher rates. In this way, there are always some TCP connections whose packets are sent at higher rates, which increases the utilization of line bandwidth.

When WRED is used, you can set the lower threshold value and maximum discarding probability for a queue. When the queue length is smaller than the lower threshold, WRED does not discard packets. When the queue length is between the higher and lower thresholds, WRED discards packets in random (the longer the queue length, the higher probability of packet discarding. There is a maximum discarding probability). When the queue length is greater than the higher threshold value, WRED discards packets at the maximum discarding probability.

Different from RED, WRED uses priorities to distinguish discarding policies. RED is a special example of WRED. When all CoS values of an interface are mapped to the same lower and higher threshold values, WRED becomes RED.

Related Configuration

Enabling the WRED function

The default packet discarding policy is Tail-Drop.

You can run the **queueing wred** command to enable the WRED function.

Configuring the lower threshold value

When 2 groups of lower thresholds in the unit of percentage are supported, the default values are 100 and 80 (the number of threshold value groups are determined by products).

In the interface configuration mode, you can run the **wrr-queue random-detect min-threshold** command to configure the lower thresholds in the unit of percentage for packets discarded by WRED in each queue. The queue value ranges from 1 to 8. The lower threshold value ranges from 1 to 100.

Configuring the maximum discarding probability

When 2 groups of maximum discarding probabilities are supported, the default values are 100 and 80 (the number of threshold value groups are determined by products).

In the interface configuration mode, you can run the **wrr-queue random-detect probability** command to configure the maximum discarding probabilities for packets discarded by WRED in each queue. The queue value ranges from 1 to 8. The maximum discarding probability ranges from 1 to 100.

Configuring the CoS-to-threshold mapping

By default, all CoS values are mapped to the first group of threshold values (the number of threshold groups is determined by products).

In the interface configuration mode, you can run the **wrr-queue cos-map** command to configure the CoS-to-threshold group mapping. The CoS value ranges from 0 to 7 and the number of threshold groups is determined by products. Multiple groups of lower threshold values and maximum discarding probabilities can be configured. By configuring the CoS-to-threshold group mapping, you can select the effective threshold group mapped to a CoS value, for example, CoS 0 mapped to the first threshold group, and CoS 1 mapped to the second threshold group. If the packets of CoS 0 and 1 are added to queue 1 for scheduling, the packets of CoS 0 are processed based on

the lower threshold values and maximum discarding probabilities in the first group and the packets of CoS 1 are processed based on the lower threshold values and maximum discarding probabilities of the second group.

When all CoS values of an interface are mapped to the same group of threshold values, the enabled WRED becomes RED.

2.4 Configuration

Configuration	Description and Command	
Configuring Stream Classification	(Optional) It is used to create stream classification information.	
	class-map	Creates a class.
	match access-group	Matches ACL rules.
	match ip precedence	Matches the PRE priorities of IP packets.
	match ip dscp	Matches the DSCP priorities of IP packets.
	policy-map	Creates a policy.
	class	Associates a class.
	police	Binds the bandwidth limit for streams and the action for processing packets out of the limit.
	set	Binds the behaviors for modifying the CoS, DSCP and VID values of streams.
	virtual-group	Creates a logical interface group and adds interfaces to the logical interface group.
	service-policy	Applies a policy to an interface.
Configuring Priority Labeling and Mapping for Packets	(Optional) It is used to configure the trust mode, default CoS value and various mappings for an interface.	
	mls qos trust	Modifies the trust mode of an interface.
	mls qos cos	Modifies the default CoS value of the interface.
	mls qos map cos-dscp	Configures the CoS-to-DSCP mapping.
	mls qos map dscp-cos	Configures the DSCP-to-CoS mapping.
	mls qos map ip-precedence-dscp	Configures the IP PRE-to-DSCP mapping.
Configuring Interface Rate Limit	(Optional) It is used to configure the rate limit for an interface.	
	rate-limit	Configures the traffic limit for an interface.
Configuring Congestion Management	(Optional) It is used to configure the CoS-to-queue mapping, queue scheduling policies and round robin weight.	
	priority-queue cos-map	Configures the CoS-to-queue mapping.
	priority-queue	Configures the output scheduling policy for a queue to SP.
	mls qos scheduler	Configures the output scheduling policy for a queue.
	wrr-queue bandwidth	Configures the round robin weight corresponding to the WRR scheduling policy for an output queue.

Configuration	Description and Command	
	drr-queue bandwidth	Configures the round robin weight corresponding to the DRR scheduling policy for an output queue.
	wfq-queue bandwidth	Configures the round robin weight corresponding to the WFQ scheduling policy for an output queue.
	qos queue bandwith	Configures the guaranteed minimum bandwidth and limited maximum bandwidth for a queue.

2.4.1 Configuring Stream Classification

Configuration Effect

- Create a class and match classification rules.
- Create a policy, bind a class and stream behaviors, and associate with an interface.

Notes

- The class and policy names cannot comprise more than 31 characters.
- Interface configurations allow for only AP and Ethernet interface configurations. Certain products support policies applied to SVI interfaces through the **service-policy** command. When both physical interfaces and SVI interfaces are configured with policies, the priority of the physical interfaces is higher than that of the SVI interfaces.
- If run the **service-policy** command in global configuration mode, policies will be applied to all interfaces which can be configured with policies.

Configuration Steps

Creating a class and matching ACL rules

- Optional.
- Create a class. In the class configuration mode, match ACL, IP PRE or DSCP.

Creating a policy

- Optional.
- Create a policy. In the policy configuration mode, bind the class and stream behaviors.

Creating a logical interface group and adding interfaces to the logical interface group

- Optional.
- Create a logical interface group and add interfaces to the logical interface group.

Applying a policy to an interface

- Optional.
- Associate a configured policy with a specified interface or logical interface group.

Verification

- Run the **show class-map** command to check whether the class is successfully created and whether rules are successfully matched.
- Run the **show policy-map** command to check whether the policy is successfully created and whether the class and stream behaviors are successfully bound.
- Run the **show mls qos interface** command to check whether the interface is associated with the policy.
- Run the **show virtual-group** command to check the interfaces in the logical interface group.
- Run the **show mls qos virtual-group** command to check whether the logical interface group is associated with the policy.

Related Commands

Creating a class

Command	**class-map** *class-map-name*
Parameter Description	*class-map-name*: Indicates the name of a class to be created. The name cannot comprise more than 31 characters.
Command Mode	Global configuration mode
Usage Guide	-

Matching an ACL

Command	**match access-group** *access-list-number*
Parameter Description	*access-list-number*: Indicates the number of the ACL to be matched.
Command Mode	Class configuration mode
Usage Guide	-

Matching PRE of IP packets

Command	**match ip precedence** *precedence-value*… [*precedence-value*…]
Parameter Description	*precedence -value*: Indicates the IP PRE to be matched, ranging from 0 to 7.
Command Mode	Class configuration mode
Usage Guide	-

Matching DSCP of IP packets

Command	**match ip dscp** *dscp-value*… [*dscp-value*…]
Parameter Description	*dscp -value*: Indicates the DSCP to be matched, ranging from 0 to 63.
Command Mode	Class configuration mode
Usage Guide	-

Creating a policy

Command	**policy-map** *policy-map-name*
Parameter Description	*policy-map-name*: Indicates the name of a policy to be created. The name cannot comprise more than 31 characters.
Command Mode	Global configuration mode
Usage Guide	-

Associating a class

Command	**class** *class-map-name*

Parameter Description	*class-map-name*: Indicates the name of a class to be associated.
Command Mode	Policy configuration mode
Usage Guide	-

Binding the behaviors for modifying the CoS, DSCP and VID values of streams

Command	**set** { **ip dscp** *new-dscp* | **cos** *new-cos* | **vid** *new-vid* }
Parameter Description	**ip dscp** *new-dscp*: Changes the DSCP value of streams to new-dscp, ranging from 0 to 63. **cos** *new-cos*: Changes the CoS value of streams to new-cos, ranging from 0 to 7. **vid** *new-vid*: Changes the VLAN ID of streams to new-vid, ranging from 1 to 4094.
Command Mode	Class configuration mode
Usage Guide	-

Binding the bandwidth limit for streams and the action for processing packets out of the limit

Command	**police** *rate-bps burst-byte* [**exceed-action** { **drop** | **dscp** *new-dscp* | **cos** *new-cos* [**none-tos**] }]
Parameter Description	*rate-bps*: Indicates the bandwidth limit per second (KBits). The value range is determined by products. *burst-byte*: Indicates the burst traffic limit (Kbytes). The value range is determined by products. **drop**: Discards packets out of the bandwidth limit. **dscp** *new-dscp*: Changes the DSCP value of packets out of the bandwidth limit to new-dscp, ranging from 0 to 63. **cos** *new-cos*: Changes the CoS value of packets out of the bandwidth limit to new-cos, ranging from 0 to 7. **none-tos**: Does not change the DSCP value of packets when changing the CoS value of the packets.
Command Mode	Class configuration mode
Usage Guide	-

Creating a logical interface group and adding interfaces to the logical interface group

Command	**virtual-group** *virtual-group-number*
Parameter Description	*virtual-group-number*: Indicates the logical interface group number, ranging from 1 to 128.
Command Mode	Create the logical interface group in the global configuration mode, add the interface to the logical interface group in the interface configuration mode. If no logical interface group exists, you need to create a logical interface group first and then add interfaces to the logical interface group.
Usage Guide	-

Applying a policy to an interface

Command	**service-policy** { **input** | **output** } *policy-map-name*
Parameter Description	**input**: Indicates the input direction of the interface. **output**: Indicates the output direction of the interface. *policy-map-name*: Indicates the name of the policy applied to the interface.
Command Mode	Interface configuration mode/Global configuration mode

Usage Guide	-

Configuration Example

Creating four stream classes and matching ACL, IP PRE and DSCP

Configuration Steps	● Create ACL rules. ● Create four stream classes and match ACL, IP PRE and DSCP.
	FS# configure terminal FS(config)# access-list 11 permit host 192.168.23.61
	FS(config)# class-map cmap1 FS(config-cmap)# match access-group 11 FS(config-cmap)# exit FS(config)# class-map cmap2 FS(config-cmap)# match ip dscp 21 FS(config-cmap)# exit FS(config)# class-map cmap3 FS(config-cmap)# match ip precedence 5 FS(config-cmap)# exit
Verification	● Check whether the created ACL rules and stream class rules are successful.
	FS# show access-lists ip access-list standard 11 10 permit host 192.168.23.61
	FS# show class-map Class Map cmap1 Match access-group 11 Class Map cmap2 Match ip dscp 21 Class Map cmap3 Match ip precedence 5

Creating a policy, binding a class and stream behaviors, and associating with an interface

Configuration Steps	● Create the stream class cmap1, and match packets whose DSCP value is 18. Create cmap2 and match packets whose IP PRE is 7. ● Create the policy pmap1, associate the policy with cmap1, and bind the behavior of changing the CoS value of the stream to 6. Associate the policy with cmap2, bind the behavior of changing the DSCP value of the stream to 16, limiting the traffic per second within 10,000 Kbits and trigger traffic within 1024 Kbits per second, and changing the DSCP value for traffic out of limit to 7. ● Apply the policy pmap1 to the output direction of the interface gigabitEthernet 0/0. ● Create virtual logical group 1, add the interfaces gigabitEthernet 0/1 and gigabitEthernet 0/2 to the group, and apply the policy pmap1 to the input interface of the virtual logical group.

	FS# configure terminal FS(config)# class-map cmap1 FS(config-cmap)# match ip dscp 18 FS(config-cmap)# exit FS(config)# class-map cmap2 FS(config-cmap)# match ip precedence 7 FS(config-cmap)# exit
	FS(config)# policy-map pmap1 FS(config-pmap)# class cmap1 FS(config-pmap-c)# set cos 6 FS(config-pmap-c)# exit FS(config-cmap)# class cmap2 FS(config-pmap-c)# set ip dscp 15 FS(config-pmap-c)# police 10000 1024 exceed-action dscp 7 FS(config-pmap-c)# exit FS(config-pmap)# exit
	FS(config)# interface gigabitEthernet 0/0 FS(config-if-GigabitEthernet 0/0)# service-policy output pmap1 FS(config-if-GigabitEthernet 0/0)# exit
	FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# virtual-group 1 FS(config-if-GigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2)# virtual-group 1 FS(config-if-GigabitEthernet 0/2)# exit FS(config)# virtual-group 1 FS(config-VirtualGroup)# service-policy input pmap1 FS(config-VirtualGroup)# exit
Verification	● Check whether the stream class rules are successfully created. ● Check whether the policy is successfully created, and whether the stream and stream behaviors are successfully bound. ● Check whether the policy is applied to the interface. ● Check whether the logical interface group is successfully created, whether interfaces are successfully associated and whether the policy is successfully applied to the interface.
	FS# show class-map Class Map cmap1 Match ip dscp 18 Class Map cmap2 Match ip precedence 7
	FS# show policy-map Policy Map pmap1 Class cmap1

	set cos 6 Class cmap2 set ip dscp 15 police 10000 1024 exceed-action dscp 7
	FS# show mls qos interface gigabitEthernet 0/0 Interface: GigabitEthernet 0/0 Ratelimit input: Ratelimit output: Attached input policy-map: Attached output policy-map: pmap1 Default trust: none Default cos: 0
	FS# show virtual-group 1 virtual-group member ---------- ---------- 1 Gi0/1 Gi0/2 FS# show mls qos virtual-group 1 Virtual-group: 1 Attached input policy-map: pmap1

2.4.2 Configuring Priority Labeling and Mapping for Packets

Configuration Effect

- Configure the trust mode and default CoS value of an interface.
- Configure the CoS-to-DSCP, DSCP-to-CoS, and IP-PRE-to-DSCP mappings.

Notes

- Interface configurations allow for only AP and Ethernet interface configurations.

Configuration Steps

Configuring the trust mode and default CoS value of an interface

- Optional.
- In the interface configuration mode, configure the trust mode and default CoS value of an interface.

Configuring the CoS-to-DSCP, DSCP-to-CoS, and IP-PRE-to-DSCP mappings

- Optional.
- Configure various mappings.

Verification

- Run the **show mls qos interface** command to display the trust mode and default CoS value of the interface.
- Run the **show mls qos maps** command to display the CoS-to-DSCP, DSCP-to-CoS and IP-PRE-to-DSCP mappings.

Related Commands

Configuring the trust mode of an interface

Command	**mls qos trust** { **cos** | **ip-precedence** | **dscp** }
Parameter Description	**cos**: Configures the trust mode of an interface to CoS. **ip-precedence**: Configures the trust mode of an interface to IP PRE. **dscp**: Configures the trust mode of an interface to DSCP.

Command Mode	Interface configuration mode
Usage Guide	-

Configuring the default CoS value of an interface

Command	**mls qos cos** *default-cos*
Parameter Description	*default-cos*: Configures the default CoS value, ranging from 0 to 7. The default value is 0.
Command Mode	Interface configuration mode
Usage Guide	-

Configuring CoS-to-DSCP MAP

Command	**mls qos map cos-dscp** *dscp1…dscp8*
Parameter Description	*dscp1….dscp8*: Indicates the DSCP values mapped to the CoS values. The default CoS values 0~7 are mapped to DSCP 0, 8, 16, 24, 32, 40, 48 and 56 respectively. The DSCP value ranges from 0 to 63.
Command Mode	Global configuration mode
Usage Guide	-

Configuring DSCP-to-CoS MAP

Command	**mls qos map dscp-cos** *dscp-list* **to** *cos*
Parameter Description	*dscp-list*: Indicates the DSCP list mapped to the CoS values. The default DSCP 0~7 are mapped to CoS 0, DSCP 8~15 mapped to CoS 1, DSCP 16~23 mapped to CoS 2, DSCP 24~31 mapped to CoS 3, DSCP 32~39 mapped to CoS 4, DSCP 40~47 mapped to CoS 5, DSCP 48~55 mapped to CoS 6, and DSCP 56~63 mapped to CoS 7. The DSCP value ranges from 0 to 63. *cos*: Indicates the CoS values mapped to the dscp-list, ranging from 0 to 7.
Command Mode	Global configuration mode
Usage Guide	-

Configuring IP-PRE-to-DSCP MAP

Command	**mls qos map ip-prec-dscp** *dscp1…dscp8*
Parameter Description	*dscp1….dscp8*: Indicates the DSCP values mapped to the IP PRE values. The default IP PRE 0~7 are mapped to DSCP 0, 8, 16, 24, 32, 40, 48 and 56 respectively. The DSCP value ranges from 0 to 63.
Command Mode	Global configuration mode
Usage Guide	-

Configuration Example

Configuring the trust mode and default CoS value of an interface

Configuration Steps	● Modify the trust mode of the interface gigabitEthernet 0/0 to DSCP. ● Change the default CoS value of the interface gigabitEthernet 0/1 to 7.

	FS# configure terminal FS(config)# interface gigabitEthernet 0/0 FS(config-if-GigabitEthernet 0/0)# mls qos trust dscp FS(config-if-GigabitEthernet 0/0)# exit FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# mls qos cos 7 FS(config-if-GigabitEthernet 0/1)# exit
Verification	● Check whether the trust mode and default CoS value are successfully configured for the interface.
	FS# show mls qos interface gigabitEthernet 0/0 Interface: GigabitEthernet 0/0 Ratelimit input: Ratelimit output: Attached input policy-map: Attached output policy-map: Default trust: dscp Default cos: 0 FS# show mls qos interface gigabitEthernet 0/1 Interface: GigabitEthernet 0/1 Ratelimit input: Ratelimit output: Attached input policy-map: Attached output policy-map: Default trust: none Default cos: 7

Configuring the CoS-to-DSCP, DSCP-to-CoS, and IP-PRE-to-DSCP mappings

Configuration Steps	● Configure CoS-to-DSCP to map CoS 0, 1, 2, 3, 4, 5, 6, and 7 to DSCP 7, 14, 21, 28, 35, 42, 49, and 56 respectively. ● Configure DSCP-to-CoS to map DSCP 0, 1, 2, 3, and 4 to CoS 4 and DSCP 11, 12, 13 and 14 to CoS 7. ● Configure IP-PRE-to-DSCP to map IP PRE 0, 1, 2, 3, 4, 5, 6, and 7 to DSCP 31, 26, 21, 15, 19, 45, 47, and 61 respectively.
	FS# configure terminal FS(config)# mls qos map cos-dscp 7 14 21 28 35 42 49 56
	FS(config)# mls qos map dscp-cos 0 1 2 3 4 to 4 FS(config)# mls qos map dscp-cos 11 12 13 14 to 7
	FS(config)# mls qos map ip-precedence-dscp 31 26 21 15 19 45 47 61
Verification	● Check whether all mappings are successfully configured.
	FS# show mls qos maps cos-dscp cos dscp --- ---- 0 7 1 14

```
2     21
3     28
4     35
5     42
6     49
7     56
```

```
FS# show mls qos maps dscp-cos
dscp cos        dscp cos        dscp cos        dscp cos
---- ---        ---- ---        ---- ---        ---- ---
 0    4           1    4           2    4           3    4
 4    4           5    0           6    0           7    0
 8    1           9    1          10    1          11    7
12    7          13    7          14    7          15    1
16    2          17    2          18    2          19    2
20    2          21    2          22    2          23    2
24    3          25    3          26    3          27    3
28    3          29    3          30    3          31    3
32    4          33    4          34    4          35    4
36    4          37    4          38    4          39    4
40    5          41    5          42    5          43    5
44    5          45    5          46    5          47    5
48    6          49    6          50    6          51    6
52    6          53    6          54    6          55    6
56    7          57    7          58    7          59    7
60    7          61    7          62    7          63    7
```

```
FS# show mls qos maps ip-prec-dscp
ip-precedence dscp
------------- ----
                0 31
                1 26
                2 21
                3 15
                4 19
                5 45
                6 47
                7 61
```

2.4.3 Configuring Interface Rate Limit

Configuration Effect

- Configure the traffic limit for an interface.

Notes

- The configuration is supported only by Ethernet and aggregate interfaces.

Configuration Steps

Configuring the traffic limit for an interface

- Optional.
- Configure the limit on the traffic and burst traffic for an interface.

Verification

- Run the **show mls qos rate-limit** command to display the rate limit information about the interface.

Related Commands

Configuring the traffic limit for an interface

Command	**rate-limit** { **input** | **output** } *bps burst-size*
Parameter Description	**input**: Indicates the input direction of the interface. **output**: Indicates the output direction of the interface. *bps*: Indicates the bandwidth limit per second (Kbits). The value range is determined by products. *burst-size*: Indicates the burst traffic limit (Kbytes). The value range is determined by products.
Command Mode	Interface configuration mode
Usage Guide	-

Configuration Example

Typical application – Interface rate limit + priority relabeling

Configuration Steps	● For Internet access by using the output interface, configure the output traffic limit on the interface G0/24, and set the bandwidth limit to 102,400 Kbits per second and burst traffic limit to 256 Kbytes per second. ● For the dormitory building, configure the input traffic limit on the interface G0/3, and set the bandwidth limit to 51,200 Kbits per second and burst traffic limit to 256 Kbytes per second. ● For the teaching building, configure the input traffic limit on the interface G0/1, and set the bandwidth limit to 30,720 Kbits per second and burst traffic limit to 256 Kbytes per second. ● For the laboratory, create the class cmap_dscp7 to match DSCP priority 7, create the policy pmap_shiyan to associate with cmap_dscp7, bind the stream behavior of changing the DSCP value for packets whose rates exceed 20M to 16, apply pmap_shiyan to the interface G0/2, and configure the interface to trusting DSCP.
	FS# configure terminal FS(config)# interface gigabitEthernet 0/24 FS(config-if-GigabitEthernet 0/24# rate-limit output 102400 256 FS(config-if-GigabitEthernet 0/24)# exit FS(config)# interface gigabitEthernet 0/3 FS(config-if-GigabitEthernet 0/3# rate-limit input 51200 256 FS(config-if-GigabitEthernet 0/3)# exit FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1# rate-limit input 30720 256 FS(config-if-GigabitEthernet 0/1)# exit
	FS(config)# class-map cmap_dscp7 FS(config-cmap)# match ip dscp 7 FS(config-cmap)# exit FS(config)# policy-map pmap_shiyan FS(config-pmap)# class cmap_dscp7 FS(config-pmap-c)# police 20480 128 exceed-action dscp 16

	FS(config-pmap-c)# exit FS(config-pmap)# exit FS(config)# interface gigabitEthernet 0/2 FS(config-if-GigabitEthernet 0/2# service-policy input pmap_shiyan FS(config-if-GigabitEthernet 0/2)# mls qos trust dscp FS(config-if-GigabitEthernet 0/2)# exit
Verification	● Check whether the interface rate limit is successfully configured. ● Check whether the class and policy are successfully created and successfully applied to the interface.
	FS# show mls qos rate-limit Interface: GigabitEthernet 0/1 rate limit input Kbps = 30720 burst = 256 Interface: GigabitEthernet 0/3 rate limit input Kbps = 51200 burst = 256 Interface: GigabitEthernet 0/24 rate limit output Kbps = 102400 burst = 256
	FS# show class-map cmap_dscp7 Class Map cmap_dscp7 Match ip dscp 7 FS# show policy-map pmap_shiyan Policy Map pmap_shiyan Class cmap_dscp7 police 20480 128 exceed-action dscp 16 FS# show mls qos interface gigabitEthernet 0/2 Interface: GigabitEthernet 0/2 Ratelimit input: Ratelimit output: Attached input policy-map: pmap_shiyan Attached output policy-map: Default trust: dscp Default cos: 0

2.4.4 Configuring Congestion Management

Configuration Effect

- Configure the CoS-to-queue mapping.
- Configure the scheduling policy and round robin weight for an output queue.
- Configure the guaranteed minimum bandwidth and limited maximum bandwidth for a queue.

Notes

- Interface configurations allow for only AP and Ethernet interface configurations.

Configuration Steps

Configuring the CoS-to-unicast and CoS-to-multicast mappings

- Optional.
- Configure the CoS-to-queue mappings. On products supporting multicast queues, you can configure the CoS-to-multicast queue mapping.

Configuring the scheduling policies and round robin weight for unicast and multicast output queues

- Optional.
- Configure the scheduling policy for an output queue and modify the round robin weight. On products supporting multicast queues, you can configure the scheduling policies and round robin weights for multicast queues.

Configuring the guaranteed minimum bandwidth and limited maximum bandwidth for a queue

- Optional.
- Configure the guaranteed minimum bandwidth and limited maximum bandwidth for a queue.

Verification

- Run the **show mls qos queueing** command to display the output queue information.
- Run the **show mls qos scheduler** command to display the scheduling policy for the output queue.
- Run the **show qos bandwidth** command to display the queue bandwidth.

Related Commands

Configuring CoS-to-Queue MAP

Command	**priority-queue cos-map** *qid cos0* [*cos1* [*cos2* [*cos3* [*cos4* [*cos5* [*cos6* [*cos7*]]]]]]]
Parameter Description	*qid*: Indicates the queue ID to be mapped, ranging from 1 to 8. *cos0~cos7*: Indicates the CoS values to be mapped to the qid. The default CoS values 0~7 are mapped to queues 1~8. The value range is 0 to 7.
Command Mode	Global configuration mode
Usage Guide	-

Configuring the scheduling policy for an output queue to SP

Command	**priority-queue**
Parameter Description	**-**
Command Mode	Global configuration mode
Usage Guide	-

Configuring the scheduling policy for an output queue

Command	**mls qos scheduler** { **sp** | **rr** | **wrr** | **drr** | **wfq** }
Parameter Description	**sp**: Sets the scheduling algorithm for an output queue to SP. rr: sets the scheduling algorithm for an output queue to RR. **wrr**: Sets the scheduling algorithm for an output queue to WRR. **drr**: Sets the scheduling algorithm for an output queue to DRR. **wfq**: Sets the scheduling algorithm for an output queue to WFQ.
Command Mode	Global configuration mode
Usage Guide	-

Configuring the scheduling policy and round robin weight for an output queue

Command	{ **drr-queue** | **wrr-queue** | **wfq-queue** } **bandwidth** *weight1...weight8*
Parameter Description	**drr-queue**: Configures the round robin weight corresponding to the DRR scheduling policy for an output queue. **wrr-queue**: Configures the round robin weight corresponding to the WRR scheduling policy for an output queue. **wfq-queue**: Configures the round robin weight corresponding to the WFQ scheduling policy for an output queue. *weight1...weight8*: Indicates the weight of queues 1 to 8. The value range is determined by products. The value 0 indicates that the queue uses the SP scheduling algorithm. The default weight for global/interface queues is 1:1.
Command Mode	Global/Interface configuration mode
Usage Guide	-

Configuring the guaranteed minimum bandwidth and limited maximum bandwidth for a queue

Command	**qos queue** *queue-id* **bandwidth** { **minimum** | **maximum** } *bandwidth*
Parameter Description	*queue-id*: Indicates the queue ID to be configured, ranging from 1 to 8. **minimum** *bandwidth*: Indicates the guaranteed minimum bandwidth Kbps. The value range is determined by products. It is not configured by default. **maximum** *bandwidth*: Indicates the limited maximum bandwidth Kbps. The value range is determined by products. It is not configured by default.
Command Mode	Interface configuration mode
Usage Guide	-

Configuration Example

Configuring the CoS-to-queue mapping and modifying the scheduling policy and its round robin weight

Configuration Steps	● Configure the CoS-to-queue mapping to the mapping from the CoS values 0, 1, 2, 3, 4, 5, 6, and 7 to queues 1, 2, 5, 5, 5, 5, 7, and 8. ● Configure the output scheduling policy for a queue to DRR and the round robin weight to 2:1:1:1:6:6:6:8.
	FS# configure terminal FS(config)# priority-queue cos-map 5 2 3 4 5 FS(config)# mls qos scheduler drr FS(config)# drr-queue bandwidth 2 1 1 1 6 6 6 8
Verification	● Check whether the CoS-to-queue mapping is successfully created, and whether the output scheduling policy and round robin weight are successfully configured for the queue.
	FS# show mls qos scheduler Global Multi-Layer Switching scheduling Deficit Round Robin FS# show mls qos queueing CoS-to-queue map: cos qid --- --- 0 1 1 2

```
2     5
3     5
4     5
5     5
6     7
7     8

wrr bandwidth weights:
qid weights
--- -------
1     1
2     1
3     1
4     1
5     1
6     1
7     1
8     1

drr bandwidth weights:
qid weights
--- -------
1     2
2     1
3     1
4     1
5     6
6     6
7     6
8     8

wfq bandwidth weights:
qid weights
--- -------
1     1
2     1
3     1
4     1
5     1
6     1
7     1
8     1
```

Typical application – Priority relabeling + queue scheduling

Configuration Steps	● Create ACLs for accessing various servers and create classes for matching these ACLs. ● Create policies for associating with the classes and specify new CoS values for packets accessing various servers. Associate the CoS values with the input interfaces for the R&D and market departments and configure the interfaces to trusting CoS. ● Configure the default CoS value for the HR management department interface to the highest priority 7 to ensure that packets from the HR management department are sent in the highest priority. ● Configure the output scheduling policy to WR and the round robin weight to 1:1:1:2:6:1:1:0 for the queues. This means that the SP scheduling algorithm is used for packets of the HR management department, and the packets of the R&D and market departments for accessing the mail database, file database and salary query database are scheduled based on the ratio of 6:2:1.
	FS# configure terminal FS(config)# ip access-list extended salary FS(config-ext-nacl)# permit ip any host 192.168.10.1 FS(config-ext-nacl)# exit FS(config)# ip access-list extended mail FS(config-ext-nacl)# permit ip any host 192.168.10.2 FS(config-ext-nacl)# exit FS(config)# ip access-list extended file FS(config-ext-nacl)# permit ip any host 192.168.10.3 FS(config-ext-nacl)# exit
	FS(config)# class-map salary FS(config-cmap)# match access-group salary FS(config-cmap)# exit FS(config)# class-map mail FS(config-cmap)# match access-group mail FS(config-cmap)# exit FS(config)# class-map file FS(config-cmap)# match access-group file
	FS(config)# policy-map toserver FS(config-pmap)# class mail FS(config-pmap-c)# set cos 4 FS(config-pmap-c)# exit FS(config-pmap)# class file FS(config-pmap-c)# set cos 3 FS(config-pmap-c)# exit FS(config-pmap)# class salary FS(config-pmap-c)# set cos 2 FS(config-pmap-c)# end
	FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# service-policy input toserver FS(config-if-GigabitEthernet 0/1)# mls qos trust cos FS(config-if-GigabitEthernet 0/1)# exit FS(config)# interface gigabitEthernet 0/2

	FS(config-if-GigabitEthernet 0/2)# service-policy input toserver FS(config-if-GigabitEthernet 0/2)# mls qos trust cos FS(config-if-GigabitEthernet 0/2)# exit
	FS(config)# interface gigabitEthernet 0/3 FS(config-if-GigabitEthernet 0/3)# mls qos cos 7
	FS(config)#wrr-queue bandwidth 1 1 1 2 6 1 1 0 FS(config)#mls qos scheduler wrr
Verification	● Check whether the ACLs are successfully created and whether the classes are successfully associated with the ACLs. ● Check whether the policies are successfully created, whether the classes and stream behaviors are successfully bound, and whether policies are successfully applied to the interfaces. ● Check whether the default CoS value is successfully configured for the interface and whether the scheduling policy and the round robin weight are successfully configured.

	FS# show access-lists
	
	ip access-list extended file
	  10 permit ip any host 192.168.10.3
	
	ip access-list extended mail
	  10 permit ip any host 192.168.10.2
	
	ip access-list extended salary
	  10 permit ip any host 192.168.10.1

	FS# show class-map
	
	  Class Map salary
	  Match access-group salary
	  Class Map mail
	  Match access-group mail
	  Class Map file
	  Match access-group file

	FS# show policy-map
	
	  Policy Map toserver
	  Class mail
	  set cos 4
	  Class file
	  set cos 3
	  Class salary
	  set cos 2

	FS# show mls qos interface gigabitEthernet 0/1 Interface: GigabitEthernet 0/1

	Ratelimit input: Ratelimit output: Attached input policy-map: toserver Attached output policy-map: Default trust: cos Default cos: 0 FS# show mls qos interface gigabitEthernet 0/2 Interface: GigabitEthernet 0/2 Ratelimit input: Ratelimit output: Attached input policy-map: toserver Attached output policy-map: Default trust: cos Default cos: 0
	FS# show mls qos interface gigabitEthernet 0/3 Interface: GigabitEthernet 0/2 Ratelimit input: Ratelimit output: Attached input policy-map: Attached output policy-map: Default trust: none Default cos: 7
	FS# show mls qos scheduler Global Multi-Layer Switching scheduling Weighted Round Robin FS# FS#show mls qos queueing CoS-to-queue map: cos qid --- --- 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 wrr bandwidth weights: qid weights --- ------- 1 1 2 1 3 1

```
4     2
5     6
6     1
7     1
8     0

drr bandwidth weights:
qid weights
--- -------
1     1
2     1
3     1
4     1
5     1
6     1
7     1
8     1

wfq bandwidth weights:
qid weights
--- -------
1     1
2     1
3     1
4     1
5     1
6     1
7     1
8     1
```

2.4.5 Configuring Congestion Mitigation

Configuration Effect

- Configure the lower threshold value for WRED. When the length of packets in a queue is smaller than the lower threshold value, WRED does not discard packets.
- Configure the maximum discarding probability. When the length of packets in the queue is between the lower and higher threshold values, WRED discards packets in random. The maximum probability for discarding packets is configured.
- Configure the CoS-to-threshold mapping.

Notes

- Interface configurations allow for only AP and Ethernet interface configurations.

Configuration Steps

Enabling the WRED function

- Optional.
- Enable the WRED function if necessary.

Configuring the lower threshold value

- Optional.
- Configure the lower threshold value if necessary.

Configuring the maximum discarding probability

- Optional.
- Configure the maximum discarding probability if necessary.

Configuring the CoS-to-threshold mapping

- Optional.
- Configure the CoS-to-threshold mapping if necessary.

Verification

- Run the **show queueing wred interface** command to display the WRED configuration.

Related Commands

Enabling the WRED function

Command	**queueing wred**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	-

Configuring the lower threshold value(in the unit of percentage)

Command	**wrr-queue random-detect min-threshold** *queue_id thr1* [*thr2*]
Parameter Description	*queue_id*: Indicates the queue ID for an interface, ranging from 1 to 8. *thrN*: Supports 2 groups of lower threshold values, ranging from 1 to the specified higher threshold.
Command Mode	Interface configuration mode
Usage Guide	Because the maximum value of the configuration range is equal to the current higher threshold, you need to pay attention to the setting of the higher threshold when configuring the lower threshold.

Configuring the maximum discarding probability

Command	**wrr-queue random-detect probability** *queue_id prob1* [*prob2*]

Parameter Description	*queue_id*: Indicates the queue ID for an interface, ranging from 1 to 8. *probN*: Supports 2 groups of maximum discarding probabilities, ranging from 1 to 100.
Command Mode	Interface configuration mode
Usage Guide	-

2.5 Monitoring

Displaying

Description	Command
Displays stream classification information.	**show class-map** [*class-map-name*]
Displays QoS policy information.	**show policy-map** [*policy-map-name* [**class** *class-map-name*]]
Displays the policy applied to an interface.	**show policy-map interface** *interface-id*
Displays logical interface group information.	**show virtual-group** [*virtual-group-number* | **summary**]
Displays the policy applied to a logical interface group.	**show mls qos virtual-group** [*virtual-group-number* | **policers**]
Displays various mappings.	**show mls qos maps** [**cos-dscp** | **dscp-cos** | **ip-prec-dscp**]
Displays interface rate limit information.	**show mls qos rate-limit** [**interface** *interface-id*]
Displays the QoS queue, scheduling policy and round robin weight information.	**show mls qos queueing**
Displays the scheduling information of an output queue.	**show mls qos scheduler**
Displays the QoS information of an interface.	**show mls qos interface** *interface-id* [**policers**]
Displays the bandwidth information of an interface.	**show qos bandwidth** [**interfaces** *interface-id*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the QoS library.	**debug qos lib** [**event** | **message**]
Debugs the QoS communication server.	**debug qos server** [**event** | **message**]
Debugs QoS user command processing.	**debug qos mls**
Debugs VMSUP configurations.	**debug qos vmsup**

3 Configuring MMU

3.1 Overview

The Memory Management Unit (MMU) means that the chip buffer is distributed reasonably so that the switching equipment can better deal with all kinds of burst flows.

Flows not steady all the time and various burst flows exist on the network. When the network flow is steady and the bandwidth is sufficient, all the data flows are processed better; when burst flows exist on the network, data flows may be discarded even if the average flow rate does not exceed the bandwidth.

Data packets that enter the switching equipment are stored in the buffer of switching equipment before being forwarded. Normally, data packets stay for a short period of time in the buffer and will be forwarded in microseconds; when there is a burst flow, if the instantaneous rate of burst flow exceeds the processing capacity of the switching equipment, the data packets that cannot be processed in time will be piled up in the switching equipment and packet loss will take place once the buffer is insufficient. In this case, the MMU can be used to reasonably configure the buffer and allocate different buffer sizes to respective services, with a view to optimizing the network.

3.2 Applications

Application	Description
Configuring Large Buffer Application Based on Egress Queue	An enterprise needs a buffer large enough in the SkyDrive service to avoid packet loss for the service flow.

3.2.1 Configuring Large Buffer Application Based on Egress Queue

Scenario

An enterprise needs a buffer large enough in the SkyDrive service to avoid packet loss for the service flow.

As shown in the following figure, equipment A is connected to 5 clients and 35 service servers, where 15 service servers virtualize 15 front end servers.

The main service flow is as follows:

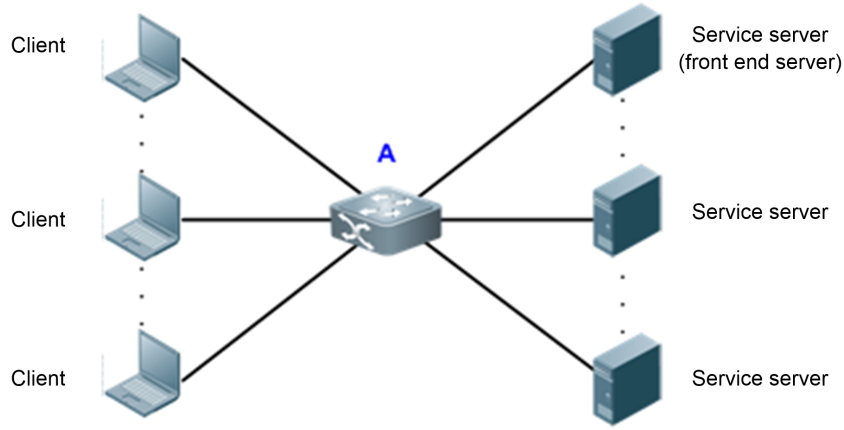
- The client server sends a request packet to the front end server.
- The front end server sends the received request packet to the service server.
- After receiving the request packet, the service server sends a response packet to the front end server.
- After receiving the response packet, the front end server sends it to the client server.
- After receiving the response packet, the client indicates that a session is created successfully.

A many-to-one flow transmission mode exists under this service model:

- The request flows of multiple clients are sent to one front end server.
- The request flows of multiple front end servers are sent to one service server.
- The response flows of multiple service servers are sent to one front end server.
- The response flows of multiple front end servers are sent to one client.

These flows are transmitted through equipment A basically, easily leading to network congestion. Such a problem can be fixed by configuring a large buffer on the equipment.

Figure 3- 1



Deployment

- In all the service ports (namely, the ports connecting clients to servers), configure the shared buffer of the queue where the service is as 100%.
- In all the service ports, configure the minimum value for the guaranteed buffer of the queue not in use.
- In all the ports not in use, configure the minimum value for the guaranteed buffers of all the queues.

For the specific configuration, see the configuration examples in "Configuration".

3.3 Features

Basic Concepts

Cell

Cell is a buffer unit, i.e., the minimum unit for the switching equipment to store packets. The size of each cell varies with the product. One packet can use multiple cells, while one cell can be used by only one packet.

Port group

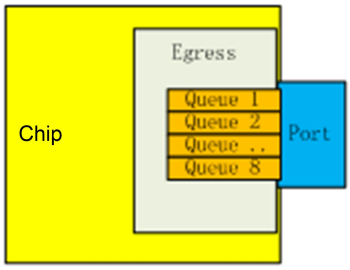
All the ports physically belonging to one switching chip are collectively called a port group, the buffer of switching equipment is managed in the port group. Take the board card M18000_40XS_CB as an example, this version has two switching chips, so there are two port groups. The first 20 ports belong to Port Group 1, and the back 20 ports belong to Port Group 2.

Egress queue

Port egress queues are classified into unicast queues and multicast queues (the number of queues depends on the product). Logically the switching chip is divided into the ingress (incoming direction) and egress (outgoing direction). The egress queue is in the egress

direction. Before packets go out of the egress, the enqueue operation needs to be performed for them at the egress queue. Some of our products implement buffer management based on the egress queue.\

Figure 3- 2



Currently there are three types of egress queue models:

- There are 8 unicast queues and 8 multicast queues at the egress. The well-known unicast packets follow the unicast queue, and all the other packets follow the multicast queue.
- There are 8 unicast queues and 4 multicast queues at the egress. The well-known unicast packets follow the unicast queue, and all the other packets follow the multicast queue.
- There are only 8 queues at the egress, without differentiating unicast and multicast.

Overview

Feature	Description
Buffer Adjustment	The buffer is adjusted based on the queue. It is the foundation of MMU.
Buffer Monitoring	Buffer monitoring actually means monitoring on the use of the buffer capacity, which facilitates buffer adjustment.
Queue Counting	The received and sent packets of each queue are counted so that the buffer adjustment result can be displayed easily.

3.3.1 Configuring Buffer Adjustment

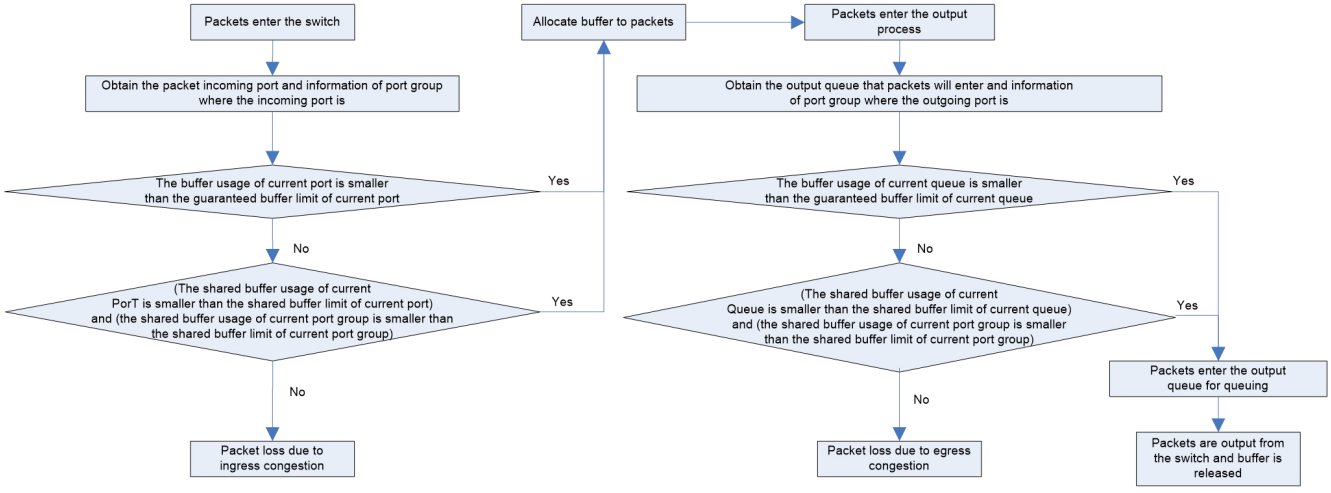
Buffer adjustment means that the queue of each service has different buffer sizes through some adjustment of the queue buffer so that each service is treated differently and services at different priorities are served differently.

Working Principle

Working mechanism of caching in hardware

In terms of hardware, the buffer is managed in the input direction and output direction. The processing mechanism is shown below:

Figure 3- 3



During buffer management, the input direction is adjusted to the maximum value to prevent packet loss in the input direction and make packet loss take place in the output direction. Therefore, adjustment is not opened for the buffer in the input direction, and CLI provides buffer adjustment in the output direction only, including the queue guaranteed buffer and queue shared buffer. Buffer adjustment configures the guaranteed buffer threshold and shared buffer threshold of queues to allocate different buffer sizes to queues.

Guaranteed buffer

Guaranteed buffer is also called exclusive buffer. This part of buffer is distributed based on each queue. The guaranteed buffer of a queue can be used by this queue only. A fixed guaranteed buffer is allocated to each queue by default. This part of queue enables this queue to forward packets at the normal line rate under the stable flow.

Shared buffer

In the total buffer of port group, the remaining part is the total shared buffer after the guaranteed buffer of each queue is deducted. The shared buffer can be used by all the queues. A shared queue threshold can be set for each queue. This threshold restricts the maximum shared buffer quantity that can be used by this queue. When the shared buffer sum configured for each queue in the port group exceeds the total shared queue of port group, the "First Come First Served" buffer occupancy mechanism is adopted.

3.3.2 Configuring Buffer Monitoring

Buffer monitoring implements monitoring on the use amount of each queue and shared buffer, with a view to providing data support for network optimization and reasonable buffer configuration.

Working Principle

Buffer monitoring adopts the polling mode to read the buffer use amount of each queue and the use situation of total buffer regularly and display the buffer use situation of current equipment in real time.

Queue buffer utilization alarm threshold

When the buffer utilization of queue exceeds this threshold, syslog will be printed to remind the user.

3.3.3 Configuring Queue Counting

Queue counting monitors the forwarding and packet loss data of each queue, and push the alarm when packet loses, so as to provide data support for network optimization and reasonable buffer configuration.

Working Principle

The queue adopts the polling mode to read the number of forwarded packets/number of bytes and the number of lost packets/number of bytes of each queue regularly, and then use the data to calculate each kind of statistics of the queue.

3.4 Configuration

Configuration	Description and Command	
Buffer Adjustment	(Optional) It is used to configure buffer.	
	mmu queue-guarantee	Configures guaranteed buffer
	mmu queue-threshold	Configures shared buffer
	mmu buffer-mode	Configures buffer mode
	mmu fc-threshold	Configure flow control threshold based on inbound port
Buffer Monitoring	(Optional) It is used to configure buffer.	
	mmu usage-warn-limit	Configures the buffer utilization alarm threshold

3.4.1 Configuring Buffer Adjustment

Configuration Effect

- Configure guaranteed buffer so that the queue can share this part of buffer exclusively.
- Configure shared buffer so as to control the shared buffer use amount of the queue.

Notes

- Configuration on the interface can be made on the physical port only.

Configuration Steps

Configuring guaranteed buffer

- Optional.
- In the interface mode, use the **mmu queue-guarantee** command to configure guaranteed buffer for each queue and ensure that the buffer configuration range varies with the product.
- Use the **no** or **default** command of this command to restore the default value of buffer.

Command	**mmu queue-guarantee output** { **unicast** } [*queue-id1* [*queue-id2* [*queue-idN*]] **set** *value*
Parameter Description	**output:** performs buffer management on the egress queue **unicast:** performs buffer management on the egress unicast queue *queue-id*: queue ID, in the range from 1 to 8 *value*: number of guaranteed buffers, in cells; the range depends on the product.
Defaults	A fixed number of guaranteed buffers are allocated to each queue by default. The specific configuration depends on the product.
Command	Interface mode

Mode	
Usage Guide	The effective way of this command varies with the equipment and depends on the product.

Configuring buffer mode

- Optional.
- Under the global configuration mode, use the **mmu buffer-mode** command to configure the buffer mode.

Command	**mmu buffer-mode** { **normal** | **burst-enhance | qos-enhance | flowctrl-enhance** }
Parameter Description	**normal:** normal buffer mode **burst-enhance:** Burst enhanced buffer mode **qos-enhance:** QoS enhanced buffer support mode **flowctrl-enhance:** flow control enhanced buffer support mode
Defaults	Normal buffer mode is applied by default.
Command Mode	Global configuration mode
Usage Guide	The effective way of this command varies with the equipment and depends on the product.

Configuring shared buffer

- Optional.
- Use the **no** or **default** command of this command to restore the default value of buffer.

Command	**mmu queue-threshold output** { **unicast** } [*queue-id1* [*queue-id2* [*queue-idN*]] **set** *thr%*
Parameter Description	**output:** performs buffer management on the egress queue **unicast:** performs buffer management on the egress unicast queue *queue-id*: queue ID, in the range from 1 to 8 *thr%*: percentage, in the range from 1 to 100
Defaults	A shared buffer use threshold is allocated to each queue by default. This threshold is a percentage. The calculation method of the maximum available shared buffer for the queue is as follows: Maximum available shared buffer for the queue = Total number of shared buffers of the port group * Threshold percentage The default value depends on the product.
Command Mode	Interface configuration mode
Usage Guide	The effective way of this command varies with the equipment and depends on the product.

Configuring flow control threshold

- Optional.
- Use the **no** or **default** form of the command to restore the default value of buffer.

Command	**mmu fc-threshold set** *thr%*
Parameter Description	*value:* flow control threshold in the unit of percentage, range: 1-100

Defaults	Vary with products
Command Mode	Interface configuration mode
Usage Guide	1. The effective way of this command varies with the product. 2. The configuration takes effect only when flow control/PFC is enabled. 3. If flow control/PFC is not enabled, the shared buffer threshold of the PG is according to the value of **ingress-threshold**. 4. The user-configured value is displayed when the **show running-config** command is executed, even if the user-configured value is the default value.

Verification

- Use the **show running** command to check whether the MMU under the corresponding interface is configured successfully.

3.4.2 Configuring Buffer Monitoring

Configuration Effect

- Configure the buffer utilization alarm threshold of queue. The log alarm will be printed when the buffer utilization of queue exceeds this configured value.

Notes

- Configuration on the interface can be made on the physical port only.

Configuration Steps

Configuring the queue buffer utilization alarm threshold

- Optional.
- In the interface configuration mode, use the **mmu usage-warn-limit** { **unicast** | **multicast** } [*queue-id1* [*queue-id2* [*queue-idN*]] **set** *value* command to configure the buffer utilization alarm threshold for each queue.
- Use the **no** or **default** command of this command to restore the default value of buffer.

Command	**mmu usage-warn-limit** { **unicast** } [*queue-id1* [*queue-id2* [*queue-idN*]] **set** *value*
Parameter Description	**unicast:** performs buffer management on the egress unicast queue *queue-id*: queue ID, in the range from 1 to 8 *value*: percentage, in the range from 1 to 100
Defaults	The default value is 0, indicating that no alarm is reported.
Command Mode	Interface configuration mode
Usage Guide	

Verification

- Use the **show running** command to check whether the MMU under the corresponding interface is configured successfully.
- Use the **show queue-buffer** command to check whether the configuration succeeds.

Configuration Examples

Configuring the buffer utilization alarm limit based on egress queue

Configuration Steps	● Configure the buffer utilization alarm threshold as 70% at the unicast queues 6 and 8 of port 1/1 on the switch.
	FS# configure terminal FS(config)# int te1/1 FS(config-if)#mmu usage-warn-limit unicast 6 8 set 70
Verification	● Check whether the created guaranteed buffer has been configured successfully.
	FS#show queue-buffer interface gigabitEthernet 0/9 Dev/slot Port-group Total-shared(%) Guarantee-used(%) Share-used(%) Available(%) Warn-limit(%) 1/- 1 74.5271 0.0822 14.7615 85.1562 NA Interface GigabitEthernet 0/9: Type Queue Admin-shared(%) Total-used(%) Available(%) Warn-limit(%) Peak-usage(%) Peak-time Unicast 1 (default) 7.4836 0.0103 NA 7.5041 2015/7/14 20:7:14 Unicast 2 (default) 0.0000 7.4938 NA 0.0000 NA Unicast 3 (default) 0.0000 7.4938 NA 0.0000 NA Unicast 4 (default) 0.0000 7.4938 NA 0.0000 NA Unicast 5 (default) 0.0000 7.4938 NA 0.0000 NA Unicast 6 (default) 0.0000 7.4938 70% 0.0000 NA Unicast 7 (default) 0.0000 7.4938 NA 0.0000 NA Unicast 8 (default) 0.0000 7.4938 70% 0.0000 NA

3.5 Monitoring

Clearing

Running the **clear** command during operation of the equipment may lead to service interruption due to loss of important information.

Description	Command
Clears the queue counter value.	**clear queue-counter**
Clears the historical buffer peak.	**clear mmu queue-buffer peaked**

Displaying

Description	Command
Displays the buffer use information of panel interface.	**show queue-buffer interface**

Displays the queue counter information of panel interface.	**show queue-counter interface**

Chapter9 Reliability Configuration

1. Configuring REUP
2. Configuring RLDP
3. Configuring DLDP
4. Configuring VRRP
5. Configuring VRRP Plus
6. Configuration BFD
7. Configuring IP Event Dampening
8. Configuring Stacking
9. Configuring RNS

1 Configuring REUP

1.1 Overview

The Rapid Ethernet Uplink Protection Protocol (REUP) provides a rapid uplink protection function.

In the dual uplink networking, REUP is used to ensure normal communication between links, block redundant links, avoid link loops, and implement fast backup.

The upstream interfaces of REUP are configured in pairs. If both interfaces are normal, an interface works in the backup state. The interface in the backup state does not forward data packets. When the interface in the forward state is faulty, the backup interface switches to the forward state immediately, and provides data transmission. In addition, REUP also sends address update packets to upstream devices so that the upstream devices can update their MAC addresses immediately. This function of REUP ensures that layer-2 data streams can be restored within 50 ms after a link is faulty.

REUP is mutually exclusive with the Spanning Tree Protocol (STP) based on interfaces. In this case, a device runs STP downward and runs REUP upward to implement backup and fault protection for the upstream link. REUP ensures that basic link redundancy is provided when STP is disabled and that millisecond-level fault recovery faster than STP is also provided.

Protocols and Standards

- REUP is a proprietary protocol of FS Network, and there is no standard and protocol for reference.

1.2 Applications

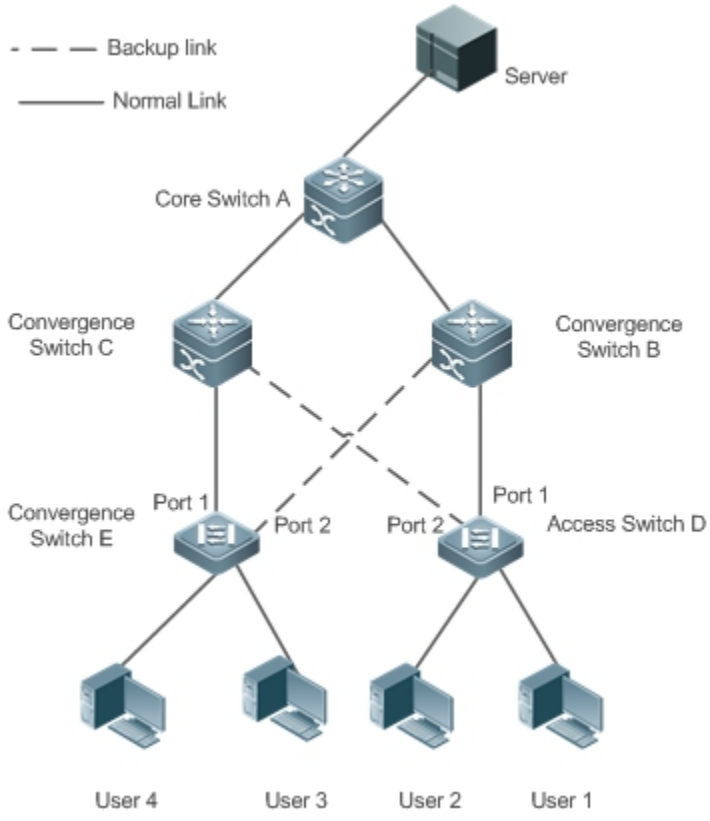
Application	Description
Communication in Dual Uplink Networking	Forward packets in the dual-uplink networking.

1.2.1 Communication in Dual Uplink Networking

Scenario

For communication in dual uplink networking, the access switch has two uplink paths, as shown in Figure 1-1.

Figure 1- 1 Dual uplink networking



Deployment

- Enable REUP on interface1 and interface2 of the access switch D/E to implement fast switching when a link is faulty.
- Enable MAC address update message receiving of REUP on the interfaces connected to switches A/B/C to rapidly clear the MAC addresses on the interfaces when a link is faulty.

1.3 Features

Basic Concepts

REUP Pair

Specify an interface as the backup interface of another interface to configure an REUP pair. One interface is the active interface and the other interface is the backup interface. When the two interfaces are normal, an interface is configured as the forward interface whereas the other interface is configured as the backup interface. You can determine the interface to be configured as the backup interface. See the related information in the section "Configuring the Preemption Mode and Delay Time of REUP".

MAC Address Update Message

MAC address update messages refer to FLUSH packets sent by FS Network to uplink devices through private multicast. When an uplink device of FS Network enables the function for receiving MAC address update messages and receives MAC address update messages, the device updates the MAC addresses of corresponding interfaces.

MAC Address Update Group

Multiple interfaces are added to a group. If one interface in the group receives a MAC address update message, the MAC addresses of

other interfaces in the group will be updated. In this case, the group is called MAC address update group.

MAC Address Update Packet

Packets sent to update MAC addresses in order to support uplink devices are called MAC address update packets.

Link Tracking Group

The uplink and downstream interfaces of a device are added to a group. If all upstream interfaces in the group are down, all downstream interfaces in this group are forced down. In this case, this group is called a link tracing group.

Overview

Feature	Description
Dual Link Backup of REUP	When a link is faulty, the other link can rapidly switch to the forward state.
Preemption Mode and Delay Time of REUP	When both links are normal, the preemption mode can be used to determine the link that is used for forwarding data and the delay time that is used to determine the waiting time before switching.
MAC Address Update	During link switching, the MAC address of an interface is updated to make packet convergence faster.
VLAN Load Balance	When the two links are normal, the utilization of link bandwidth can be maximized.
Link State Tracking	When the upstream link is faulty, the downstream link is switched.

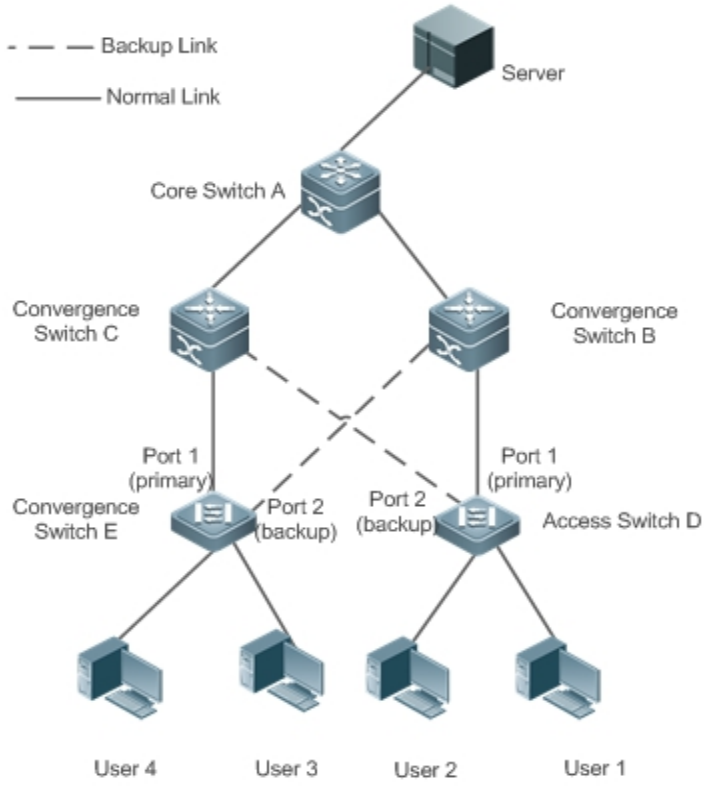
1.3.1 Dual Link Backup of REUP

When an active link is faulty, the link in the backup state will rapidly switch to the forward state and start forwarding data, minimizing the service interruption caused by link failure.

Working Principle

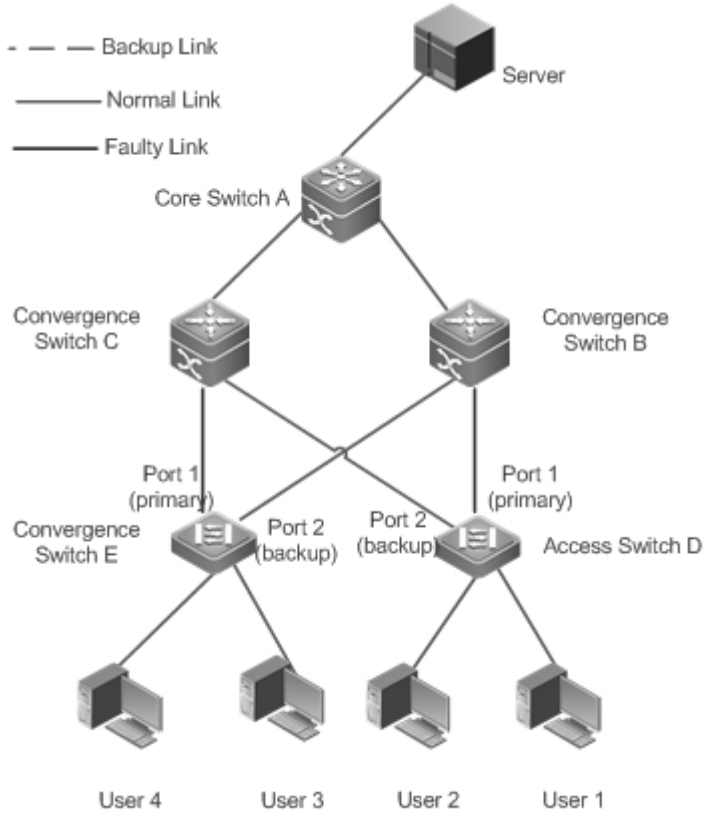
Specify an interface as the backup interface of another interface to configure an REUP pair. When the two interfaces are normal, a link is in the forward state (forwarding data packets) and the other link is in the backup state (not forwarding data). When the active link is faulty, the link in the backup state rapidly switches to the forward state and starts forwarding data. When the faulty link is recovered, the link enters the backup state and does not forward data packets. Of course, you can configure the preemption mode to specify whether a link recovered from failure preempts the link that is in the forward state currently.

Figure 1- 2 A topology with two normal links



As shown in Figure 1-2, connect interfaces 1 and 2 of switch D (E) to the uplink switches B and C (C and B) and configure REUP on interfaces 1 and 2. When the links are normal, interface 1 is in the forward state and forwards data packets and interface 2 is in the backup state and does not forward data packets.

Figure 1- 3 A topology with interface 1 of switch D (E) faulty



Once interface 1 is faulty, interface 2 immediately starts forwarding data packets and recovers the uplink transmission of the switch. In the non-preemption mode, when the link of interface 1 is recovered, interface 1 is in the backup state and does not forward data packets whereas interface 2 continues forwarding data packets.

Related Configuration

Enabling Dual Link Backup on an interface

By default, dual link backup on an interface is disabled.

You can run the **switchport backup interface** command to configure a layer-2 physical interface (or layer-2 AP interface) as a backup interface and enable the dual link backup function of REUP.

You must enable the dual link backup function of REUP on an interface. The function involves the link switching of REUP only when an interface is faulty.

REUP, ERPS, and RERP do not share interfaces.

Devices enabled with REUP must disable the storm control function of all layer-2 interfaces.

1.3.2 Preemption Mode and Delay Time of REUP

Working Principle

You can determine which link should be used first by configuring the preemption mode of REUP. If the preemption mode is set to

bandwidth first, REUP selects a link with a high bandwidth first. You can also set the preemption mode to forced to select a stable and reliable link first forcibly.

To avoid frequent active/backup link switching caused by abnormal faults, REUP provides a preemption delay function. When the two links are recovered, link switching is performed when the faulty link becomes stable after a delay (35s by default).

Related Configuration

Configuring the Preemption Mode and Delay Time of REUP

By default, the preemption mode is disabled and the delay time is 35s.

You can run the **switchport backup interface preemption mode** command to configure the preemption mode.
You can run the **switchport backup interface preemption delay** command to configure the delay time.

A smaller delay means more frequent preemption switching after the faulty link is recovered.

REUP uses the value of the **Bandwidth** attribute for an AP interface as the actual bandwidth of the AP interface, which is equal to the value of the **Speed** attribute (the number of link up member interfaces x the number of member interfaces).

When an uplink enables STP, the preemption delay time of REUP is greater than 35s.

1.3.3 MAC Address Update

During link switching, the MAC address of an interface is updated to make packet convergence faster.

Working Principle

As shown in Figure 1-2, interface 1 and interface 2 of switch D (E) are enabled with dual link backup of REUP. Interface 1 works as the active interface. During normal communication, switch A learns the MAC addresses of users 1 and 2 (users 3 and 4) from the interfaces connecting to switch B (C).

When interface 1 of switch D (E) is faulty, interface 2 rapidly switches to the forward state and starts forwarding data packets. In this case, switch A does not learn the MAC addresses of users 1 and 2 (users 3 and 4) on the interfaces connecting to switch B (C). The data packets sent by the server to users 1 and 2 (users 3 and 4) are forwarded to switch C (B) by switch A, causing that the packets from the server to users 1 and 2 (users 3 and 4) are lost.

To avoid the preceding problems, you can enable the MAC address update function on switch D (E). When interface 2 starts forwarding packets, switch D (E) sends a MAC address update message to interface 2. After receiving the MAC address update message, switch A updates the MAC address on the interface of switch A. In this way, switch A forwards the packets sent by the server to the users to the interfaces of switch B (C) to make packet convergence faster.

In addition, import the setting of a MAC address update group, that is, classify multiple interfaces into the same group. When an interface in this group receives a MAC address update message, the MAC addresses on other interfaces in the group are updated to reduce the side effect of flooding caused by MAC address update.

To be compatible with upstream devices not supporting MAC address update messages, switch D (E) will send MAC address update packets for users 1 and 2 (users 3 and 4) upward when interface 2 switches to the forward state. In this way, switch A can update the MAC addresses of users 1 and 2 (users 3 and 4) to the corresponding interfaces and recover the downlink data transmission of switch A.

Related Configuration

Enabling Sending of MAC Address Update Messages on an interface

By default, sending of MAC address update messages is disabled on an interface.

You can run the **mac-address-table move update transit** command to enable sending of MAC address updates on all interfaces of a device.

If sending of MAC address update messages is not enabled, MAC address update messages will not be sent when dual link backup switching of REUP is performed.

Enabling Receiving of MAC Address Update Messages on an interface

By default, receiving of MAC address update messages is disabled on an interface.

You can run the **mac-address-table move update receive** command to enable receiving of MAC address updates on all interfaces of a device.

If receiving of MAC address update messages is not enabled, a device cannot receive MAC address update messages from downlink devices during dual link backup switching of REUP and will not update the MAC addresses.

Configuring a VLAN for Sending MAC Address Update Messages

By default, a VLAN for sending MAC address update messages is the default VLAN to which an interface belongs.

You can run the **mac-address-table move update transit vlan** command to configure the VLAN in which interfaces send MAC address update messages.

If the VLAN in which interfaces send MAC address update messages is configured, the messages are sent in the configured VLAN; otherwise, the messages are sent in the default VLAN to which the interface belongs.

Configuring a VLAN for Receiving MAC Address Update Messages

By default, MAC address update messages are received in all VLANs.

You can run the **no mac-address-table move update receive vlan** command to configure a VLAN in which interfaces do not receive MAC address update messages. MAC address update messages are received in remaining VLANs.

If no VLAN in which interfaces receive MAC address update messages is configured, MAC address update messages are received in all the configured VLANs; otherwise, MAC address update messages are received in the remaining VLANs.

Configuring a MAC Address Update Group

By default, there is no MAC address update group.

You can run the **mac-address-table update group** command to add an interface to the MAC address update group. The interface is added to the first update group by default.

If no MAC address update group is configured, MAC address update will not be performed when MAC address update packets are received.

Configuring the Maximum Number of MAC Address Update Packets Sent Per Second

By default, the maximum number of MAC address update packets sent per second is 150.

You can run the **mac-address-table move updatemax-update-rate** command to configure the maximum number of MAC address update packets sent per second.

The larger the number of packets, the more CPU time used for sending the packets, and the fewer downlink packets are lost.

1.3.4 VLAN Load Balance

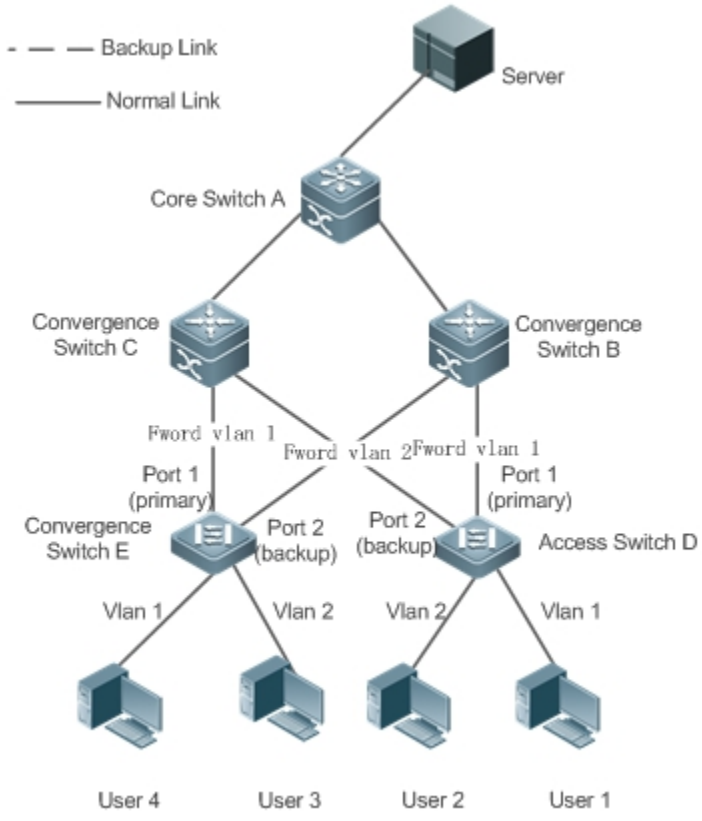
Working Principle

The VLAN load balance function allows REUP to forward data packets of mutually exclusive VLANs for two interfaces to make full use of the link bandwidth.

As shown in Figure 1-4, configure dual link backup of REUP and enable VLAN load balance of REUP on interface 1 and interface 2 of switch D, and map VLAN 1 to instance 1 and VLAN 2 to instance 2. Data of VLAN 1 (instance 1) is transmitted through interface 1 and all the other data of VLAN 2 (instance 2) is transmitted through interface 2. Perform the same processing on switch E.

When an interface is faulty, the other interface takes over the transmission of all VALNs. When the faulty interface is recovered and does not become faulty within the preemption delay, the transmission of VLANs is switched back to the recovered interface.

Figure 1- 4 A topology with two normal links of load balance



Related Configuration

Enabling VLAN Load Balance on an interface

By default, the VLAN load balance function on an interface is disabled.

You can run the **switchport backup interface prefer instance** command to enable the VLAN load balance function.

If this function is not enabled, the link bandwidth cannot be fully used when packets are forwarded when the two links are normal. You must enable the VLAN load balance function on a port so that the interface can be involved in VLAN load balance.

The instance mapping of REUP VLAN load balance is controlled by the MSTP module in a unified manner. For details about how to

configure the instances, see the description in the *Configuring MSTP*.

The VLAN load balance function can be configured only on trunk, uplink or hybrid interfaces.

1.3.5 Link State Tracking

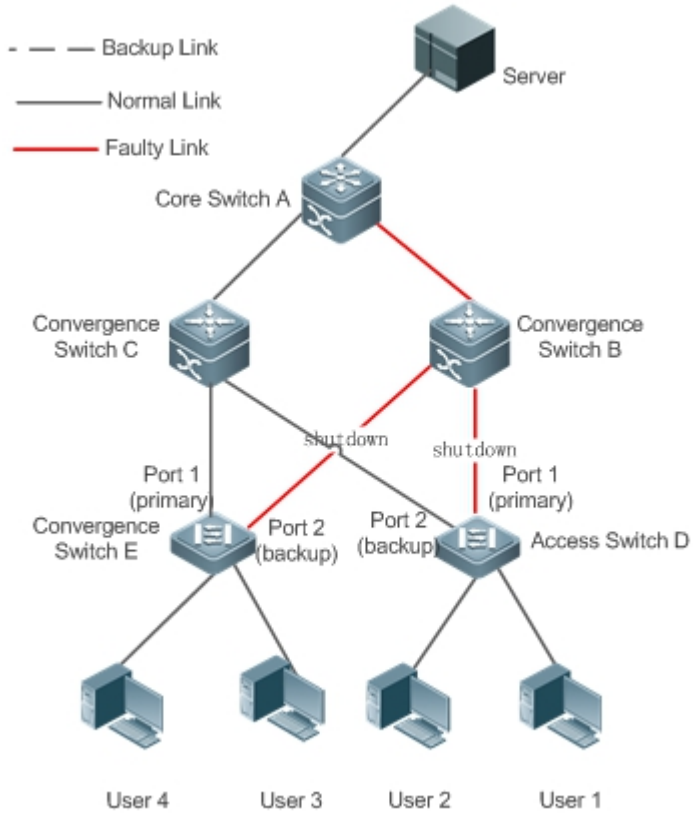
Link tracking means that when the upstream link is faulty, services are switched to the downstream link so that the backup interface can continue forwarding packets.

Working Principle

Link state tracking provides the function of notifying downlink devices for link switching when the upstream link is faulty. You can configure the uplink and downstream interfaces of a link state tracking group and bind the link status of multiple downstream interfaces to the interfaces of multiple upstream links to implement link status synchronization. When all upstream links in a tracking group are faulty, the interfaces of the downstream links are shut down forcibly to ensure that the transmission of the downstream links is switched from the active link to the backup link.

As shown in Figure 1-5, when the upstream link of switch B is faulty, link state tracking rapidly shuts down the downstream interface of switch B so that the uplink transmission of switch D is switched to switch C.

Figure 1- 5 A topology where the upstream link of the active link is faulty



Related Configuration

Enabling Link Tracking

Link tracking is disabled by default.

You can run the **link state track** [*number*] command to enable a link tracking group. The value of **number** ranges from 1 to 2. The first

link tracking group is enabled by default (the default value of **number** is 1).

If link tracking is not enabled, the status of a corresponding upstream interface cannot be detected and packet forwarding switching cannot be implemented in time.

Adding an interface to a Link Tracking Group

By default, an interface is not added to a link tracking group.

You can run the **link state group** [*number*] {**upstream** | **downstream**} command to set upstream interfaces and downstream interfaces of the link tracking group. The value of **number** ranges from 1 to 2. An interface is added to the first link tracking group by default (the default value of number is **1**).

If an interface is not added to a tracking group, the status of a corresponding upstream interface cannot be detected and packet forwarding switching cannot be implemented in time.

1.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of REUP	(Mandatory) It is used to enable dual link backup of REUP.	
	switchport backup interface	Enables dual link backup of REUP.
Configuring the Preemption Mode and Delay Function of REUP	(Optional) It is used to determine the preemption mode and delay time. The default values are used if they are not configured.	
	switchport backup interface preemption mode	Sets the preemption mode.
	switchport backup interface preemption delay	Sets the delay time for preemption.
Configuring MAC Address Update	(Optional) It is used to enable rapid update of MAC addresses.	
	mac-address-table update group	Sets the MAC address update group ID of a switch.
	mac-address-table move update transit	Enables sending of MAC address update messages.
	mac-address-table move update transit vlan	Enables sending of the VLAN ID of MAC address update messages.
	mac-address-table move update	Configures the maximum number of MAC address update packets sent per second. The value ranges from 0 to 32000. The default value is 150.
	mac-address-table move update receive	Enables receiving of MAC address update messages.
	mac-address-table move update receive vlan	Configures the VLAN range for processing MAC address update messages.
Configuring VLAN Load Balance	(Optional) It is used to enable VLAN load balance.	
	switchport backup interface prefer instance	Configures the link VLAN load balance of REUP.
Configuring Link Tracking	(Optional) It is used to enable link tracking.	

Configuration	Description and Command	
	link state track up-delay	Enables the downlink delay up for a link state tracking group.
	link state track	Enables a link state tracking group.
	link state group	Add an interface as an upstream interface or a downstream interface of a specified link state tracking group.

1.4.1 Configuring Basic Functions of REUP

Configuration Effect

- When a link is faulty, the other normal link is switched to the forward state immediately for forwarding packets.

Notes

- An interface belongs to only one REUP pair. Each active link has only one backup link. A backup link can be used as the backup link of only one active link. The active and backup links must use different interfaces.
- REUP supports layer-2 physical interfaces and AP interfaces, but does not support AP member interfaces.
- The active and backup interfaces may be of different types and have different rates. For example, an AP interface can be used as the active interface whereas a physical interface is configured as the backup interface.
- Interfaces configured with REUP are not involved in STP calculation.
- Each device can be configured with a maximum of 16 REUP pairs.
- Interfaces successfully configured with REUP cannot change interfaces to layer-3 interfaces or be added to an AP.

Configuration Steps

Enabling Dual Link Backup of REUP

- Mandatory.
- If there is no special requirement, dual link backup of REUP should be enabled on an interface of the receiving switch.

Verification

Run the **show interfaces switchport backup [detail]** command to check whether dual link backup of REUP is configured.

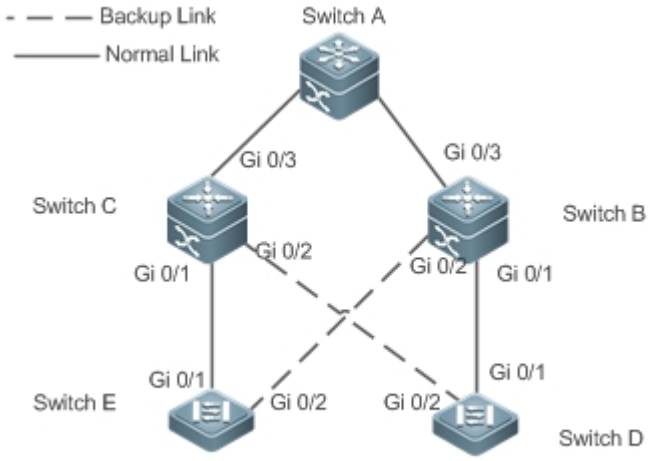
Related Commands

Enabling Dual Link Backup of REUP

Command	**switchport backup interface** *interface-id*
Parameter Description	*interface-id*: Indicates the backup interface ID.
Command Mode	Interface configuration mode
Usage Guide	If the interface where the mode resides is the active interface, the interface corresponding to the **interface-id** parameter is the backup interface. When the active link is faulty, rapidly recover the transmission of the backup link.

Configuration Example

Enabling Dual Link Backup of REUP

Scenario **Figure 1- 6 Dual uplink networking**	As shown in Figure 1-6, there are two upstream links from switch D to switch A, which are switch D > switch B > switch A and switch D > switch C > switch A. There are two upstream links from switch E to switch A, which are switch E > switch B > switch A and switch E > switch C > switch A. 
Configuration Steps	● Configure dual link backup (the interface Gi0/1 is the active interface and Gi0/2 is the backup interface) of REUP on the access switch D (E).
D	SwitchD> enable SwitchD# configure terminal SwitchD(config)# interface GigabitEthernet 0/1 SwitchD(config-if-GigabitEthernet 0/1)# switchport mode trunk SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface GigabitEthernet 0/2 SwitchD(config-if-GigabitEthernet 0/1)# exit
E	SwitchE> enable SwitchE# configure terminal SwitchE(config)# interface GigabitEthernet 0/1 SwitchE(config-if-GigabitEthernet 0/1)# switchport mode trunk SwitchE(config-if-GigabitEthernet 0/1)#switchport backup interface GigabitEthernet 0/2 SwitchD(config-if-GigabitEthernet 0/1)# exit
Verification	● Check the dual link backup information configured for switch D (E).

D	SwitchD#show interfaces switchport backup detail Switch Backup Interface Pairs: Active Interface Backup Interface State ---------- Gi0/1 Gi0/2 Active Up/Backup Standby Interface Pair : Gi0/1, Gi0/2 Preemption Mode : off Preemption Delay : 35 seconds Bandwidth : Gi0/1(100000 Mbits), Gi0/2(100000 Mbits)
E	SwitchE#show interfaces switchport backup detail Switch Backup Interface Pairs: Active Interface Backup Interface State ---------- Gi0/1 Gi0/2 Active Up/Backup Standby Interface Pair : Gi0/1, Gi0/2 Preemption Mode : off Preemption Delay : 35 seconds Bandwidth : Gi0/1(100000 Mbits), Gi0/2(100000 Mbits)

Common Errors

- Other REUP pairs are configured on a configured interface.
- A configured interface is not a layer-2 physical interface or AP interface.

1.4.2 Configuring the Preemption Mode and Delay Function of REUP

Configuration Effect

- Restrict the preemption mode and preemption delay time for REUP link switching.

Notes

- Dual link backup of REUP must be configured.

Configuration Steps

- Optional.
- If the active link needs to always forward packets or the link bandwidth needs to be used to determine the link for forwarding packets, the corresponding preemption mode and delay time must be configured.

Verification

Run the **show interfaces switchport backup [detail]** command to check whether the preemption mode and delay time are consistent with the configurations.

Related Commands

Configuring the Preemption Mode of REUP

Command	**switchport backup interface** *interface-id* **preemption mode** {**forced**|**bandwidth**|**off**}
Parameter Description	*interface-id*: Indicates the backup interface ID. **mode**: Sets the preemption mode: **forced**: Indicates the forced mode. **bandwidth**: Indicates the bandwidth mode. **off**: Indicates that the preemption mode is off.
Command Mode	Interface configuration mode
Usage Guide	The preemption modes include forced, bandwidth and off. In the bandwidth mode, an interface with a high bandwidth is selected first to transmit data; in the forced mode, the active interface is selected first to transmit data; in the off mode, no preemption is performed. The default mode is off.

Configuring the Delay Time of REUP

Command	**switchport backup interface** *interface-id***preemption delay** *delay-time*
Parameter Description	*interface-id*: Indicates the backup interface ID. *delay-time*: Indicates the delay time.
Command Mode	Interface configuration mode
Usage Guide	Preemption delay indicates the delay time after a faulty link is recovered to the time when link switching is performed again.

Configuration Example

Configuring the Preemption Mode and Delay Time of REUP

Scenario	As shown in Figure 1-6, there are two upstream links from switch D to switch A, which are switch D > switch B > switch A and switch D > switch C > switch A. There are two upstream links from switch E to switch A, which are switch E > switch B > switch A and switch E > switch C > switch A.
Configuration Steps	● Configure the preemption mode to bandwidth on the access switch D (E) and the delay time to 40s.
D	SwitchD> enable SwitchD# configure terminal SwitchD(config)# interface GigabitEthernet 0/1 SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface gi 0/2 preemption mode bandwidth SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface gi 0/2 preemption delay 40 SwitchD(config-if-GigabitEthernet 0/1)# exit

E	SwitchE> enable SwitchE# configure terminal SwitchD(config)# interface GigabitEthernet 0/1 SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface gi 0/2 preemption mode bandwidth SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface gi 0/2 preemption delay 40 SwitchD(config-if-GigabitEthernet 0/1)# exit
Verification	● Check the dual link backup information configured for switch D (E).
D	SwitchD#show interfaces switchport backup detail Switch Backup Interface Pairs: Active Interface Backup Interface State ---------- Gi0/1 Gi0/2 Active Up/Backup Standby Interface Pair : Gi0/1, Gi0/2 Preemption Mode : bandwidth Preemption Delay : 40 seconds Bandwidth : Gi0/1(100000 Mbits), Gi0/2(100000 Mbits)
E	SwitchE#show interfaces switchport backup detail Switch Backup Interface Pairs: Active Interface Backup Interface State ---------- Gi0/1 Gi0/2 Active Up/Backup Standby Interface Pair : Gi0/1, Gi0/2 Preemption Mode : bandwidth Preemption Delay : 40 seconds Bandwidth : Gi0/1(100000 Mbits), Gi0/2(100000 Mbits)

Common Errors

- A configured interface is not a layer-2 physical interface or AP interface.

1.4.3 Configuring MAC Address Update

Configuration Effect

- Rapidly delete and update MAC addresses of an interface during link switching to make packet convergence faster.

Notes

- Dual link backup of REUP must be configured.
- Each device can be configured with a maximum of 8 address update groups. Each address update group can have a maximum of 8 member interfaces and an interface can belong to multiple address update groups.

Configuration Steps

- Mandatory.
- If there is no special requirement, the MAC address update function should be configured.

Verification

Run the **show mac-address-table update group** [**detail**] command to view the update group configuration.

Related Commands

Configuring the MAC Address Update Group ID of a Switch

Command	**mac-address-table update group** [*group-num*]
Parameter Description	*group-num*: Indicates the MAC address update group ID.
Command Mode	Interface configuration mode
Usage Guide	In order to reduce large flooding caused by MAC address update which may affect normal data transmission of the switch, we add a setting of a MAC address update group. Only after all interfaces on a switching path are added to the same MAC address update group, transmission of downlink data can be rapidly recovered.

Enabling Sending of MAC Address Update Messages

Command	**mac-address-table move update transit**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	To reduce link switching and loss of downlink data streams, you need to enable sending of MAC address update messages on a switch that performs switching.

Enabling Sending of the VLAN ID of MAC Address Update Messages

Command	**mac-address-table move update transit vlan***vid*
Parameter Description	*vid*: Indicates the VLAN ID for sending MAC address update messages.
Command	Interface configuration mode

Mode	
Usage Guide	After sending of MAC address update messages is enabled, MAC address update messages can be sent to uplink devices during link switching.

Configure the maximum number of MAC address update packets sent per second.

Configuring the Maximum Number of MAC Address Update Packets Sent Per Second

Command	**mac-address-table move update max-update-rate***pkts-per-second*
Parameter Description	*pkts-per-second*: Indicates the maximum number of MAC address update packets sent per second. The value ranges from 0 to 32000. The default value is 150.
Command Mode	Global configuration mode
Usage Guide	During link switching, REUP sends MAC address update packets of a specified quantity to uplink devices per second to recover the downlink data transmission of the uplink device.

Enabling Receiving of MAC Address Update Messages

Command	**mac-address-table move update receive**
Parameter Description	-
Command Mode	Global configuration mode
Usage Guide	During switching of dual link backup, downlink data streams may be lost since the MAC address table of the uplink switch is not updated in real time. In order to reduce loss of layer-2 data streams, you need to update the MAC address table of the uplink switch. In this case, you need to enable receiving of MAC address update messages on the uplink switch.

Configuring the VLAN Range for Processing MAC Address Update Messages

Command	**mac-address-table move update receive vlan***vlan-range*
Parameter Description	*vlan-range*: Indicates the VLAN range for processing MAC address update messages.
Command Mode	Global configuration mode
Usage Guide	This command is used to disable the function for processing MAC address update messages on certain VLANs. For a VLAN disabled with the function for processing MAC address update messages, MAC address update packets can be used to recover the downlink transmission of uplink devices; however, the convergence performance for link faults will be decreased.

Configuration Example

Configuring MAC Address Update

Scenario	As shown in Figure 1-6, there are two upstream links from switch D to switch A, which are switch D > switch B > switch A and switch D > switch C > switch A. There are two upstream links from switch E to switch A, which are switch E > switch B > switch A and switch E > switch C > switch A.

Configuration Steps	● Enable sending of MAC address update messages on the access switch D (E). ● Enable receiving of MAC address update packets on switch B (C). ● Add all interfaces on the REUP switching path to the same MAC address update group. ● In the environment, Gi0/1 and Gi0/3 of switch B are the interfaces on the switching path of switch D's uplink, and Gi0/3 and Gi0/2 are the interfaces on the switching path of switch E's uplink. You can add interfaces Gi0/1, Gi0/2 and Gi0/3 to the same address update group. Similarly, you can obtain the configuration of switch C. ● Enable receiving of MAC address update packets on switch A. ● Add all interfaces on the REUP switching path of switch A to the same MAC address update group.
D	SwitchD> enable SwitchD# configure terminal SwitchD(config)# mac-address-table move update transit SwitchD(config)# exit
E	SwitchE> enable SwitchE# configure terminal SwitchE((config)# mac-address-table move update transit SwitchE(config)# exit
B	SwitchB# configure terminal SwitchB(config)# mac-address-table move update receive SwitchB(config)# interface range gigabitEthernet 0/1 -3 SwitchB(config-if-range)#switchport mode trunk SwitchB(config-if-range)# mac-address-table update group 1 SwitchB(config-if-range)# end
C	SwitchB# configure terminal SwitchB(config)# mac-address-table move update receive SwitchB(config)# interface range gigabitEthernet 0/1 -3 SwitchB(config-if-range)#switchport mode trunk SwitchB(config-if-range)# mac-address-table update group 1 SwitchB(config-if-range)# end

A	SwitchA# configure terminal SwitchA(config)# mac-address-table move update receive SwitchA(config)# interface range gigabitEthernet 0/1 -2 SwitchA(config-if-range)# switchport mode trunk SwitchA(config-if-range)# mac-address-table update group 1 SwitchA(config-if-range)# end
Verification	Check the information about the address update groups on switches D, E, C, B and A.
D	SwitchD# show run | incl mac-ad mac-address-table move update transit
E	SwitchE# show run | incl mac-ad mac-address-table move update transit
B	SwitchB# show mac-address-table update group detail show mac-address-table update group detailMac-address-table Update Group:1 Received mac-address-table update message count:0 Group member Receive Count Last Receive Switch-ID Receive Time ---------- Gi0/1 0 0000.0000.0000 Gi0/2 0 0000.0000.0000 Gi0/3 0 0000.0000.0000
C	SwitchC# show mac-address-table update group detail Mac-address-table Update Group:1 Received mac-address-table update message count:0 Group member Receive Count Last Receive Switch-ID Receive Time ---------- Gi0/1 0 0000.0000.0000 Gi0/2 0 0000.0000.0000 Gi0/3 0 0000.0000.0000
A	SwitchA# show mac-address-table update group detail Mac-address-table Update Group:1 Received mac-address-table update message count:0 Group member Receive Count Last Receive Switch-ID Receive Time

	---------- Gi0/1 0 0000.0000.0000 Gi0/2 0 0000.0000.0000

Common Errors

- A configured interface is not a layer-2 physical interface or AP interface.

1.4.4 Configuring VLAN Load Balance

Configuration Effect

- Maximize the utilization of link bandwidth.

Notes

- Dual link backup of REUP must be configured.
- The Access interface cannot be shared by VLAN load balance and STP.
- For interfaces successfully configured with VLAN load balance, you cannot modify the attributes of the interfaces but can modify the VLAN attributes of the interfaces.

Configuration Steps

- If maximizing bandwidth utilization is not required, this configuration is optional.
- If there is a requirement for VLAN load balance, corresponding configuration must be performed.

Verification

Run the **show interfaces switchport backup [detail]** command to check whether VLAN load balance is configured.

Related Commands

Configuring VLAN Load Balance

Command	**switchport backup interface** *interface-id* **prefer instance** *instance-range*
Parameter Description	*interface-id*: Indicates the backup interface ID. *instance-range*: Indicates the load instance range of the backup interface.
Command Mode	Interface configuration mode
Usage Guide	You can modify the mapping between instances and VLANs by using the instance mapping function of MSTP.

Configuration Example

Configuring VLAN Load Balance

Scenario	As shown in Figure 1-6, there are two upstream links from switch D to switch A, which are switch D > switch B > switch A and switch D > switch C > switch A. There are two upstream links from switch E to switch A, which are switch E > switch B > switch A and switch E > switch C > switch A.
Configuration	● Configure instance mappings on switch D (E) to map VLAN 1 to instance 1, VLAN 2 to instance 2, VLAN 3 to

Steps	instance 3, and VLAN 4 to instance 4. For details, see the *MSTP Configuration Guide*. ● Configure the VLAN load balance function on switch D (E).
D	SwitchD> enable SwitchD# configure terminal SwitchD(config)# interface GigabitEthernet 0/1 SwitchD(config-if-GigabitEthernet 0/1)# switchport mode trunk SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface gi0/2 prefer instance 2 SwitchD(config-if-GigabitEthernet 0/1)# exit
E	SwitchE> enable SwitchE# configure terminal SwitchE(config)# interface GigabitEthernet 0/1 SwitchE(config-if-GigabitEthernet 0/1)# switchport mode trunk SwitchD(config-if-GigabitEthernet 0/1)#switchport backup interface gi0/2 prefer instance 4 SwitchD(config-if-GigabitEthernet 0/1)# exit
Verification	● Check the dual link backup information configured for switch D (E).
D	SwitchD#show interfaces switchport backup detail Switch Backup Interface Pairs: Active Interface Backup Interface State ---------- Gi0/1 Gi0/2 Active Up/Backup Up Instances Preferred on Active Interface: Instance 0-1,3-64 Mapping VLAN 1,3-4094 Instances Preferred on Backup Interface: Instance 2 Mapping VLAN 2 Interface Pair : Gi0/1, Gi0/2 Preemption Mode : balance Preemption Delay : 35 seconds Bandwidth : Gi0/1(800 kbits), Gi0/2(100000 kbits)

E	SwitchE#show interfaces switchport backup detail Switch Backup Interface Pairs: Active Interface Backup Interface State ---------- Gi0/1 Gi0/2 Active Up/Backup Up Instances Preferred on Active Interface: Instance 0-3,5-64 Mapping VLAN 1-3,5-4094 Instances Preferred on Backup Interface: Instance 4 Mapping VLAN 4 Interface Pair : Gi0/1, Gi0/2 Preemption Mode : balance Preemption Delay : 35 seconds Bandwidth : Gi0/1(800 kbits), Gi0/2(100000 kbits)

Common Errors

- The mappings between VLAN IDs and instances are not configured.

1.4.5 Configuring Link Tracking

Configuration Effect

- After detecting that the upstream link is disconnected, forcibly disconnect the downstream link so that link switching can be performed.

Notes

- Dual link backup of REUP must be configured.
- For the link state tracking function, each interface belongs to only one link state tracking group and each device can be configured with up to 2 link state tracking groups. Each link state tracking group can have 8 upstream interfaces and 256 downstream interfaces.

Configuration Steps

- Mandatory.
- If there is no special requirement, the uplink tracking function should be configured.

Verification

Run the **show link state group** command to view the configured link tracking information.

Related Commands

Enabling a Link State Tracking Group

Command	**link state track** [*num*]
Parameter Description	*num*: Indicates the ID of a link state tracking group.
Command Mode	Global configuration mode
Usage Guide	You can create a link tracking group and then add an interface to the specified tracking group.

Adding an interface to a Link Tracking Group

Command	**ink stategroup** *num* {**upstream | downstream**}
Parameter Description	*num*: Indicates the ID of a link state tracking group. **upstream**: Adds the interface as an upstream interface of the tracking group. **downstream**: Adds the interface as a downstream interface of the tracking group.
Command Mode	Interface configuration mode
Usage Guide	You can create a link tracking group and then add an interface to the specified tracking group.

Configuration Example

Configuring a Link Tracking Group

Scenario	As shown in Figure 1-6, there are two upstream links from switch D to switch A, which are switch D > switch B > switch A and switch D > switch C > switch A. There are two upstream links from switch E to switch A, which are switch E > switch B > switch A and switch E > switch C > switch A.
Configuration Steps	● Create link tracking group 1 on switch B (C). ● On switch B (C), add the interfaces Gi0/1 and Gi0/2 as downstream interfaces of the link tracking group and add the interface Gi0/3 as an upstream interface of the link tracking group.

B	SwitchB> enable SwitchB# configure terminal SwitchB(config)# link state track 1 SwitchB(config)# interface GigabitEthernet 0/1 SwitchB(config-if-GigabitEthernet 0/1)#link state group 1 downstreamSwitchB(config-if-GigabitEthernet 0/1)#exit SwitchB(config)# interface GigabitEthernet 0/2 SwitchB(config-if-GigabitEthernet 0/2)# link state group 1 downstream SwitchB(config-if-GigabitEthernet 0/2)#exit SwitchB(config)# interface GigabitEthernet 0/3 SwitchB(config-if-GigabitEthernet 0/3)#link state group 1 upstream SwitchB(config-if-GigabitEthernet 0/3)#exit
C	SwitchC> enable SwitchC# configure terminal SwitchC(config)# link state track 1 SwitchC(config)# interface GigabitEthernet 0/1 SwitchC(config-if-GigabitEthernet 0/1)#link state group 1 downstreamSwitchC(config-if-GigabitEthernet 0/1)#exit SwitchC(config)# interface GigabitEthernet 0/2 SwitchC(config-if-GigabitEthernet 0/2)# link state group 1 downstream SwitchC(config-if-GigabitEthernet 0/2)#exit SwitchC(config)# interface GigabitEthernet 0/3 SwitchC(config-if-GigabitEthernet 0/3)#link state group 1 upstream SwitchC(config-if-GigabitEthernet 0/3)#exit
Verification	Check the link tracking group information configured for switch B (C).
B	SwitchB#show link state group Link State Group:1 Status: enabled, Down Upstream Interfaces :Gi0/3(Down) Downstream Interfaces : Gi0/2(Down)

Common Errors

- Interfaces are added to a link tracking group when the link tracking group is not enabled.

1.5 Monitoring

Displaying

Description	Command
Displays the dual link backup information of REUP.	**show interfaces**[*interface-id*]**switchport backup** [**detail**]
Displays the configurations of an MAC address update group.	**show mac-address-table update group** [**detail**]
Displays the REUP statistics about sent MAC address update messages.	**show mac-address-table move update**
Displays the information about a link state tracking group.	**show link state group**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Enables all REUP debugging.	**debug reup all**
Debugs the normal running process of REUP.	**debug reup process**
Debugs MAC address update messages of REUP.	**debug reup packet**
Debugs MAC address update packets of REUP.	**debug reup macupdt**
Debugs hot backup.	**debug reup ha**
Debugs errors occurring in REUP running.	**debug reup error**
Debugs received events.	**debug reup evnet**
Debugs statistics when **show** operations are performed.	**debug reup status**

2 Configuring RLDP

2.1 Overview

The Rapid Link Detection Protocol (RLDP) achieves rapid detection of unidirectional link failures, directional forwarding failures and downlink loop failures of an Ethernet. When a failure is found, relevant ports will be closed automatically according to failure treatment configuration or the user will be notified to manually close the ports to avoid wrong flow forwarding or an Ethernet layer-2 loop.

2.2 Applications

Application	Description
Unidirectional Link Detection	Detect a unidirectional link failure.
Bidirectional Forwarding Detection	Detect a bidirectional link failure.
Downlink Loop Detection	Detect a link loop.

2.2.1 Unidirectional Link Detection

Scenario

As shown in the following figure, A is connected to B via optical fiber. The two lines are the Tx and Rx lines of optical fiber. Unidirectional link detection is enabled on A and B. If any of the Tx of Port A, Rx of Port B, Tx of Port B and Rx of Port A fails, a unidirectional failure will be detected and treated under the RLDP. If the failure is eliminated, the administrator may manually restore the RLDP on A and B and resume detection.

Figure 2- 1

Remarks	A and B are layer-2 or layer-3 switches. The Tx of Port A of A is connected to the Rx of Port B of B. The Rx of Port A of A is connected to the Tx of Port B of B.

Deployment

- Global RLDP is enabled.
- Configure unidirectional link detection under Port A and Port B and define a method for failure treatment.

2.2.2 Bidirectional Forwarding Detection

Scenario

As shown in the following figure, A is connected to B via optical fiber, and the two lines are Tx and Rx lines of optical fiber. Unidirectional link detection is enabled on A and B. If the Tx of Port A, Rx of Port B, Rx of Port A and Tx of Port B all fail, a bidirectional failure will be detected and treated under the RLDP. If the failure is eliminated, the administrator may manually restore the RLDP on A and B and

resume detection.

Figure 2- 2

Remarks	A and B are layer-2 or layer-3 switches. The Tx of Port A of A is connected to the Rx of Port B of B. The Rx of Port A of A is connected to the Tx of Port B of B.

Deployment

- Global RLDP is enabled.
- Configure BFD under Port A and Port B and define a method for failure treatment.

2.2.3 Downlink Loop Detection

Scenario

As shown in the following figure, A, B and C are connect into a loop. Downlink loop detection is enabled on A, and a loop is detected and treated.

Figure 2- 3

Remarks	A, B and C are layer-2 or layer-3 switches. A, B and C are interconnected via exchange ports.

Deployment

- Global RLDP is enabled on A.
- Configure downlink loop detection on the Gi 2/0/1 and Gi 2/0/9 ports of A, and define a method for failure treatment.

2.3 Features

Most Ethernet link detection mechanisms detect link connectivity through automatic physical-layer negotiation. However, in some cases devices are connected on the physical layer and operate normally but layer-2 link communication is disabled or abnormal. The RLDP recognizes a neighbor device and detects a link failure through exchanging Prob packets, Echo packets or Loop packets with the device.

Basic Concepts

Unidirectional Link Failure

A unidirectional link failure occurs in case of a cross-connected optical fiber, a disconnected optical fiber, an open-circuit optical fiber, one open-circuit line in a twisted-pair cable, or unidirectional open circuit of an intermediate device between two devices. In such cases, one end of a link is connected and the other disconnected so that flow is forwarded wrongly or a loop guard protocol (for example, the STP) fails.

Bidirectional Link Failure

A bidirectional link failure occurs in case of two optical fibers, two open-circuit lines in a twisted-pair cable, or bidirectional open circuit of an intermediate device between two devices. In such cases, the both ends of a link are disconnected so that flow is forwarded wrongly.

Loop Failure

A downlink device is wrongly connected to form a loop, resulting in a broadcast storm.

RLDP Packet

The RLDP defines three types of packets: Prob packets, Echo packets and Loop packets.

- Prob packets are layer-2 multicast packets for neighbor negotiation, and unidirectional or bidirectional link detection. The default encapsulation format is SNAP, which changes automatically to EthernetII if a neighbor sends EthernetII packets.
- Echo packets are layer-2 unicast packets as response to Prob packets and used for unidirectional or bidirectional link detection. The default encapsulation format is SNAP, which changes automatically to EthernetII if a neighbor sends EthernetII packets.
- Loop packets are layer-2 multicast packets for downlink loop detection. They can only be received. The default encapsulation format is SNAP.

RLDP Detection Interval and Maximum Detection Times

A detection interval and the maximum detection times can be configured for the RLDP. A detection interval determines the period of sending Prob packets and Loop packets. When a device receives a Prob packet, it replies with an Echo packet immediately. A detection interval and the maximum detection times determine the maximum detection time (equal to a detection interval × the maximum detection times + 1) for unidirectional or bidirectional link detection. If neither Prob nor Echo packet from a neighbor can be received within the maximum detection time, the treatment of unidirectional or bidirectional failure will be triggered.

RLDP Neighbor Negotiation

When configured with unidirectional or bidirectional link detection, a port can learn a peer-end device as its neighbor. One port may learn one neighbor, which is variable. If negotiation is enabled, unidirectional or bidirectional link detection starts after a port finds a neighbor through negotiation, which succeeds when a port receives a Prob packet from the neighbor. However, if the RLDP is enabled under a failure, the port cannot learn a neighbor so that detection cannot start. In this case, recover the link state before enabling the RLDP.

Treatment for Failed Port under RLDP

- Warning: Only print Syslog to indicate a failed port and a failure type.

- Shutdown SVI: Print Syslog, and then inquire an SVI according to the Access VLAN or Native VLAN of a port and shut down the SVI if the port is a physical exchange port or layer-2 AP member port.
- Port violation: Print Syslog, and configure a failed port as in violation state, and the port will enter Linkdown state physically.
- Block: Print Syslog, and configure the forward state of a port as Block, and the port will not forward packets.

Recovery of Failed Port under RLDP

- Manual reset: Manually reset all failed ports to initialized state and restart link detection.
- Manual or automatic errdisable recovery: Recover all failed ports to initialized state manually or regularly (30s by default and configurable) and restart link detection.
- Automatic recovery: Under unidirectional or bidirectional link detection, if the treatment for failed ports is not specified as port violation, recover ports to initialized state based on Prob packets and restart link detection.

Port State under RLDP

- normal: Indicates the state of a port after link detection is enabled.
- error: Indicates the state of a port after a unidirectional or bidirectional link failure or a loop failure is detected.

Overview

Feature	Description
Deploying RLDP Detection	Enable unidirectional or bidirectional link detection or downlink loop detection for failures and implement treatment.

2.3.1 Deploying RLDP Detection

The RLDP provides unidirectional link detection, bidirectional forwarding detection and downlink loop detection.

Working Principle

Unidirectional Link Detection

When this function is enabled, a port sends Prob packets and receives Echo packets from a neighbor regularly as well as receiving Prob packets from a neighbor and replying with Echo packets. Within the maximum detection time, if the port receives Prob packets but no Echo packets, or none of them, treatment for a unidirectional failure will be triggered and detection will stop.

Bidirectional Forwarding Detection

When this function is enabled, a port sends Prob packets and receives Echo packets from a neighbor regularly as well as receiving Prob packets from a neighbor and replying with Echo packets. Within the maximum detection time, if the port receives neither Prob packets nor Echo packets from a neighbor, treatment for a bidirectional failure will be triggered and detection will stop.

Downlink Loop Detection

When this function is enabled, a port sends Loop packets regularly. In the following cases, a loop failure will be triggered after the same port or a different port receives the packets: in one case, the egress and ingress ports are the same routed port or layer-3 AP member port; in another case, the egress and ingress ports are exchange ports or layer-2 AP member ports in a same default VLAN and in Forward state. Treatment for the failure will be implemented and detection will stop.

Related Configuration

- Configuring RLDP Detection

By default, RLDP detection is disabled.

You may run the global command **rldp enable** or the interface command **rldp port** to enable RLDP detection and specify a detection type and treatment.

You may run the **rldp neighbor-negotiation** command to neighbor negotiation, the **rldp detect-interval** to specify a detection interval, the **rldp detect-max** to specify detection times, the **rldp error-recover interval** to regularly recover the state of the failed port, or the **rldp reset** to recover a failed port.

2.4 Configuration

Configuration	Description and Command	
Configuring Basic RLDP Functions	(Mandatory) It is used to enable RLDP detection under global configuration mode.	
	rldp enable	Enables global RLDP detection on all ports.
	(Mandatory)It is used to specify under interface configuration mode a detection type and failure treatment for an interface.	
	rldp port	Enables RLDP detection on a port and specifies a detection type and failure treatment.
	(Optional)It is used to configure a detection interval, detection times and neighbor negotiation under global configuration mode.	
	rldp detect-interval	Modifies global RLDP parameters on all ports, such as the detection interval, maximum detection times and neighbor negotiation.
	rldp detect-max	
	rldp neighbor-negotiation	
	rldp error-recover interval	Configures the global time interval for recovering the state of the failed port.
	(Optional) It is used under privileged mode.	
	rldp reset	Recovers all ports.

2.4.1 Configuring Basic RLDP Functions

Configuration Effect

- Enable RLDP unidirectional link detection, bidirectional forwarding detection, downlink loop detection, or VLAN-based loop detection to discover loop failures.

Notes

- Loop detection is effective to all member ports of an AP when configured on one of the ports. Unidirectional link detection and bidirectional forwarding detection are effective only on an AP member port.
- The loop detection on a physical port added to an AP shall be configured the same as that of the other member ports. There are

three cases. First, if loop detection is not configured on a newly-added port but on the existing member ports, the new port adopts the configuration and detection results of the existing ports. Second, if a newly-added port and the existing member ports have different loop detection configuration, the new port adopts the configuration and detection results of the existing ports.

- When configuring the RLDP on an AP port, you may configure failure treatment only as "shutdown-port", to which other configurations will be modified.
- When "shutdown-port" is configured on a port, RLDP detection cannot be restored in case of a failure. After troubleshooting, you may run the **rldp reset** or **errdisable recovery** command to restore the port and resume detection. For configuration of the **errdisable recovery** command, please refer to the *Configuring Interface*.
- The VLAN-based loop detection function (configured by the **rldp port vlan-loop-detec** command) can be configured only on multi-VLAN ports such as trunk and hybrid ports. It is recommended not to configure this function on access or routed ports. This function does not take effect when configured on routed ports.
- It is recommended not to configure the **rldp port vlan-loop-detect** command and the **rldp port loop-detect** command on the same port. The **rldp port loop-detect** command is mainly configured on access ports to detect downlink loops.
- The VLAN list for the VLAN-based loop detection function needs to be configured in consideration of performance, and should include only VLANs that possibly encounter loops to reduce invalid detection and improve detection performance.
- The VLAN-based loop detection function can be configured only on AP member ports. It is recommended to configure the same detection VLAN on all member ports of an AP to ensure normal loop detection.

Configuration Steps

Enabling RLDP

- Mandatory.
- Enable RLDP detection on all ports under global configuration mode.

Enabling Neighbor Negotiation

- Optional.
- Enable the function under global configuration mode, and port detection will be started under successful neighbor negotiation.

Configuring Detection Interval

- Optional.
- Configure a detection interval under global configuration mode.

Configuring Maximum Detection Times

- Optional.
- Specify the maximum detection times under global configuration mode.

Configuring a Time Interval for Recovering State of the Failed Port

- Optional.
- Configure the time interval for recovering the state of the failed port under global configuration mode.

Configuring Detection under Port

- Mandatory.
- Configure unidirectional RLDP detection, bidirectional RLDP detection or downlink loop detection under interface configuration mode, and specify failure treatment.

Restoring All Failed Ports

- Optional.
- Enable this function under privileged mode to restore all failed ports and resume detection.

Verification

- Display the information of global RLDP, port and neighbor.

Related Commands

Enabling Global RLDP Detection

Command	**rldp enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	Enable global RLDP detection.

Enabling RLDP Detection on Interface

Command	**rldp port { unidirection-detect | bidirection-detect | loop-detect } { warning | shutdown-svi | shutdown-port | block }** **rldp port vlan-loop-detect { warning | isolate-vlan} vlan** *vlan-list*
Parameter Description	**unidirection-detect**: Indicates unidirectional link detection. **bidirection-detect**: Indicates bidirectional forwarding detection. **loop-detect**: Indicates downlink loop detection. **vlan-loop-detect**: Indicates VLAN-based loop detection. **warning**: Indicate the failure treatment is warning. **shutdown-svi**: Indicate the failure treatment is closing the SVI that the interface is on. **shutdown-port**: Indicates the failure treatment is port violation. **block**: Indicates the failure treatment is disabling learning and forwarding of a port. **isolate-vlan**: Indicates that the failure treatment is isolation for the faulty VLAN. **vlan-list**: Indicates the VLANs to be detected.
Command Mode	Interface configuration mode
Usage Guide	The interfaces include layer-2 switch ports, layer-3 routed ports, layer-2 AP member ports, and layer-3 AP member ports.

Modifying Global RLDP Detection Parameters

Command	**rldp** {**detect-interval** *interval* | **detect-max** *num* | **neighbor-negotiation** }
Parameter Description	**detect-interval** *interval*: Indicates a detection interval. **detect-max** *num*: Indicates detection times. **neighbor-negotiation**: Indicates neighbor negotiation.
Command Mode	Global configuration mode
Usage Guide	Modify all RLDP parameters on all ports when necessary.

Recovering Failed Port

Command	**rldp reset**
Parameter Description	N/A
Command Mode	Privileged mode
Usage Guide	Recover all failed ports to initialized state and resume detection.

Displaying RLDP State Information

Command	**show rldp** [**interface** *interface-name*]
Parameter Description	*interface-name*: Indicates the interface to display information of.
Command Mode	Privileged mode, global configuration mode, or interface configuration mode
Usage Guide	Display RLDP state information.

Configuration Example

Enabling RLDP Detection in Ring Topology

Scenario **Figure 2- 4**	As shown in the following figure, the aggregation and access sections are in a ring topology. The STP is enabled on all devices to prevent loop and provide redundancy protection. To avoid a unidirectional or bidirectional link failure resulting in STP failure, RLDP unidirectional and bidirectional link detection is enabled between aggregation devices as well as between an aggregation device and the access device. To avoid loop due to wrong downlink connection of the aggregation devices, enable RLDP downlink loop detection on the downlink ports of the aggregation devices and of the access device. To avoid loop due to wrong downlink connection of the access device, enable RLDP downlink loop detection on the downlink ports of the access device.

Configuration Steps	● SW A and SW B are aggregation devices, and SW C is an access device. Users connected to SW C. SW A, SW B and SW C are structured in a ring topology, and the STP is enabled on each of them. For STP configuration, refer to relevant configuration guide. ● Enable the RLDP on SW A, enable unidirectional and bidirectional link detection on the two ports, and enable loop detection on the downlink port. ● Enable the RLDP on SW B, enable unidirectional and bidirectional link detection on the two ports, and enable loop detection on the downlink port. ● Enable the RLDP on SW C, enable unidirectional and bidirectional link detection on the two uplink ports, and enable loop detection on the two downlink ports.
A	A#configure terminal A(config)#rldp enable A(config)#interface GigabitEthernet 2/0/1 A(config-if-GigabitEthernet 2/0/1)#rldp port unidirection-detect shutdown-port A(config-if-GigabitEthernet 2/0/1)#rldp port bidirection-detect shutdown-port A(config-if-GigabitEthernet 2/0/1)# exit A(config)#interface GigabitEthernet 2/0/9 A(config-if-GigabitEthernet 2/0/1)#rldp port unidirection-detect shutdown-port A(config-if-GigabitEthernet 2/0/1)#rldp port bidirection-detect shutdown-port A(config-if-GigabitEthernet 2/0/1)#rldp port loop-detect shutdown-port A(config-if-GigabitEthernet 2/0/1)#exit
B	Apply the configuration on SW A.
C	C#configure terminal C(config)#rldp enable C(config)#interface GigabitEthernet 0/49 C(config-if-GigabitEthernet 0/49)#rldp port unidirection-detect shutdown-port C(config-if-GigabitEthernet 0/49)#rldp port bidirection-detect shutdown-port C(config-if-GigabitEthernet 0/49)# exit

	C(config)#interface GigabitEthernet 0/50 C(config-if-GigabitEthernet 0/50)#rldp port unidirection-detect shutdown-port C(config-if-GigabitEthernet 0/50)#rldp port bidirection-detect shutdown-port C(config-if-GigabitEthernet 0/50)#exit C(config)#interface GigabitEthernet 0/1 C(config-if-GigabitEthernet 0/1)# rldp port loop-detect shutdown-port C(config-if-GigabitEthernet 0/1)#exit C(config)#interface GigabitEthernet 0/2 C(config-if-GigabitEthernet 0/2)# rldp port loop-detect shutdown-port C(config-if-GigabitEthernet 0/2)#exit
Verification	● Check the RLDP information on SW A, SW B and SW C. Take SW A for example.
A	A#show rldp rldp state : enable rldp hello interval: 3 rldp max hello : 2 rldp local bridge : 00d0.f822.33aa ---------- Interface GigabitEthernet 2/0/1 port state : normal neighbor bridge : 00d0.f800.51b1 neighbor port : GigabitEthernet 2/0/1 unidirection detect information: action: shutdown-port state : normal bidirection detect information: action: shutdown-port state : normal Interface GigabitEthernet 2/0/9 port state : normal neighbor bridge : 00d0.f800.41b0 neighbor port : GigabitEthernet 0/49 unidirection detect information: action: shutdown-port

	state : normal bidirection detect information: action: shutdown-port state : normal loop detect information: action: shutdown-port state : normal

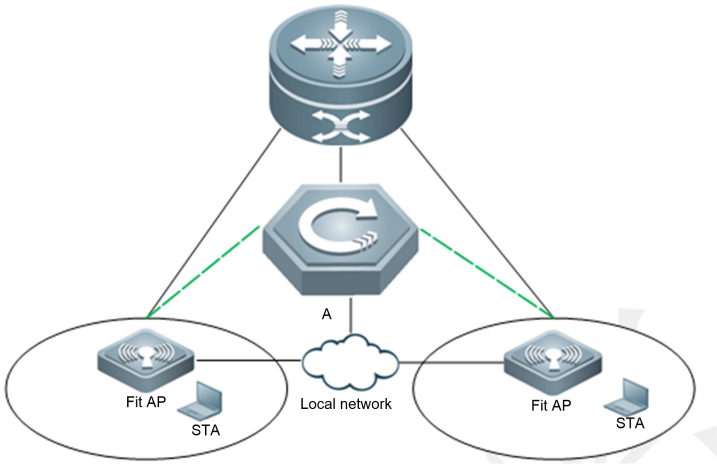
Common Errors

- RLDP functions and private multicast address authentication or TPP are enabled at the same time.
- Neighbor negotiation is not enabled when configuring unidirectional or bidirectional link detection. The RLDP should be enabled on a neighbor device, or otherwise a unidirectional or bidirectional failure will be detected.
- If RLDP detection is configured to be implemented after neighbor negotiation while configuring unidirectional or bidirectional link detection, detection cannot be implemented as no neighbor can be learned due to a link failure. In this situation, you are suggested to recover the link state first.
- You are suggested not to specify the failure treatment as Shutdown SVI under a routed port.
- You are suggested not to specify the failure treatment as Block for a port, on which a loop protection protocol is enabled, for example, the STP.

Common Errors

- When the **exec-cmd** command is executed for interface configuration, the input of the corresponding AP wired port is incorrect.
- When the RLDP loop detection configurations are modified, the **no exec-cmd** command is not executed to delete the original configurations or the **exec-cmd** command is not re-executed to cancel the configurations.

Configuring RLDP Loop Detection on Wireless APs

Scenario **Figure 2-5**	As shown in the following figure, a large number of APs exist in the wireless AP scenario. If the RLDP loop detection function is configured and modified on APs one by one, the workload is heavy. The RLDP loop detection configurations can be pushed from the AC device to all online APs (or an independent AP). 

Configuration Steps	● Log in to the AC device and enter the AP configuration mode. ● Enable the RLDP loop detection function on the wired ports of the corresponding AP. ● Enable the RLDP function on corresponding APs in global configuration mode. ● On corresponding APs, configure the recovery time for the RLDP violated port.
A	A#configure terminal A(config)#ap-config all A(config-ap)#exec-cmd mode "int gi 0/1" cmd "rldp port loop-detect shutdown-port" A(config-ap)#exec-cmd mode configure cmd "rldp enable" A(config-ap)#exec-cmd mode configure cmd "errdisable recovery interval 600"
Verification	● On the AC device, check the RLDP loop detection configurations.

A	A# show run
	!
	ap-config all
	  exec-cmd mode "int gi 0/1" cmd "rldp port loop-detect shutdown-port"
	  exec-cmd mode configure cmd "rldp enable"
	  exec-cmd mode configure cmd "errdisable recovery interval 600"
	!

Common Errors

- When the **exec-cmd** command is executed for interface configuration, the input of the corresponding AP wired port is incorrect.
- When the RLDP loop detection configurations are modified, the **no exec-cmd** command is not executed to delete the original configurations or the **exec-cmd** command is not re-executed to cancel the configurations.

2.5 Monitoring

Displaying

Description	Command
Displays RLDP state.	**show rldp** [**interface** *interface-name*]

3 Configuring DLDP

3.1 Overview

The Data Link Detection Protocol (DLDP) is a protocol used to quickly detect faulty Ethernet links.

A typical Ethernet link detection mechanism detects physical link connectivity through autonegotiation at the physical layer. Such a mechanism has limitations when detecting Layer-3 data communication exceptions despite normal physical connections.

DLDP provides reliable Layer-3 link detection information. After detecting a faulty link, DLDP shuts down the logical state of Layer-3 ports to realize fast Layer-3 protocol convergence.

3.2 Applications

Application	Description
Intra-Network Segment DLDP Detection	The source IP address of the detected port and the detected IP address are in the same network segment.
Inter-Network Segment DLDP Detection	The source IP address of the detected port and the detected IP address are in different network segments.

3.2.1 Intra-Network Segment DLDP Detection

Scenario

This section describes the basic DLDP application scenario where the source IP address of the detected port and the detected IP address are in the same network segment.

In Figure 3- 1, the Gi 0/1 Layer-3 port on Switch A and the Gi 0/2 Layer-3 port on Switch C are in the same network segment. To detect the Layer-3 link connectivity from Gi 0/1 to Gi 0/2, enable DLDP on Gi 0/1 or Gi 0/2.

Figure 3- 1

Remarks	Gi 0/1 and Gi 0/2 are Layer-3 ports in the same network segment. B is a network in the same network segment as Gi 0/1 and Gi 0/2.

Deployment

- Enable DLDP on Gi 0/1 or Gi 0/2.

3.2.2 Inter-Network Segment DLDP Detection

Scenario

This section describes the DLDP application scenario where the source IP address of the detected port and the detected IP address are in different network segments.

In Figure 3- 2, the Gi 0/1 Layer-3 port on Switch A and the Gi 0/4 Layer-3 port on Switch D are in different network segments. To detect the Layer-3 link connectivity from Gi 0/1 to Gi 0/4, enable DLDP on Gi 0/1 and configure the DLDP next-hop IP address (IP address of the Gi 0/2 port on Switch B).

Figure 3- 2

Remarks	Gi 0/1 and Gi 0/4 are Layer-3 ports in different network segments.

Deployment

- Enable DLDP on Gi 0/1 and configure the DLDP next-hop IP address.

3.3 Features

Basic Concepts

DLDP Detection Interval and Retransmission Times

Detection interval: Indicates the interval at which DLDP detection packets (ICMP echo) are transmitted.

Retransmission times: Indicate the maximum times DLDP detection packets can be retransmitted in the case of a DLDP detection failure.

When a network device does not receive a reply packet from the peer end within the period of the detection interval multiplied by the retransmission times, the device determines that a Layer-3 link failure occurs and shuts down the logical state of its Layer-3 port (despite the normal physical link connection). When Layer-3 link connectivity is recovered, the device restores its Layer-3 port to Up logical state.

DLDP Detection Modes

Active mode and passive mode are two DLDP detection modes.

Active mode (default): ICMP detection packets are sent actively.

Passive mode: ICMP detection packets are received passively.

DLDP Next Hop

Next hop: Indicates the next node connected to the detected IP address in inter-network segment DLDP detection.

In some cases, DLDP needs to detect IP reachability in non-directly connected network segments. You need to configure the next-hop IP address for the detected port to allow DLDP to obtain the next-hop MAC address through an ARP packet before sending a correct ICMP packet.

In this situation, you need to avoid the return of the reply packet from another link; otherwise, DLDP will misjudge that the detected port does not receive an ICMP reply.

DLDP Recovery Times

Recovery times: Indicate the times DLDP needs to receive consecutive reply packets (ICMP reply) before it can determine link failure recovery.

In some cases, link detection may be unstable. For example, a link is only intermittently pingable. In this case, DLDP repeatedly changes the link status between Up and Down, which may further destabilize the ring network.

Recovery times indicate the times DLDP needs to receive consecutive reply packets before DLDP can set the link in Down state to Up. The default recovery times are three times, indicating that the link needs to be successfully pinged three times before it is set to Up. The recovery times setting reduces link detection sensitivity but increases stability. Related parameters are adjustable according to the network condition.

DLDP Bound MAC Address

Bound MAC address: Indicates the MAC address bound to the detected IP address.

In a complex network environment, DLDP may obtain an invalid MAC address if the detected link has abnormal ARP packets transmitted (causing ARP spoofing), which will make DLDP detection abnormal.

To address this problem, you can bind the detected IP address (or next-hop IP address) to a static MAC address to avoid a DLDP failure in the case of ARP spoofing.

Overview

Feature	Description
DLDP Detection	Detects Layer-3 link connectivity. When a Layer-3 link is abnormal, DLDP shuts down the Layer-3 port.
MAC Address Binding	Binds the detected IP address to the MAC address of the detected device to avoid DLDP exceptions otherwise caused by ARP spoofing.
Passive DLDP Detection	When both ends of the detected link are enabled with DLDP, you can configure one end in passive mode to save bandwidth and CPU resources.

3.3.1 DLDP Detection

DLDP detects Layer-3 link connectivity. When a Layer-3 link is abnormal, DLDP shuts down the corresponding Layer-3 port.

Working Principle

After DLDP detection is enabled, DLDP sends an ARP packet to obtain the MAC address and outbound port of the detected device or the next-hop device. Then DLDP periodically sends IPv4 ICMP echo packets to the MAC address and outbound port to detect link connectivity. If DLDP does not receive an IPv4 ICMP reply packet from the detected device within a specific period, DLDP determines that the link is abnormal and sets the Layer-3 port to Down.

Related Configuration

- Enabling DLDP Detection

By default, DLDP detection is disabled on ports.

Run the **dldp** command with the detected IP address specified to enable DLDP detection.

You can configure the next-hop IP address, MAC address of the detected device, transmission interval, retransmission times, and recovery times based on the actual environment.

3.3.2 MAC Address Binding

The MAC address binding feature is used to bind the detected IP address (or next-hop IP address) to the MAC address of the detected device (or next-hop device) to avoid DLDP exceptions otherwise caused by ARP spoofing

Working Principle

You can bind the detected IP address (or next-hop IP address) to a static MAC address to avoid a DLDP failure in the case of ARP spoofing.

Related Configuration

By default, no MAC address is bound in DLDP detection.

Bind the MAC address of the detected device when you run the **dldp** command to enable DLDP detection. If the next-hop IP address is specified, bind the MAC address of the next-hop device.

After DLDP detection is enabled, DLDP sends ARP packets and ICMP packets with a fixed destination IP address and a fixed destination MAC address. If the source IP address and MAC address in the received packet do not match the bound IP address and MAC address, DLDP will not process the packet.

3.3.3 Passive DLDP Detection

When both ends of the detected link are enabled with DLDP, you can configure one end in passive mode to save bandwidth and CPU resources.

Working Principle

After the device at the local end sends an ICMP echo packet, the peer device determines link connectivity according to the packet reception time by using specific detection parameters, which are the same as those at the local end, thus saving bandwidth and CPU resources.

Related Configuration

By default, passive DLDP detection is disabled.

Run the **dldp passive** command to enable passive DLDP detection.

After passive DLDP detection is enabled, DLDP will return an ICMP reply packet upon receiving an ICMP echo packet, instead of actively sending ICMP echo packets to the peer end. If DLDP does not receive an ICMP echo packet within a specific period, it determines that the link to the peer port is abnormal.

3.4 Configuration

Configuration	Description and Command	
Enabling DLDP Detection	(Mandatory) It is used to enable DLDP detection in interface configuration mode.	
	dldp	Enables DLDP detection.
	(Mandatory) It is used to enable passive DLDP detection in interface configuration mode.	
	dldp passive	Enables passive DLDP detection.
	(Optional) It is used to configure the detection interval, retransmission times, and recovery times of DLDP detection in global configuration mode.	
	dldp interval	Modifies the DLDP parameters globally to apply the modifications to DLDP detection on all ports.
	dldp retry	
	dldp resume	

3.4.1 Enabling DLDP Detection

Configuration Effect

- Detect Layer-3 link connectivity. When a Layer-3 link is abnormal, DLDP shuts down the Layer-3 port.

Notes

- DLDP supports the configuration of multiple IP addresses on a Layer-3 port. DLDP sets the port to Down when none of the IP addresses receives an ICMP reply. If one IP address resumes communication, DLDP sets the port to Up again.
- DLDP uses the first IP address of the Layer-3 port as the source IP address of detection packets.

Configuration Steps

Enabling DLDP Detection

- Mandatory.
- When you enable DLDP detection in interface configuration mode, you can configure the next-hop IP address, MAC address, transmission interval, retransmission times, and recovery times based on the actual environment.

Configuring a DLDP Detection Mode

- Optional.
- You can configure active or passive DLDP detection in interface configuration mode based on the actual environment.
- If DLDP detection needs to be enabled at both ends of a Layer-3 link, you can configure passive DLDP detection at one end to save bandwidth and CPU resources.

Configuring DLDP Parameters Globally

- Optional.
- You can modify the parameters of DLDP detection on all ports in global configuration mode based on requirements. The parameters include the packet transmission interval, packet retransmission times, and recovery times.

Verification

- Display the device DLDP information, including the status and statistics of DLDP detection on all ports.

Related Commands

Enabling DLDP Detection

Command	**dldp** *ip-address* [*next-hop-ip*] [**mac-address** *mac-addr*] [**interval** *tick*] [**retry** *retry-num*] [**resume** *resume-num*]
Parameter Description	*ip-address*: Indicates the detected IP address. ***next-hop-ip***: Indicates the next-hop IP address. *mac-addr*: Indicates the MAC address of the detected device to be bound. If the next-hop IP address is specified, bind the MAC address of the next-hop device. *tick*: Indicates the interval at which detection packets are transmitted. The value ranges from 5 to 6,000 ticks (1 tick = 10 ms). The default value is 100 ticks (1s).

	retry-num: The value ranges from 1 to 3,600. The default value is 4. *resume-num*: Indicates the recovery times. The value ranges from 1 to 200. The default value is 3.
Command Mode	Interface configuration mode
Usage Guide	The port to be enabled with DLDP detection must be a Layer-3 port, such as a router port, L3AP port, and SVI port.

Configuring a DLDP Detection Mode

Command	**dldp passive**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	You need to enable DLDP detection before configuring a DLDP detection mode.

Modifying DLDP Detection Parameters Globally

Command	**dldp** { **interval** *tick* | **retry** *retry-num* | **resume** *resume-num* }
Parameter Description	*tick*: Indicates the interval at which detection packets are transmitted. The value ranges from 5 to 6,000 ticks (1 tick = 10 ms). The default value is 100 ticks (1s). *retry-num*: Indicates the interval at which detection packets are retransmitted. The value ranges from 5 to 3,600. The default value is 4. *resume-num:* Indicates the recovery times. The value ranges from 1 to 200. The default value is 3.
Command Mode	Global configuration mode
Usage Guide	Use this command to quickly modify the parameters of DLDP detection on all ports when the actual environment is changed.

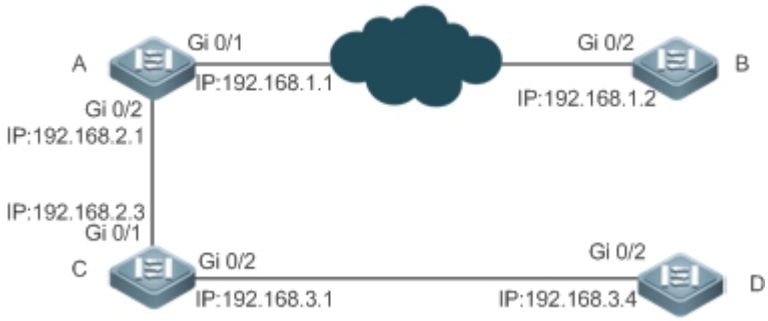
Displaying the DLDP Status

Command	**show dldp statistic** [**interface** *interface-name*]
Parameter Description	*interface-name*: Indicates the Layer-3 port on which the DLDP status will be displayed.
Command Mode	Privileged mode, global configuration mode, and interface configuration mode
Usage Guide	Use this command to display the DLDP status on a specific port. You can also use this command to display the DLDP status on all ports.

Configuration Example

Enabling DLDP Detection on Layer-3 Ports on Device A and Device B in a Layer-3 Network

Scenario **Figure 3- 3**	



Verification

- Enable DLDP detection on the Gi 0/1 and Gi 0/2 router ports on Device A to detect the Layer-3 link connectivity between Device A and Device B and that between Device A and Device D.
- To control the Gi 0/2 router port of Device B, enable passive DLDP detection on the port.

A

```
A#configure terminal
A(config)#interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#dldp 192.168.1.2
A(config-if-GigabitEthernet 0/1)# exit
A(config)#interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/1)#dldp 192.168.3.4 192.168.2.3
```

B

```
B#configure terminal
B(config)#interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/1)#dldp 192.168.1.1
B(config-if-GigabitEthernet 0/1)#dldp passive
```

Verification

- Display the DLDP status on Device A and Device B to check whether DLDP detection is enabled and works normally.

A

```
A# show dldp
Interface   Type        Ip          Next-hop     Interval   Retry   Resume   State
---------   -------     -----------  -----------  --------   -----   ------   ------
Gi0/1       Active      192.168.1.2                100        4       3        Up
Gi0/1       Active      192.168.3.4  192.168.2.3   100        4       3        Up
```

B

```
B# show dldp
Interface   Type        Ip          Next-hop     Interval   Retry   Resume   State
---------   -------     -----------  -----------  --------   -----   ------   ------
Gi0/2       Passive     192.168.1.1                100        4       3        Up
```

Common Errors

- An unreachable IPv4 unicast route is misjudged as a DLDP detection failure.
- DLDP detection fails because the peer device does not support ARP/ICMP replies.

- No next-hop IP address is configured in inter-network segment DLDP detection.

3.5 Monitoring

Clearing

Description	Command
Clears DLDP statistics.	**clear dldp** [**interface** *interface-name* [*ip-address*]]

Displaying

Description	Command
Displays the DLDP status.	**show dldp** [**interface** *interface-name*]
Displays the DLDP statistics on the Up/Down port sates.	**show dldp statistic**

4 Configuring VRRP

4.1 Overview

Virtual Router Redundancy Protocol (VRRP) is a fault-tolerant routing protocol.

VRRP adopts the master-backup design to ensure migration of functions from a Master router to a Backup one when the Master failed, without influencing internal and external data communication or modifying Local Area Network (LAN) configuration. A VRRP group maps multiple routers into a virtual router. VRRP ensures only one router at a moment on behalf of a virtual router transfers packets, which is the elected Master. If the Master fails, one of the Backup routers will replace it. Under VRRP, it seems that a host in a LAN uses only one router and the routing remains functional even when the first-hop router fails.

- VRRP is applicable to LAN scenarios which require the redundancy of routing egresses.

Protocols and Standards

- RFC2338: Virtual Router Redundancy Protocol
- RFC3768: Virtual Router Redundancy Protocol (VRRP)
- RFC5798: Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6

4.2 Applications

Application	Description
Routing Redundancy	Configure routers in a LAN as one VRRP group to achieve simple routing redundancy.
Load Balancing	Configure routers in a LAN as multiple VRRP groups to achieve traffic load balancing.

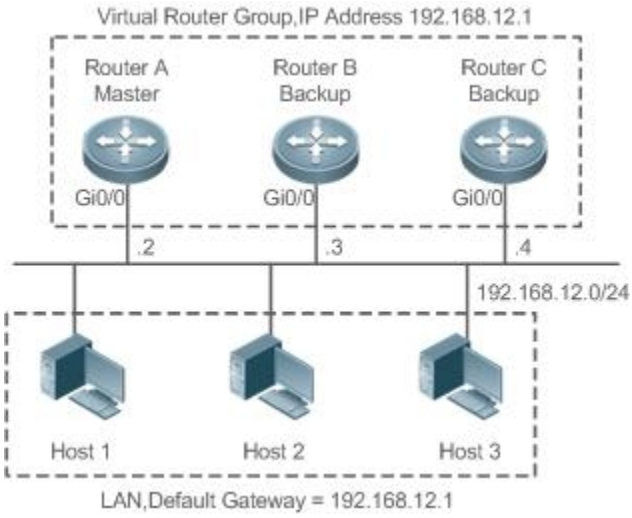
4.2.1 Routing Redundancy

Scenario

Configure routers in a LAN as one VRRP group, where hosts take the virtual IP address of this group as the default gateway address.

- Packets from Host 1, Host 2 and Host 3 to other networks are forwarded by the elected Master router (Router A in Figure 4- 1).
- If Router A fails, the Master will be re-elected between Router B and Router C to forward packets, achieving simple routing redundancy.

Figure 4- 1



Deployment

- Router A, Router B and Router C are connected to the LAN via Ethernet interfaces.
- On Router A, Router B and Router C, VRRP is configured on the Ethernet interfaces connected to the LAN.
- These Ethernet interfaces are in the same VRRP group whose virtual IP address is 192.168.12.1.
- The gateway address for Host 1, Host 2 and Host 3 is the IP address of the VRRP group, namely 192.168.12.1.

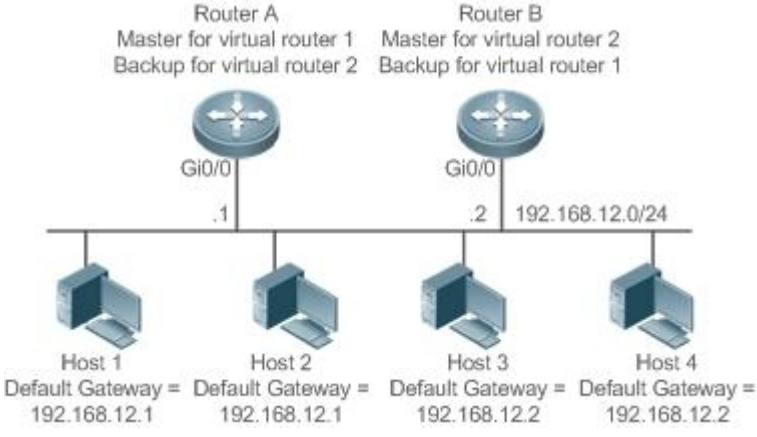
4.2.2 Load Balancing

Scenario

Configure routers in a LAN as multiple VRRP groups. Hosts in the LAN take virtual IP addresses of the groups as their gateways, and each router backs up for other routers in different group.

- Packets from Host 1 and Host 2 to other networks with the default gateway address as the virtual IP address of virtual router 1 are forwarded by the Master of virtual router 1 (Router A in Figure 4- 2).
- Packets from Host 3 and Host 4 to other networks with the default gateway address as the virtual IP address of virtual router 2 are forwarded by the Master of virtual router 2 (Router B in Figure 4- 2).
- Routing redundancy is achieved on Router A and Router B, and the LAN traffic is shared to achieve load balancing.

Figure 4- 2



Deployment

- Router A and Router B are connected to the LAN via Ethernet interfaces.
- On Router A and Router B, two virtual routers are configured on the Ethernet interfaces connected to the LAN.
- Router A takes the IP address 192.168.12.1 of Ethernet interface Gi0/0 as the IP address of virtual router 1. Thus for virtual router 1, Router A becomes the Master and Router B becomes the Backup.
- Router B takes the IP address 192.168.12.2 of Ethernet interface Gi0/0 as the IP address of virtual router 2. Thus for virtual router 2, Router B becomes the Master and Router A becomes the Backup.
- In the LAN, Host 1 and Host 2 take the IP address 192.168.12.1 of virtual router 1 as the default gateway address, while Host 3 and Host 4 take the IP address 192.168.12.2 of virtual router 2 as the default gateway address.

4.3 Features

Basic Concepts

Virtual Router

A virtual router, also called a VRRP group, is regarded as a default gateway for hosts in a LAN. A VRRP group contains a Virtual Router Identifier (VRID) and a set of virtual IP addresses.

Virtual IP Address

Indicates the IP address of a virtual router. A virtual router can be configured with one or multiple IP addresses.

IP Address Owner

If a VRRP group has the virtual IP address as that of an Ethernet interface on one real router, the router is regarded as the virtual IP address owner. In such case, the router priority is 255. If the owned Ethernet interface is available, the VRRP group will be in Master state automatically. The IP address owner receives and processes the packets with the destination IP address as that of the virtual router.

Virtual MAC Address

The virtual MAC address of a VRRP group is an IEEE 802 MAC address, formatted as **00-00-5E-00-01-{VRID}** with the first five octets assigned and the last two as a group VRID. A VRRP group responds to an Address Resolution Protocol (ARP) request with its virtual MAC address instead of a real MAC address.

Master Router

In a VRRP group, only the Master router answers ARP requests and forwards IP packets. If a real router is the IP Address Owner, it becomes the Master router.

Backup Router

In a VRRP group, Backup routers only monitor the state of the Master but do not respond to ARP requests or forward IP packets. When the Master fails, Backup routers will take the chance to compete for the position.

Preemption Mode

If a VRRP group runs in Preemption mode, a higher priority Backup router will replace the lower priority Master router.

Overview

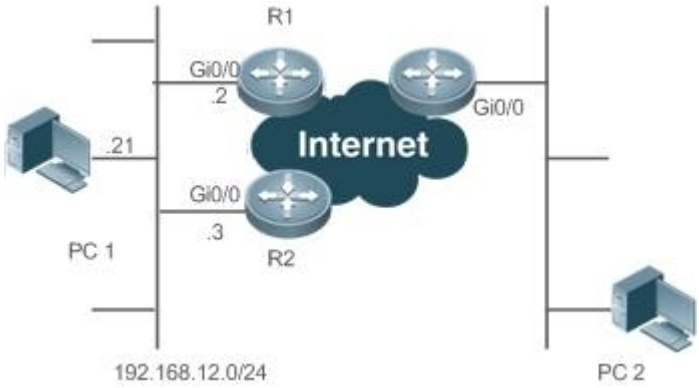
Feature	Description
VRRP	VRRP achieves redundancy for the default gateways of terminals on a multi-access media (for example, Ethernet). It enables a Backup router to forward packets when the Master router is down, providing transparent routing switch and promoting network service quality.

4.3.1 VRRP

In case that the Master router is faulty, VRRP achieves migration of functions from the Master router to a Backup one without influencing internal and external data communication or modifying LAN configuration.

Working Principle

Figure 4- 3 Working Principle of VRRP



Working Mode of VRRP

The RFC2338, RFC3768 and RFC5798 protocols define the format and operating mechanism of VRRP packets. Multicast VRRP packets are sent periodically with specified destination addresses by the Master router to advertise normal operation or for Master election. VRRP allows a router in a LAN to automatically replace the Master who forwards IP packets when the latter fails. This helps achieve hot backup and fault tolerance of IP-based routing as well as ensure communication continuity and reliability for hosts in the LAN. A VRRP group achieves redundancy through multiple real routers. However, only one router acts as the Master to forward packets while the others are Backup routers. Router switching in a VRRP group is completely transparent to hosts in a LAN.

Master Election Process

The RFC standards stipulate the master election process as follows:

- VRRP provides a simple mechanism for Master election. First, compare the VRRP priorities configured on the interfaces of the routers in a VRRP group. The router with the highest priority is elected as the Master. If these priorities are equal, compare the primary IP addresses of these routers. The router with the biggest IP address is elected as the Master.
- After the Master router is elected, the other routers become Backup routers (and enter the **Backup** state) and monitor the state of the master router through the VRRP packets the master router sends. If the master router is operational, it regularly sends VRRP multicast packets known as Advertisement packets to notify the Backup routers of its status. If the Backup routers do not receive such packets within a set period, all of them will enter the Master state. In such case, the previous step of Master election is repeated. In this way, a router with the highest priority will be elected as a new master, achieving VRRP backup.

Once the Master router of a VRRP group is elected, it is responsible to forward packets for hosts in a LAN.

Communication Process

The VRRP communication process can be explained by Figure 1-3. The routers R1 and R2 are connected to the LAN segment 192.168.12.0/24 via the VRRP-enabled Ethernet interfaces Gi0/0. Hosts in the LAN take the virtual IP address of the VRRP group as the default gateway address. Only the virtual router is recognized by the hosts. The Master router in the group, however, is unknown. For example, when PC 1 plan to communicate with PC 2, PC 1 sends packets to the default gateway with the virtual IP address; The Master router in the group receives the packets and forwards them to PC 2. In this process, PC 1 only senses the virtual router instead of R1 or R2. The Master router in the group is elected between R1 and R2. When the Master fails, it will be replaced automatically by the other router.

Related Configuration

Enabling VRRP

By default, VRRP is disabled on an interface.

In the interface configuration mode, run the **vrrp** *group* **ip** *ipaddress* [**secondary**] or **vrrp** *group* **ipv6** *ipv6-address* command to set the VRID and virtual IP address to enable VRRP.

VRRP must be enabled on an interface.

Configuring the IPv4 VRRP Authentication String

By default, VRRP is in non-authentication mode.

Run the **vrrp group authentication** *string* command to set an authentication string in MD5 authentication mode or a plain text password in plain text mode for an IPv4 VRRP group. In the plain text authentication mode, a password contains 8 bytes at most.

Members of a VRRP group can communicate with each other only when they are in the same authentication mode. In the plain text authentication mode, all routers in a VRRP group should have the same authentication password. The plain text authentication password cannot guarantee security but only prevents/prompts wrong VRRP configurations.

Configuring the VRRP Advertisement Interval

By default, the advertisement interval of the Master router is 1 second.

Run the **vrrp** [**ipv6**] *group* **timers advertise** *{ advertise-interval* | **csec** *centisecond-interval }* command to change the interval.

When VRRP learning timer is not configured, the same advertisement interval should be set for a VRRP group, otherwise routers in **Backup** state will discard received VRRP packets.

Configuring the VRRP Preemption Mode

By default, a VRRP group operates in the Preemption mode.

To enable the Preemption mode for a VRRP group, run the **vrrp** [**ipv6**] *group* **preempt** [**delay** *seconds*] command. The optional parameter **delay** *seconds* is 0 by default.

If a VRRP group operates in the Preemption mode, a router will become the Mater of the group when it finds that its priority is higher than that of the current Master. If a VRRP group operates in Non-preemption mode, a router will not become the Master even when it finds that its priority is higher than that of the current Master. It makes little sense to configure the Preemption mode when the VRRP group uses the IP address of an Ethernet interface, in which case the group has the highest priority and automatically becomes the Master in the group. The optional parameter **Delay** *Seconds* defines the delay before a backup VRRP router declares its Master identity.

Enabling the IPv6 VRRP Accept Mode

By default, the Accept mode is disabled for an IPv6 VRRP group.

To enable the Accept mode, run the **vrrp ipv6** *group* **accept_mode** command.

After the Accept mode is enabled, an IPv6 VRRP virtual router in **Master** state receives and processes packets with the virtual router IP address as the destination; when the Accept mode is disabled, the virtual router discards such packets except Neighbor Advertisement (NA) packets and Neighbor Solicitation (NS) packets. Besides, an IPv6 VRRP master virtual router in **Owner** state receives and processes packets with the virtual router IP address as the destination by default no matter whether the Accept mode is configured or not.

Configuring the VRRP Router Priority

By default, the router priorities in a VRRP group are all 100.

To adjust the priority, run the **vrrp** [**ipv6**] *group* **priority** *level* command.

If a router in the Preemption mode owns the group's virtual IP address and the highest priority, it becomes the group Master, while the other routers with lower priorities in the group become Backup (or monitoring) routers.

Configuring the VRRP Tracked Interface

By default, no interface is tracked by a VRRP group.

To configure such an interface, run the **vrrp** *group* **track** { *interface-type interface-number* } [*priority*] or **vrrp ipv6** ***group*** **track** *interface-type interface-number* **[** *priority* **]** command.

After an interface is configured for a VRRP group to monitor, the router priority will be adjusted dynamically based on the interface state. Once the interface becomes unavailable, the priority of the router in the group will be reduced by a set value, and another functional and higher priority router in this group will become the Master.

Configuring the VRRP Tracked IP Address

By default, no IP address is tracked by a VRRP group.

To configure such an address, run the **vrrp** *group* **track** *ip-address* [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*] or **vrrp ipv6** *group* **track {** *ipv6-global-address* | { *ipv6-linklocal-address interface-type interface-number* } } [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*] command.

After an IP address is configured for a VRRP group to monitor, the router priority will be adjusted dynamically based on the address accessibility. Once the address is inaccessible (the **ping** command fails), the priority of the router in the group will be reduced by a set value, and another higher priority router in this group will become the Master.

Configuring the VRRP Learning Timer

By default, the learning timer is disabled for a VRRP group.

To enable it, run the **vrrp** [**ipv6**] *group* **timers learn** command.

After the learning timer is configured, a VRRP Backup router learns the advertisement interval of NA packets from the Master. Based on this instead of a locally set interval, the Backup router calculates the interval for determining a failure of the Master. This command achieves the synchronization of advertisement intervals between Backup routers and the Master.

Configuring the VRRP Group Description

By default, no description is configured for a VRRP group.

To configure such a string, run the **vrrp** [**ipv6**] *group* **description** *text* command.

A VRRP description helps distinguishing VRRP groups. A description has 80 bytes at most, otherwise wrong configuration is prompted.

Configuring the VRRP Delay

By default, no delay is configured for a VRRP group.

To enable it, run the **vrrp delay** { **minimum** *min-seconds* | **reload** *reload-seconds* } command. The two types of delay range from 0 to 60 seconds.

The command configures the delay of starting a VRRP group on an interface. There are two types of VRRP delay: the delay after system startup and the delay after an interface resumes. You may configure them respectively or simultaneously. After the delay is configured for a VRRP group on an interface, the VRRP group starts after the delay instead of immediately upon system startup or the interface's resumption, ensuring non-preemption. If the interface receives a VRRP packet during the delay, the delay will be canceled and the VRRP will be started immediately. This configuration will be effective for both IPv4 and IPv6 VRRP groups of an interface.

Configuring the IPv4 VRRP Version

By default, IPv4 adopts the VRRPv2 standard.

To specify the version for IPv4 VRRP, run the **vrrp** *group* **version** { **2 | 3** } command.

When the parameter value is set to 2, VRRPv2 is adopted; when the parameter value is set to 3, VRRPv3 is adopted.

Specifying a Sub VLAN of a Super VLAN to Receive the IPv4 VRRP Packets

By default, IPv4 VRRP packets are sent to the first **Up** Sub VLAN interface of a Super VLAN.

To specify the first Sub VLAN in Up state of a Super VLAN to receive IPv4 VRRP packets, run the **vrrp detection-vlan first-subvlan** command; to specify a Sub VLAN, run the **vrrp detection-vlan** *subvlan-id* command. If VRRP and VRRP Plus are enabled simultaneously on a Super VLAN interface, VRRP packets are sent to all Up interfaces of the Sub VLANs under the Super VLAN.

Both the above configurations reduce VRRP packets and avoid influencing router performance and occupying network bandwidth. Yet the routers constituting an IPv4 VRRP group should be interconnected within the first UP Sub VLAN interface or a specified Sub VLAN of the Super VLAN.

Configuring the BFD Support for IPv4 VRRP on an Interface

By default, the Bidirectional Forwarding Detection (BFD) protocol support for VRRP is not enabled on an interface.

To enable it, run the **vrrp** *group* **bfd** *ip-address* command.

For a Backup router, run this command to correlate an IPv4 VRRP group with BFD without caring the configured IP address. For the Master, as the primary IP address of a Backup router is not known, the router IP address can only be specified by the administrator.

To enable the BFD support, make sure that IP and BFD session parameters are configured on the target interface.

After the BFD support is enabled for a specified IPv4 VRRP group, when the Master fails, a Backup router may detect it within one second.

Configuring Global IPv4 VRRP BFD

By default, the VRRP does not adopt the global IPv4 VRRP BFD mode in detecting the state of the Master.

To enable global IPv4 VRRP BFD, run the **vrrp bfd** *interface-type interface-number ip-address* command.

After global IPv4 VRRP BFD is enabled, multiple IPv4 VRRP groups may share BFD sessions, achieving fast detection and master-backup failover.

To enable the BFD support, make sure that IP and BFD session parameters are configured on the target interface.

4.4 Configuration

Configuration	Description and Command

Configuration	Description and Command	
Configuring IPv4 VRRP	(Mandatory) It is used to enable IPv4 VRRP.	
	vrrp *group* **ip** *ipaddress* [**secondary**]	Enables IPv4 VRRP.
	(Optional) It is used to configure IPv4 VRRP parameters.	
	vrrp *group* **authentication** *string*	Configures the IPv4 VRRP authentication string.
	vrrp *group* **timers advertise** { *advertise-interval* | **csec** *centisecond-interval* }	Configures the IPv4 VRRP advertisement interval and timeout times.
	vrrp *group* **preempt** [**delay** *seconds*]	Configures the IPv4 VRRP Preemption mode.
	vrrp *group* **priority** *level*	Configures the IPv4 VRRP router priority.
	vrrp group track { interface-type interface-number | bfd interface-type interface-number ipv4-addess } [priority]	Configures the IPv4 VRRP tracked interface.
	vrrp *group* **track** *ip-address* [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*]	Configures the IPv4 VRRP tracked IP address.
	vrrp *group* **timers learn**	Configures the IPv4 VRRP learning timer.
	vrrp *group* **description** *text*	Configures the IPv4 VRRP group description.
	vrrp delay { **minimum** *min-seconds* | **reload** *reload-seconds* }	Configures the IPv4 VRRP delay.
	vrrp *group* **version** { **2 | 3** }	Configures the IPv4 VRRP version.
	vrrp detection-vlan {first-subvlan | *subvlan-id***}**	Specifies a sub VLAN of a super VLAN to receive the IPv4 VRRP packets.
	vrrp *group* **bfd** *ip-address*	Configures the BFD support for IPv4 VRRP on an Interface.
	vrrp bfd *interface-type interface-number ip-address*	Configures global IPv4 VRRP BFD.
Configuring IPv6 VRRP	(Mandatory) It is used to enable IPv6 VRRP.	
	vrrp *group* **ipv6** *ipv6-address*	Enables IPv6 VRRP in interface configuration mode.
	(Optional) It is used to configure IPv6 VRRP parameters.	
	vrrp ipv6 *group* **timers advertise** { *advertise-interval* | **csec** *centisecond-interval* }	Configures the IPv6 advertisement interval and timeout times.
	vrrp ipv6 *group* **preempt** [**delay** *seconds*]	Configures the IPv6 VRRP Preemption mode.
	vrrp ipv6 *group* **accept_mode**	Enables the Accept mode for an IPv6 VRRP group.
	vrrp ipv6 *group* **priority** *level*	Configures the IPv6 VRRP router priority.
	vrrp ipv6 *group* **track** *interface-type interface-number* [*interface–priority*]	Configures the IPv6 VRRP tracked interface.

Configuration	Description and Command	
	vrrp ipv6 *group* **track** { *ipv6-global-address* | { *ipv6-linklocal-address interface-type interface-number* } } [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*]	Configures the IPv6 VRRP tracked IP address.
	vrrp ipv6 *group* **timers learn**	Configures the IPv6 VRRP learning timer.
	vrrp ipv6 *group* **description** *text*	Configures the IPv6 VRRP group description.
	vrrp delay { **minimum** *min-seconds* | **reload** *reload-seconds* }	Configures the IPv6 VRRP delay.
Configuring VRRP-MSTP	The configuration is the same as IPv4 VRRP configuration.	

4.4.1 Configuring IPv4 VRRP

Configuration Effect

- Configure a VRRP group on an interface of a specific LAN segment by setting the VRID and virtual IP address.
- Configure multiple VRRP groups on an interface to achieve load balancing and offer more stable and reliable network services.
- Configure the VRRP tracked interfaces to monitor real-time failures, change interface priorities and realize master-backup failover dynamically.

Notes

- To achieve VRRP, the routers in a VRRP group should be configured with the same virtual IPv4 address.
- To achieve mutual backup between multiple IPv4 VRRP groups, configure multiple IPv4 VRRP groups with identical VRRP configuration on different interface and configure different priorities for them so that they act as the master and backup groups mutually.
- Enable VRRP on Layer-3 interfaces.

Configuration Steps

Enabling IPv4 VRRP

- By default, IPv4 VRRP is disabled on an interface. You can enable it based on your demand.

Configuring the IPv4 VRRP Authentication String

- By default, VRRP is in non-authentication mode. You can enable plain text authentication mode based on your demand.

Configuring the IPv4 VRRP Advertisement Interval

- By default, the Master router sends advertisement packets every one second. You can modify the interval based on your demand.

Configuring the IPv4 VRRP Preemption Mode

- By default, a VRRP group operates in Preemption mode with a zero-second delay.

Configuring the IPv4 VRRP Router Priority

- The default router priority for a VRRP group is 100. You can modify the priority based on your demand.

Configuring the IPv4 VRRP Tracked Interface

- By default, an IPv4 VRRP group monitors no interface and the value of priority change is 10. To achieve fault monitoring through interface monitoring, please configure this item.

Configuring the IPv4 VRRP Learning Timer

- By default, the learning timer is disabled for a VRRP group. Enable this function if the Backup routers need to learn the Master's advertisement interval.

Configuring the IPv4 VRRP Group Description

- By default, no description is configured for a VRRP group. To distinguish VRRP groups clearly, configure descriptions.

Configuring the IPv4 VRRP Delay

- By default, the IPv6 VRRP delay is not configured. To guarantee an effective non-preemption mode, configure the delay.

Configuring the IPv4 VRRP Version

- By default, IPv4 adopts the VRRPv2 standard. To change it, use the corresponding command.

Specifying a Sub VLAN of a Super VLAN to Receive the IPv4 VRRP Packets

- By default, IPv4 VRRP packets are only sent to the first **UP** Sub VLAN interface of a Super VLAN, but you may configure a specific Sub VLAN.

Configuring the BFD Support for IPv4 VRRP on an Interface

- By default, the BFD support is not configured on an interface. To configure it, use the corresponding command.

Configuring Global IPv4 VRRP BFD

- By default, global IPv4 VRRP BFD is not enabled. To implement it, use the corresponding command.

Verification

- Run the **show vrrp** command to verify the configuration.

Related Commands

Enabling IPv4 VRRP

Command	**vrrp** *group* **ip** *ipaddress* [**secondary**]
Parameter Description	*group*: Indicates the VRID of a VRRP group, the range of which varies with product models. *ipaddress*: Indicates the IP address of a VRRP group. **secondary**: Indicates the secondary IP address of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	If no virtual IP address is specified, routers cannot join a VRRP group. If no secondary IP address is applied, the configured IP address will be the primary IP address of a VRRP group.

Configuring the IPv4 VRRP Authentication String

Command	**vrrp** *group* **authentication** *string*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *string*: Indicates the authentication string of a VRRP group (a plain text password consists of 8 bytes at most).
Command Mode	Interface configuration mode
Usage Guide	In a VRRP group, the same authentication password should be configured for routers. The plain text authentication password cannot guarantee security but only prevents/prompts wrong VRRP configurations. This command is only applicable to VRRPv2 instead of VRRPv3. Authentication is abolished for VRRPv3 (IPv4 VRRP and IPv6 VRRP) packets. If VRRPv2 is chosen for an IPv4 VRRP group, the command is effective; if VRRPv3 is chosen, the command is ineffective.

Configuring the IPv4 VRRP Advertisement Interval

Command	**vrrp** *group* **timers advertise** { *advertise-interval* | **csec** *centisecond-interval* }
Parameter Description	*group*: Indicates the VRID of a VRRP group. *advertise-interval*: Indicates the advertisement interval of a VRRP group (unit: second). ***csec*** *centisecond-interval:* An interval for a master router in a backup group to send VRRP packets. It is an integer from 50 to 99. The unit is centisecond. No default value is provided. The command is only effective for VRRPv3 packets. If it is configured for VRRPv2 packets, the default interval is one second.
Command Mode	Interface configuration mode
Usage Guide	If a router is elected as the Master in a VRRP group, it sends VRRP advertisement packets at the set interval to announce its VRRP state, priority and other information. According to the RFC standards, if an IPv4 VRRP group adopts VRRPv3 for sending multicast packets, the maximum advertisement interval is 40 seconds. Therefore, if the interval is set longer than 40 seconds, this maximum interval will be applied, though the configuration is effective.

Configuring the IPv4 VRRP Preemption Mode

Command	**vrrp** *group* **preempt** [**delay** *seconds*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. **delay** *seconds*: Indicates the preemption delay for the Master router to claim its status. The default value is 0 second.
Command Mode	Interface configuration mode
Usage Guide	If a VRRP group runs in Preemption mode, a higher priority router will take the place of the lower priority Master. If a VRRP group runs in Non-preemption mode, a router with the priority higher than that of the Master remains Backup. It makes little sense to configure the Preemption mode when the VRRP group uses the IP address of an Ethernet interface, in which case the group has the highest priority and automatically becomes the Master in the group.

Configuring the IPv4 VRRP Router Priority

Command	**vrrp** *group* **priority** *level*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *level*: Indicates the priority of an interface in a VRRP group.

Command Mode	Interface configuration mode
Usage Guide	This command is used to manually configure the VRRP router priority.

Configuring the IPv4 VRRP Tracked Interface

Command	**rrp group track {interface-type interface-number | bfd interface-type interface-number ipv4-addess } [priority]**
Parameter Description	*group*: Indicates the VRID of a VRRP group. *interface-type interface-number*: Indicates the interface to be tracked. *priority*: Indicates the scale of VRRP priority change when the state of a monitored interface changes. The default value is 10.
Command Mode	Interface configuration mode
Usage Guide	A tracked interface must be a routable Layer-3 logic interface (for example, a Routed port, an SVI interface, or a Loopback interface). The priority of the router owns the virtual IP address associated with a VRRP group must be 255, and no tracked interface can be configured on it.

Configuring the IPv4 VRRP Tracked IP Address

Command	**vrrp** *group* **track** *ipv4-address* [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. *ipv4-addres*: Indicates the IPv4 address to be tracked. ***interval*** *interval-value*: Indicates the probe interval. The unit is second. Unless configured manually, the value is 3 seconds by default. ***timeout*** *timeout-value*: Indicates the probe timeout of waiting for responses. If no response is received when the timeout is up, it is regarded that the destination is inaccessible. The unit is second. Unless configured manually, the value is 1 second by default. **retry** *retry-value*: Indicates the probe retries. If the probe packet is sent continually for retry-value times but no response is received, it is regarded that the destination is inaccessible. The unit is times. Unless configured, the value is 3 times by default. *priority*: Indicates the scale of VRRP priority change when the state of a monitored interface changes. The default value is 10.
Command Mode	Interface configuration mode
Usage Guide	To monitor a host, specify its IPv4 address for an IPv4 VRRP group. If a VRRP group owns the actual IP address of an Ethernet interface, the group priority is 255, and no monitored IP address can be configured.

Configuring the IPv4 VRRP Learning Timer

Command	**vrrp** *group* **timers learn**
Parameter Description	*group*: Indicates the VRID of a VRRP group.
Command Mode	Interface configuration mode

Usage Guide	Once the learning timer is enabled on a VRRP router, a Backup router learns the advertisement interval of the Master during the timer. Based on this, the Backup router calculates the interval for determining the Master router as failed instead of using the locally configured advertisement interval. This command achieves synchronization with the learning timer between the Master and Backup routers.

Configuring the IPv4 VRRP Group Description

Command	**vrrp** *group* **description** *text*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *text*: Indicates the description of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	A VRRP description helps distinguishing VRRP groups. A description has 80 bytes at most, otherwise wrong configuration is prompted.

Configuring the IPv4 VRRP Delay

Command	**vrrp delay** { **minimum** *min-seconds* | **reload** *reload-seconds* }
Parameter Description	**minimum** *min-seconds*: Indicates the VRRP delay after an interface state changes. **reload** *reload-seconds*: Indicates the VRRP delay after the system starts.
Command Mode	Interface configuration mode
Usage Guide	After the delay is configured for a VRRP group on an interface, the VRRP group starts after the delay instead of immediately upon system startup or the interface's resumption, ensuring non-preemption. If the interface receives a VRRP packet during the delay, the delay will be canceled and the VRRP will be started immediately. The two types of delay share a value range of 0 to 60 seconds. This configuration will be effective for both IPv4 and IPv6 VRRP groups of an interface.

Configuring the IPv4 VRRP Version

Command	**vrrp** *group* **version** { **2** | **3** }
Parameter Description	**2**: Indicates VRRPv2. **3**: Indicates VRRPv3.
Command Mode	Interface configuration mode
Usage Guide	Considering the compatibility between VRRPv2 and VRRPv3, specify a standard for IPv4 VRRP based on the actual network condition. VRRPv2 is developed in RFC3768, while VRRPv3 is described in RFC5798. This command is only applicable to IPv4 VRRP.

Specifying a Sub VLAN of a Super VLAN to Receive the IPv4 VRRP Packets

Command	**vrrp detection-vlan {first-subvlan |** *subvlan-id***}**
Parameter Description	**first-subvlan**: Sends IPv4 VRRP packets only to the first UP Sub VLAN interface in a Super VLAN. *subvlan-id*: Sends IPv4 VRRP packets to a specified Sub VLAN.
Command Mode	Interface configuration mode
Usage Guide	This command is used to specify a Sub VLAN of a Super VLAN to receive the IPv4 VRRP packets. IPv4 VRRP packets are

	sent in a Super VLAN using the following three methods. Packets are sent to the first UP Sub VLAN interface in a Super VLAN, or to a specified Sub VLAN interface in a Super VLAN, or to all the Sub VLAN interfaces in a Super VLAN. If VRRP and VRRP Plus are enabled simultaneously on a Super VLAN interface, VRRP packets are sent to all Up interfaces of the Sub VLANs under the Super VLAN. This command is configured on a VLAN interface and effective only to Super VLAN interfaces.

Configuring the BFD Support for IPv4 VRRP on an Interface

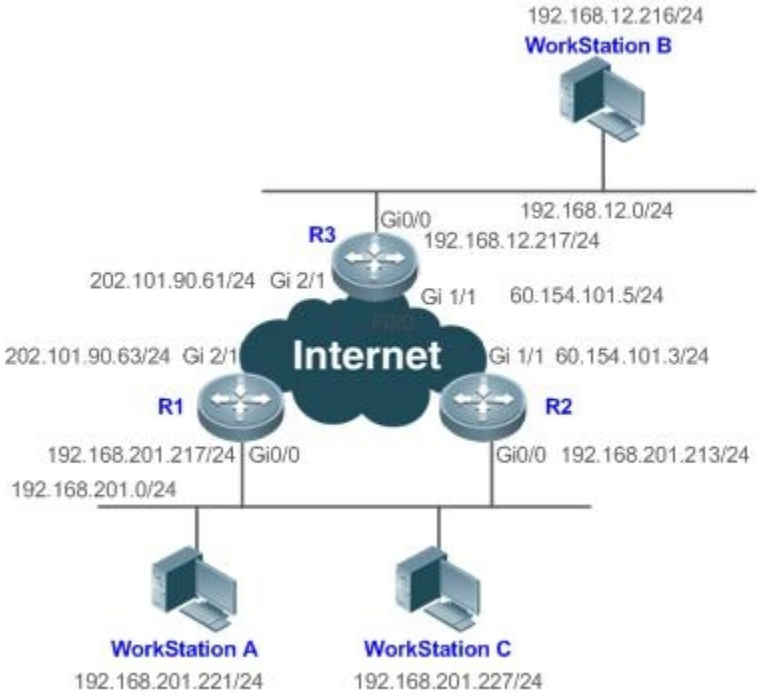
Command	**vrrp** *group* **bfd** *ip-address*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *ip-address*: Indicates the interface IP address.
Command Mode	Interface configuration mode
Usage Guide	For a Backup router, run this command to correlate an IPv4 VRRP group with BFD without caring the configured IP address. For the Master, as the primary IP address of a Backup router is not known, the router IP address can only be specified by the administrator. If global IPv4 VRRP BFD is configured, this configuration cannot be performed. To enable the BFD support, make sure that IP and BFD session parameters are configured on the target interface.

Configuring Global IPv4 VRRP BFD

Command	**vrrp bfd** *interface-type interface-number ip-address*
Parameter Description	*interface-type interface-number*: Indicates interface type and ID. *ip-address*: Indicates the interface IP address.
Command Mode	Global configuration mode
Usage Guide	If global IPv4 VRRP BFD is configured, the configured BFD support will be deleted. To enable the BFD support, make sure that IP and BFD session parameters are configured on the target interface. A global IPv4 VRRP BFD session is only applicable to an IPv4 VRRP group consisting of two routers.

Configuration Example

Configuring an IPv4 VRRP Group and Tracked Interface

Scenario **Figure 4- 4**	
Configuration Steps	● The cluster of Work Station A and Work Station B (192.168.201.0/24) uses the virtual IP address 192.168.201.1 of the VRRP group constituted by the routers R1 and R2 as the gateway address to communicate with Work Station B (192.168.12.0 /24). ● GigabitEthernet 2/1 on R1 is configured as the tracked interface. ● No VRRP but an ordinary routing function is configured on R3.
R3	R3#configure terminal R3(config)#interface GigabitEthernet 0/0 // The command "no switchport" is only required for a switch. R3(config-if-GigabitEthernet 0/0)#no switchport R3(config-if-GigabitEthernet 0/0)#ip address 192.168.12.217 255.255.255.0 R3(config-if-GigabitEthernet 0/0)#exit R3(config)#interface GigabitEthernet 1/1 // The command "no switchport" is only required for a switch. R3(config-if-GigabitEthernet 1/1)#no switchport R3(config-if-GigabitEthernet 1/1)#ip address 60.154.101.5 255.255.255.0 R3(config-if-GigabitEthernet 1/1)#exit R3(config)#interface GigabitEthernet 2/1 // The command "no switchport" is only required for a switch. R3(config-if-GigabitEthernet 2/1)#no switchport R3(config-if-GigabitEthernet 2/1)#ip address 202.101.90.61 255.255.255.0 R3(config-if-GigabitEthernet 2/1)#exit

	R3(config)#router ospf R3(config-router)#network 202.101.90.0 0.0.0.255 area 10 R3(config-router)#network 192.168.12.0 0.0.0.255 area 10 R3(config-router)#network 60.154.101.0 0.0.0.255 area 10
R1	R1#configure terminal R1(config)#interface GigabitEthernet 0/0 R1(config-if-GigabitEthernet 0/0)#ip address 192.168.201.217 255.255.255.0 R1(config-if-GigabitEthernet 0/0)#vrrp 1 priority 120 R1(config-if-GigabitEthernet 0/0)#vrrp 1 version 3 R1(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3 R1(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1 R1(config-if-GigabitEthernet 0/0)#vrrp 1 track GigabitEthernet 2/1 30 R1(config-if-GigabitEthernet 0/0)#exit R1(config)#interface GigabitEthernet 2/1 R1(config-if-GigabitEthernet 2/1)#ip address 202.101.90.63 255.255.255.0 R1(config-if-GigabitEthernet 2/1)#exit R1(config)#router ospf R1(config-router)#network 202.101.90.0 0.0.0.255 area 10 R1(config-router)#network 192.168.201.0 0.0.0.255 area 10
R2	R2#configure terminal R2(config)#interface GigabitEthernet 0/0 R2(config-if-GigabitEthernet 0/0)#ip address 192.168.201.213 255.255.255.0 R2(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1 R2(config-if-GigabitEthernet 0/0)#vrrp 1 version 3 R2(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3 R2(config-if-GigabitEthernet 0/0)#exit R2(config)#interface GigabitEthernet 1/1 // The command “no switchport” is only required for a switch. R2(config-if-GigabitEthernet 1/1)#no switchport R2(config-if-GigabitEthernet 1/1)#ip address 60.154.101.3 255.255.255.0 R2(config-if-GigabitEthernet 1/1)#exit R2(config)#router ospf R2(config-router)#network 60.154.101.0 0.0.0.255 area 10

	R2(config-router)#network 192.168.201.0 0.0.0.255 area 10
Verification	Run the **show vrrp** command to verify the configuration. ● Check whether R1, which acts as the Master, reduces its VRRP priority from 120 to 90 when GigabitEthernet2/1 connected to the Wide Area Network (WAN) is unavailable. If yes, R2 becomes the Master. ● Check whether R1 resumes its VRRP priority from 30 to 120 when GigabitEthernet 2/1 connected to the WAN recovers. If yes, R1 is re-elected as the Master.
R1	R1#show vrrp GigabitEthernet 0/0 - Group 1 State is Master Virtual IP address is 192.168.201.1 configured Virtual MAC address is 0000.5e00.0101 Advertisement interval is 3 sec Preemption is enabled min delay is 0 sec Priority is 120 Master Router is 192.168.201.217 (local), priority is 120 Master Down interval is 10.59 sec Tracking state of 1 interface, 1 up: up GigabitEthernet 2/1 priority decrement=30
R2	R2#show vrrp GigabitEthernet 0/0 - Group 1 State is Backup Virtual IP address is 192.168.201.1 configured Virtual MAC address is 0000.5e00.0101 Advertisement interval is 3 sec Preemption is enabled min delay is 0 sec Priority is 100 Master Router is 192.168.201.217 , priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.82 sec

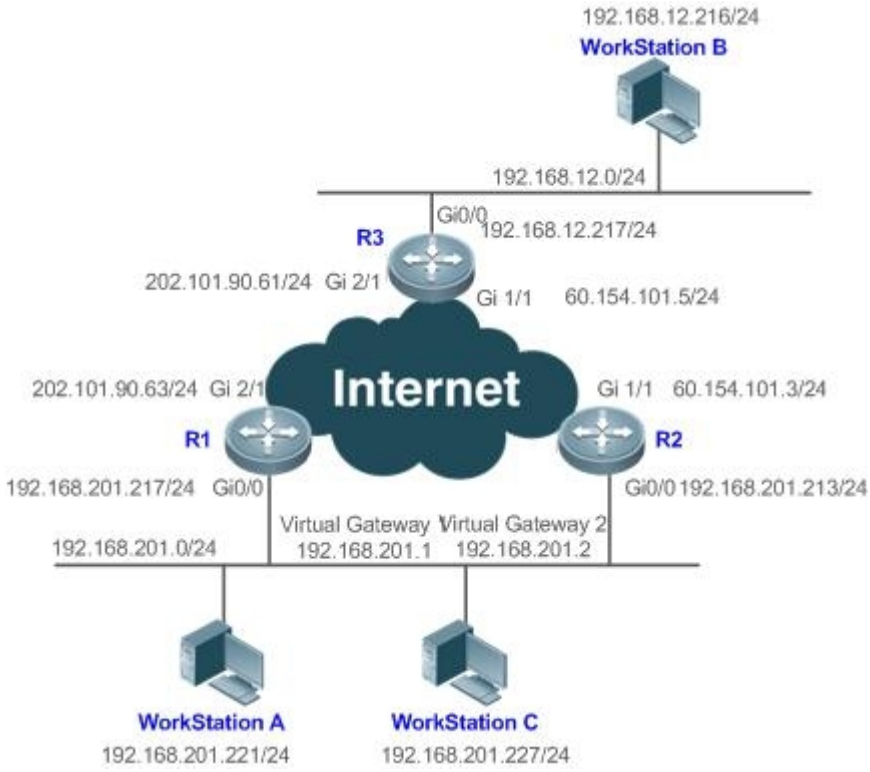
Common Errors

- Different virtual IP addresses are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.

- Different VRRP advertisement intervals are configured on the routers in a VRRP group and the learning timer is not configured, resulting in multiple Master routers in the group.
- Different VRRP versions are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- For VRRPv2, the Ethernet interfaces of the routers in a VRRP group are all in plain text authentication mode but inconsistent in authentication strings, resulting in multiple Master routers in the group.

Configuration Example

Configuring Multiple IPv4 VRRP Groups

Scenario **Figure 4- 5**	
Configuration Steps	• The user workstation cluster (192.168.201.0/24) uses the backup group constituted by the routers R1 and R2. The gateway for partial workstations (A for example) points to the virtual IP address 192.168.201.1 of the backup group 1, while that for other partial workstations (C for example) points to the virtual IP address 192.168.201.2 of the backup group 2. IPv4 multicast routing is enabled on all the routers. • R1 acts as the master router in the group 2 and as a backup router in the group 1. • R2 acts as a backup router in the group 2 and as a master router in the group 1.
R3	R3#configure terminal R3(config)#interface GigabitEthernet 0/0 // The command "no switchport" is only required for a switch. R3(config-if-GigabitEthernet 0/0)#no switchport R3(config-if-GigabitEthernet 0/0)#ip address 192.168.12.217 255.255.255.0 R3(config-if-GigabitEthernet 0/0)#exit

	R3(config)#interface GigabitEthernet 1/1 // The command "no switchport" is only required for a switch. R3(config-if-GigabitEthernet 1/1)#no switchport R3(config-if-GigabitEthernet 1/1)#ip address 60.154.101.5 255.255.255.0 R3(config-if-GigabitEthernet 1/1)#exit R3(config)#interface GigabitEthernet 2/1 // The command "no switchport" is only required for a switch. R3(config-if-GigabitEthernet 2/1)#no switchport R3(config-if-GigabitEthernet 2/1)#ip address 202.101.90.61 255.255.255.0 R3(config-if-GigabitEthernet 2/1)#exit R3(config)#router ospf R3(config-router)#network 202.101.90.0 0.0.0.255 area 10 R3(config-router)#network 192.168.12.0 0.0.0.255 area 10 R3(config-router)#network 60.154.101.0 0.0.0.255 area 10
R1	R1#configure terminal R1(config)#interface GigabitEthernet 0/0 R1(config-if-GigabitEthernet 0/0)#ip address 192.168.201.217 255.255.255.0 R1(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3 R1(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1 R1(config-if-GigabitEthernet 0/0)#vrrp 2 priority 120 R1(config-if-GigabitEthernet 0/0)#vrrp 2 timers advertise 3 R1(config-if-GigabitEthernet 0/0)#vrrp 2 ip 192.168.201.2 R1(config-if-GigabitEthernet 0/0)#vrrp 2 track GigabitEthernet 2/1 30 R1(config-if-GigabitEthernet 0/0)#exit R1(config)#interface GigabitEthernet 2/1 R1(config-if-GigabitEthernet 2/1)#ip address 202.101.90.63 255.255.255.0 R1(config-if-GigabitEthernet 2/1)#exit R1(config)#router ospf R1(config-router)#network 202.101.90.0 0.0.0.255 area 10 R1(config-router)#network 192.168.201.0 0.0.0.255 area 10
R2	R2#configure terminal R2(config)#interface GigabitEthernet 0/0 R2(config-if-GigabitEthernet 0/0)#ip address 192.168.201.213 255.255.255.0

	R2(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1 R2(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3 R2(config-if-GigabitEthernet 0/0)#vrrp 1 priority 120 R2(config-if-GigabitEthernet 0/0)#vrrp 2 ip 192.168.201.2 R2(config-if-GigabitEthernet 0/0)#vrrp 2 timers advertise 3 R2(config-if-GigabitEthernet 0/0)#exit R2(config)#interface GigabitEthernet 1/1 R2(config-if-GigabitEthernet 1/1)#ip address 60.154.101.3 255.255.255.0 R2(config-if-GigabitEthernet 1/1)#exit R2(config)#router ospf R2(config-router)#network 60.154.101.0 0.0.0.255 area 10 R2(config-router)#network 192.168.201.0 0.0.0.255 area 10
Verification	Run the **show vrrp** command to verify the configuration. ● Check whether R1, which acts as a master router in the group 2, reduces its VRRP group priority from 30 to 90 when it finds that the interface GigabitEthernet 2/1 connected to a WAN is unavailable. If yes, R2 in the group 2 becomes a master router. ● Check whether R1 increases its VRRP group priority from 30 to 120 when it finds the interface GigabitEthernet 2/1 connected to a WAN becomes available again. If yes, R1 becomes a master router again in the group 2.
R1	R1#show vrrp GigabitEthernet 0/0 - Group 1 State is Backup Virtual IP address is 192.168.201.1 configured Virtual MAC address is 0000.5e00.0101 Advertisement interval is 3 sec Preemption is enabled min delay is 0 sec Priority is 100 Master Router is 192.168.201.213 , priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.82 sec GigabitEthernet 0/0 - Group 2 State is Master Virtual IP address is 192.168.201.2 configured Virtual MAC address is 0000.5e00.0102

	Advertisement interval is 3 sec Preemption is enabled min delay is 0 sec Priority is 120 Master Router is 192.168.201.217 (local), priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.59 sec Tracking state of 1 interface, 1 up: up GigabitEthernet 2/1 priority decrement=30
R2	R2#show vrrp GigabitEthernet 0/0 - Group 1 State is Master Virtual IP address is 192.168.201.1 configured Virtual MAC address is 0000.5e00.0101 Advertisement interval is 3 sec Preemption is enabled min delay is 0 sec Priority is 120 Master Router is 192.168.201.213 (local), priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.59 sec GigabitEthernet 0/0 - Group 2 State is Backup Virtual IP address is 192.168.201.2 configured Virtual MAC address is 0000.5e00.0102 Advertisement interval is 3 sec Preemption is enabled min delay is 0 sec Priority is 100 Master Router is 192.168.201.217 , priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.82 sec

Common Errors

- Different virtual IP addresses are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- Different VRRP advertisement intervals are configured on the routers in a VRRP group and the learning timer is not configured, resulting in multiple Master routers in the group.
- Different VRRP versions are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- For VRRPv2, the Ethernet interfaces of the routers in a VRRP group are all in plain text authentication mode but inconsistent in authentication strings, resulting in multiple Master routers in the group.

4.4.2 Configuring IPv6 VRRP

Configuration Effect

- Configure an IPv6 VRRP group on an interface of a specific LAN segment by setting the VRID and virtual IPv6 address.
- Configure multiple IPv6 VRRP groups on an interface to achieve load balance and achieve more stable and reliable network services.
- Configure the VRRP tracked interfaces to monitor real-time failures, change interface priorities and realize master-backup failover dynamically.

Notes

- To achieve VRRP, the routers in a VRRP group should be configured with the same virtual IPv6 address.
- To achieve mutual backup for multiple IPv6 VRRP backup groups, you need to configure multiple IPv6 VRRP groups with identical VRRP configuration on an interface and configure different priorities for them to make routers master and backup mutually.
- VRRP must be enabled on Layer-3 interfaces.

Configuration Steps

Enabling IPv6 VRRP in Interface Configuration Mode

- By default, IPv6 VRRP is not enabled on an interface. You can enable it based on your demand.

Configuring the IPv6 VRRP Advertisement Interval

- By default, the Master router sends advertisement packets every one second. You can modify the interval based on your demand.

Configuring the IPv6 VRRP Preemption Mode

- By default, a VRRP group operates in Preemption mode with a zero-second delay.

Enabling the Accept Mode for an IPv6 VRRP Group

- By default, the Accept mode is disabled for an IPv6 VRRP group. To require an IPv6 VRRP VRRP group in Master state to receive and process packets with the destination IP address as that of the virtual router, enable Accept mode.

Configuring the IPv6 VRRP Router Priority

- The default router priority for a VRRP group is 100. You can modify the priority based on your demand.

Configuring the IPv6 VRRP Tracked Interface

- By default, no tracked interface is configured. You can modify the interval based on your demand.

Configuring the IPv6 VRRP Tracked IP Address

- By default, no tracked IPv6 address is configured and the value of priority change is 10. You can configure this function based on your demand.

Configures the IPv6 VRRP Learning Timer

- By default, the learning timer is disabled for a VRRP group. Enable this function if the Backup routers need to learn the Master's advertisement interval.

Configuring the IPv6 VRRP Group Description

- By default, no description is configured for a VRRP group. To distinguish VRRP groups clearly, configure descriptions.

Configuring the IPv4 VRRP Delay

- By default, the IPv6 VRRP delay is not configured. To guarantee an effective non-preemption mode, configure the delay.

Verification

- Run the **show ipv6 vrrp** command to verify the configuration.

Related Commands

Enabling IPv6 VRRP

Command	**vrrp** *group* **ipv6** *ipv6-address*
Parameter Description	*group*: Indicates the VRID of a VRRP group, the range of which varies with product models. *ipv6-address*: Indicates the IPv6 address of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	IPv6 VRRP groups and IPv4 VRRP groups share a VRID range from 1 to 255. One VRID is applicable to an IPv4 VRRP group and an IPv6 VRRP group at the same time. The first configured address should be a link-local address, which can be deleted only after other virtual addresses.

Configuring the IPv6 VRRP Advertisement Interval

Command	**vrrp ipv6** *group* **timers advertise** { *advertise-interval* | **csec** *centisecond-interval* }
Parameter Description	*group*: Indicates the VRID of a VRRP group. *advertise-interval*: Indicates the advertisement interval of a VRRP group (unit: second). ***csec*** *centisecond-interval:* An interval for a master router in a backup group to send VRRP packets. It is an integer from 50 to 99. The unit is centisecond. No default value is provided. The command is only effective for VRRPv3 packets. If it is configured for VRRPv2 packets, the default interval is one second.
Command Mode	Interface configuration mode
Usage Guide	If a router is elected as the Master in a VRRP group, it sends VRRP advertisement packets at the set interval to announce its VRRP state, priority and other information. According to the RFC standards, if an IPv6 VRRP group adopts VRRPv3 for sending multicast packets, the maximum advertisement interval is 40 seconds. Therefore, if the interval is set longer than 40 seconds, this maximum interval

	will be applied, though the configuration is effective.

Configuring the Preemption Mode

Command	**vrrp ipv6** *group* **preempt** [**delay** *seconds*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. **delay** *seconds*: Indicates the preemption delay for the Master router to claim its status. The default value is 0 second.
Command Mode	Interface configuration mode
Usage Guide	If a VRRP group runs in Preemption mode, a higher priority router will take the place of the lower priority Master. If a VRRP group runs in Non-preemption mode, a router with the priority higher than that of the Master remains Backup. It makes little sense to configure the Preemption mode when the VRRP group uses the IP address of an Ethernet interface, in which case the group has the highest priority and automatically becomes the Master in the group.

Enabling the Accept Mode for an IPv6 VRRP Group

Command	**vrrp ipv6** *group* **accept_mode**
Parameter Description	*group*: Indicates the VRID of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	By default, an IPv6 VRRP group in **Master** state is not permitted to receive packets with the destination IPv6 address as that of the VRRP group. However, it receives NA and NS packets no matter whether Accept mode is configured. Besides, the IP Address Owner in Master state receives and processes the packets with the destination IPv6 address as that of the VRRP group no matter whether Accept mode is configured or not.

Configuring the IPv6 VRRP Router Priority

Command	**vrrp ipv6** *group* **priority** *level*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *level*: Indicates the priority of a VRRP router.
Command Mode	Interface configuration mode
Usage Guide	This command is used to manually configure the VRRP router priority.

Configuring the IPv6 VRRP Tracked Interface

Command	**vrrp ipv6** *group* **track** *interface-type interface-number* [*priority*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. *interface-type interface-number*: Indicates the interface to be tracked. *priority*: Indicates the scale of VRRP priority change when the state of a monitored interface changes. The default value is 10.
Command Mode	Interface configuration mode
Usage Guide	A tracked interface must be a routable Layer-3 logic interface (for example, a Routed port, an SVI interface, a Loopback interface, or a Tunnel interface). The priority of the router owns the virtual IP address associated with a VRRP group must be 255, and no tracked

	interface can be configured on it.

Configuring the IPv6 VRRP Tracked IP Address

Command	**vrrp ipv6** *group* **track** { *ipv6-global-address* | *ipv6-linklocal-address interface-type interface-number* } [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. *ipv6-global-addres*: Indicates the IPv6 global unicast address. *ipv6-linklocal-address*: Indicates the IPv6 link-local address. *interface-type interface-number*: Indicates the interface to be tracked. ***interval*** *interval-value*: Indicates the probe interval. The unit is second. Unless configured manually, the value is 3 seconds by default. ***timeout*** *timeout-value*: Indicates the probe timeout of waiting for responses. If no response is received when the timeout is up, it is regarded that the destination is inaccessible. The unit is second. Unless configured manually, the value is 1 second by default. **retry** *retry-value*: Indicates the probe retries. If the probe packet is sent continually for *retry-value* times but no response is received, it is regarded that the destination is inaccessible. The unit is times. Unless configured, the value is 3 times by default. *priority*: Indicates the scale of VRRP priority change when the state of a monitored interface changes. The default value is 10.
Command Mode	Interface configuration mode
Usage Guide	To monitor a host, specify its IPv6 address for an IPv6 VRRP group. If the host IP address being tracked is a link-local address, specify a network interface. If a VRRP group owns the actual IP address of an Ethernet interface, the group priority is 255, and no monitored IP address can be configured.

Configures the IPv6 VRRP Learning Timer

Command	**vrrp ipv6** *group* **timers learn**
Parameter Description	*group*: Indicates the VRID of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	Once the learning timer is enabled on a VRRP router, a Backup router learns the advertisement interval of the Master during the timer. Based on this, the Backup router calculates the interval for determining the Master router as failed instead of using the locally configured advertisement interval. This command achieves synchronization with the learning timer between the Master and Backup routers.

Configuring the IPv6 VRRP Group Description

Command	**vrrp ipv6** *group* **description** *text*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *text*: Indicates the description of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	A VRRP description helps distinguishing VRRP groups. A description has 80 bytes at most, otherwise wrong

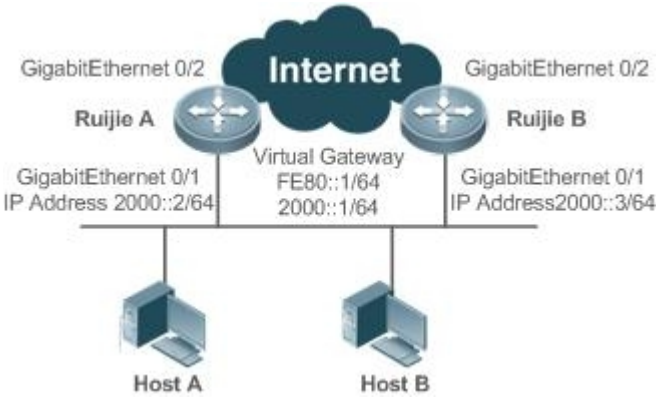
	configuration is prompted.

Configuring the IPv4 VRRP Delay

Command	**vrrp delay** { **minimum** *min-seconds* | **reload** *reload-seconds* }
Parameter Description	**minimum** *min-seconds*: Indicates the VRRP delay after an interface state changes. **reload** *reload-seconds*: Indicates the VRRP delay after the system starts.
Command Mode	Interface configuration mode
Usage Guide	After the delay is configured for a VRRP group on an interface, the VRRP group starts after the delay instead of immediately upon system startup or the interface's resumption, ensuring non-preemption. If the interface receives a VRRP packet during the delay, the delay will be canceled and the VRRP will be started immediately. The two types of delay share a value range of 0 to 60 seconds. This configuration will be effective for both IPv4 and IPv6 VRRP groups of an interface.

Configuration Example

Configuring an IPv6 VRRP Group and Tracked Interface

Scenario **Figure 4- 6**	
Configuration Steps	● Host A and Host B access the Internet resources through the default gateway 2000::1/64. ● FS A and FS B belong to the IPv6 VRRP group 1, and their virtual addresses are 2000::1/64 and FE80::1 respectively. ● FS A tracks the interface GigabitEthernet 0/2 connected to the Internet. When GigabitEthernet 0/2 is unavailable, FS A reduces its priority and FS B acts as a gateway.
FSA	FSA#configure terminal FSA(config)#interface GigabitEthernet 0/1 FSA(config-if-GigabitEthernet 0/1)#no switchport FSA(config-if-GigabitEthernet 0/1)#ipv6 address 2000::2/64 FSA(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 FE80::1 FSA(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 2000::1 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 priority 120 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 timers advertise 3

	FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 track GigabitEthernet 0/2 50 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 accept_mode
FSB	FSB#configure terminal FSB(config)#interface GigabitEthernet 0/1 FSB(config-if-GigabitEthernet 0/1)#no switchport FSB(config-if-GigabitEthernet 0/1)#ipv6 address 2000::3/64 FSB(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 FE80::1 FSB(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 2000::1 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 priority 100 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 timers advertise 3 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 accept_mode
Verification	Run the **show vrrp** command to verify the configuration. ● Check whether FS A, which acts as the Master router, reduces its VRRP group priority from 120 to 70 when it finds that the interface GigabitEthernet 0/2 connected to WAN is unavailable. If yes, FS B becomes the Master. ● Check whether FS A increases its VRRP group priority from 50 to 120 when it finds the interface GigabitEthernet 0/2 connected to WAN becomes available again. If yes, FS A becomes the Master again.
FSA	FSA#show ipv6 vrrp 1 GigabitEthernet 0/1 - Group 1 State is Master Virtual IPv6 address is as follows: FE80::1 2000::1 Virtual MAC address is 0000.5e00.0201 Advertisement interval is 3 sec Accept_Mode is enabled Preemption is enabled min delay is 0 sec Priority is 120 Master Router is FE80::1234 (local), priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.59 sec Tracking state of 1 interface, 1 up: up GigabitEthernet 0/2 priority decrement=50

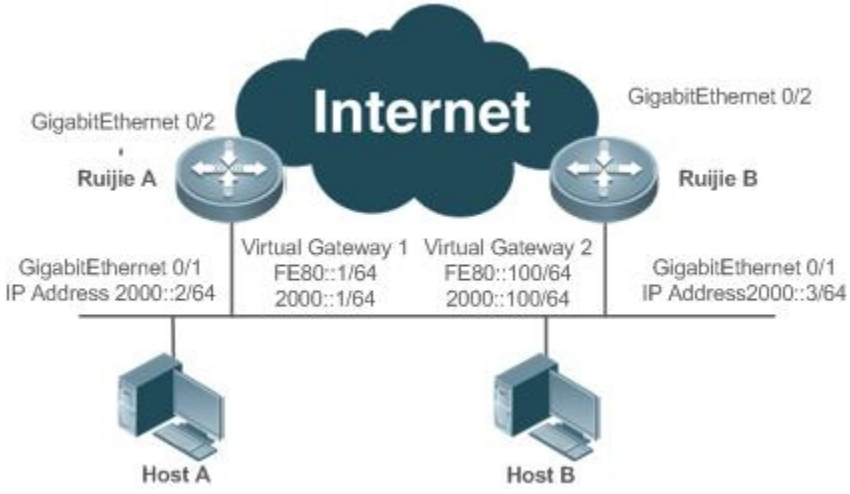
FSB	FSB#show ipv6 vrrp 1 GigabitEthernet 0/1 - Group 1 State is Backup Virtual IPv6 address is as follow: FE80::1 2000::1 Virtual MAC address is 0000.5e00.0201 Advertisement interval is 3 sec Accept_Mode is enabled Preemption is enabled min delay is 0 sec Priority is 100 Master Router is FE80::1234, priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.82 sec

Common Errors

- Different virtual IPv6 addresses are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- Different VRRP advertisement intervals are configured on the routers in a VRRP group and the learning timer is not configured, resulting in multiple Master routers in the group.

Configuration Example

Multiple VRRP Backup Groups (under IPv6)

Scenario **Figure 4- 7**	
Configuration Steps	● Host A and Host B access the Internet resources through the gateways 2000::1/64 and 2000::100/64 respectively. ● FS A and FS B belong to the IPv6 VRRP group 1, and their virtual addresses are 2000::1/64 and FE80::1

	respectively. ● FS A and FS B belong to the backup group 2 of a virtual IPv6 router, and their virtual addresses are 2000::100/64 and FE80::100 respectively. ● FS A and FS B act as gateways and forward flows, being a backup router to each other.
FSA	FSA#configure terminal FSA(config)#interface GigabitEthernet 0/1 FSA(config-if-GigabitEthernet 0/1)#no switchport FSA(config-if-GigabitEthernet 0/1)#ipv6 address 2000::2/64 FSA(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 FE80::1 FSA(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 2000::1 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 priority 120 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 timers advertise 3 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 accept_mode FSA(config-if-GigabitEthernet 0/1)#vrrp 2 ipv6 FE80::100 FSA(config-if-GigabitEthernet 0/1)# vrrp 2 ipv6 2000::100 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 2 priority 100 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 2 timers advertise 3 FSA(config-if-GigabitEthernet 0/1)#vrrp ipv6 2 accept_mode
FSB	FSB#configure terminal FSB(config)#interface GigabitEthernet 0/1 FSB(config-if-GigabitEthernet 0/1)#no switchport FSB(config-if-GigabitEthernet 0/1)#ipv6 address 2000::3/64 FSB(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 FE80::1 FSB(config-if-GigabitEthernet 0/1)#vrrp 1 ipv6 2000::1 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 priority 100 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 timers advertise 3 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 1 accept_mode FSB(config-if-GigabitEthernet 0/1)#vrrp 2 ipv6 FE80::100 FSB(config-if-GigabitEthernet 0/1)# vrrp 2 ipv6 2000::100 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 2 priority 120 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 2 timers advertise 3 FSB(config-if-GigabitEthernet 0/1)#vrrp ipv6 2 accept_mode
Verification	Run the **show vrrp** command to verify the configuration.

FSA	FSA#show ipv6 vrrp GigabitEthernet 0/1 - Group 1 State is Master Virtual IPv6 address is as follows: FE80::1 2000::1 Virtual MAC address is 0000.5e00.0201 Advertisement interval is 3 sec Accept_Mode is enabled Preemption is enabled min delay is 0 sec Priority is 120 Master Router is FE80::1234 (local), priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.59 sec GigabitEthernet 0/1 - Group 2 State is Backup Virtual IPv6 address is as follows: FE80::100 2000::100 Virtual MAC address is 0000.5e00.0202 Advertisement interval is 3 sec Accept_Mode is enabled Preemption is enabled min delay is 0 sec Priority is 100 Master Router is FE80::5678, priority is 120 Master Advertisement interval is 3 sec Master Down interval is 10.82 sec
FSB	FSB#show ipv6 vrrp GigabitEthernet 0/1 - Group 1 State is Backup Virtual IPv6 address is as follow:

```
    FE80::1
    2000::1
  Virtual MAC address is 0000.5e00.0201
  Advertisement interval is 3 sec
  Accept_Mode is enabled
  Preemption is enabled
    min delay is 0 sec
  Priority is 100
  Master Router is FE80::1234, priority is 120
  Master Advertisement interval is 3 sec
  Master Down interval is 10.82 sec
GigabitEthernet 0/1 - Group 2
  State is Master
  Virtual IPv6 address is as follows:
    FE80::100
    2000::100
  Virtual MAC address is 0000.5e00.0202
  Advertisement interval is 3 sec
  Accept_Mode is enabled
  Preemption is enabled
    min delay is 0 sec
  Priority is 120
  Master Router is FE80::5678(local), priority is 120
  Master Advertisement interval is 3 sec
  Master Down interval is 10.59 sec
```

Common Errors

- Different virtual IPv6 addresses are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- Different VRRP advertisement intervals are configured on the routers in a VRRP group and the learning timer is not configured, resulting in multiple Master routers in the group.

4.4.3 Configuring VRRP-MSTP

Configuration Effect

- Link-level and gateway-level backup are achieved and network robustness is improved greatly when MTSP and VRRP are applied

simultaneously.

Notes

- configure the routers in a VRRP backup group with the same virtual IPv4 address.
- Enabled VRRP on a Layer 3 interface.

Configuration Steps

Enabling IPv4 VRRP

- By default, IPv4 VRRP is not enabled on an interface. To enable IPv4 VRRP, please configure this item.

Configuring the IPv4 VRRP Authentication String

- By default, VRRP is in a non-authentication mode. To enable plain text password authentication for VRRP, please configure this item.

Configuring the IPv4 VRRP Advertisement Interval

- By default, a master router sends VRRP GWADV packets at an interface of one second. To manually set a value, please configure this item.

Configuring the IPv4 VRRP Preemption Mode

- By default, VRRP groups work in the preemption mode with zero-second delay.

Configuring the IPv4 VRRP Router Priority

- The default router priority for a VRRP group is 100. You can modify the priority based on your demand.

Configuring the IPv4 VRRP Tracked Interface

- By default, an IPv4 VRRP group monitors no interface. To achieve fault monitoring through monitoring an interface, please configure this item.

Configuring the IPv4 VRRP Learning Timer

- By default, timed learning is not enabled for a VRRP backup group. To enable backup routers to learn the VRRP GWADV packets from a master router, please configure this item.

Configuring the IPv4 VRRP Group Description

- By default, no description is configured for a VRRP group. To distinguish VRRP groups conveniently, please configure this item.

Configuring the IPv4 VRRP Delay

- By default, the VRRP delay for a VRRP group is not configured. Configure the delay to guarantee a stable transition from Non-preemption mode to Preemption mode.

Configuring the IPv4 VRRP Version

- By default, the VRRPv2 standard is adopted for IPv4 VRRP packets. To modify it manually, please configure this item.

Specifying a Sub VLAN of a Super VLAN to Receive the IPv4 VRRP Packets

- By default, IPv4 VRRP packets are only sent to the first UP Sub VLAN interface in a Super VLAN, but you may configure a specific Sub

VLAN interface to send such packets.

Configuring the BFD Support for IPv4 VRRP on an Interface

- By default, the linkage between an IPv4 VRRP and BFD is not configured on an interface. To enable such linkage, please configure this item.

Configuring Global IPv4 VRRP BFD

- By default, global IPv4 VRRP BFD is not used to detect whether a master router is active. To enable this, please configure this item.

Verification

- Run the **show vrrp** command to verify the configuration.

Related Commands

Enabling IPv4 VRRP

Command	**vrrp** *group* **ip** *ipaddress* [**secondary**]
Parameter Description	*group*: Indicates the VRID of a VRRP group, the range of which varies with product models. *Ipaddress*: The IP address of a VRRP group. **secondary**: Indicates the secondary IP address of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	If no virtual IP address is specified, routers cannot join a VRRP group. If no secondary IP address is applied, the configured IP address will be the primary IP address of a VRRP group.

Configuring the IPv4 VRRP Authentication String

Command	**vrrp** *group* **authentication** *string*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *string*: Indicates the authentication string of a VRRP group (a plain text password consists of 8 bytes at most).
Command Mode	Interface configuration mode
Usage Guide	In a VRRP group, the same authentication password should be configured for routers. The plain text authentication password cannot guarantee security but only prevents/prompts wrong VRRP configurations. This command is only applicable to VRRPv2 instead of VRRPv3. Authentication is abolished for VRRPv3 packets. If VRRPv2 is chosen for an IPv4 VRRP group, the command is effective; if VRRPv3 is chosen, the command is ineffective.

Configuring the IPv4 VRRP Advertisement Interval

Command	**vrrp** *group* **timers advertise** { *advertise-interval* | **csec** *centisecond-interval* }
Parameter Description	*group*: Indicates the VRID of a VRRP group. *advertise-interval*: Indicates the advertisement interval of a VRRP group (unit: second). ***csec*** *centisecond-interval:* An interval for a master router in a backup group to send VRRP packets. It is an integer from 50 to 99. The unit is centisecond. No default value is provided. The command is only effective for VRRPv3 packets. If it is configured for VRRPv2 packets, the default interval is one second.

Command Mode	Interface configuration mode
Usage Guide	If a router is elected as the Master in a VRRP group, it sends VRRP advertisement packets at the set interval to announce its VRRP state, priority and other information. According to the RFC standards, if an IPv4 VRRP group adopts VRRPv3 for sending multicast packets, the maximum advertisement interval is 40 seconds. Therefore, if the interval is set longer than 40 seconds, this maximum interval will be applied, though the configuration is effective.

Configuring the IPv4 VRRP Preemption Mode

Command	**vrrp** *group* **preempt** [**delay** *seconds*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. **delay** *seconds*: Indicates the preemption delay for the Master router to claim its status. The default value is 0 second.
Command Mode	Interface configuration mode
Usage Guide	If a VRRP group runs in Preemption mode, a higher priority router will take the place of the lower priority Master. If a VRRP group runs in Non-preemption mode, a router with the priority higher than that of the Master remains Backup. It makes little sense to configure the Preemption mode when the VRRP group uses the IP address of an Ethernet interface, in which case the group has the highest priority and automatically becomes the Master in the group.

Configuring the IPv4 VRRP Router Priority

Command	**vrrp** *group* **priority** *level*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *level*: Indicates the priority of an interface in a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	This command is used to manually configure the priority of a VRRP group.

Configuring the IPv4 VRRP Tracked Interface

Command	**vrrp** *group* **track** { *interface-type interface-number* | **bfd** *interface-type interface-number ipv4-addess* } [*priority*]
Parameter Description	*group*: Indicates the VRID of a VRRP group. *interface-type interface-number*: Indicates the interface to be tracked. **bfd** *interface-type interface-number ipv4-address*: A specified adjacent IP address tracked through BFD. *priority*: Indicates the scale of VRRP priority change when the state of a monitored interface changes. The default value is 10.
Command Mode	Interface configuration mode
Usage Guide	A tracked interface must be a routable Layer-3 logic interface (for example, a Routed port, an SVI interface, a Loopback interface, or a Tunnel interface). The priority of the router owns the virtual IP address associated with a VRRP group must be 255, and no tracked interface can be configured on it.

Configuring the IPv4 VRRP Tracked IP Address

Command	**vrrp** *group* **track** *ipv4-address* [**interval** *interval-value*] [**timeout** *timeout-value*] [**retry** *retry-value*] [*priority*]

Parameter Description	*group*: Indicates the VRID of a VRRP group. *ipv4-addres*: Indicates the IPv4 address to be tracked. ***interval*** *interval-value*: Indicates the probe interval. The unit is second. Unless configured manually, the value is 3 seconds by default. ***timeout*** *timeout-value*: Indicates the probe timeout of waiting for responses. If no response is received when the timeout is up, it is regarded that the destination is inaccessible. The unit is second. Unless configured manually, the value is 1 second by default. **retry** *retry-value*: Indicates the probe retries. If the probe packet is sent continually for retry-value times but no response is received, it is regarded that the destination is inaccessible. The unit is times. Unless configured, the value is 3 times by default. *priority*: Indicates the scale of VRRP priority change when the state of a monitored interface changes. The default value is 10.
Command Mode	Interface configuration mode
Usage Guide	To monitor a host, specify its IPv4 address for an IPv4 VRRP group. If a VRRP group owns the actual IP address of an Ethernet interface, the group priority is 255, and no monitored IP address can be configured.

Configuring the IPv4 VRRP Learning Timer

Command	**vrrp** *group* **timers learn**
Parameter Description	*group*: Indicates the VRID of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	Once the learning timer is enabled on a VRRP router, a Backup router learns the advertisement interval of the Master during the timer. Based on this, the Backup router calculates the interval for determining the Master router as failed instead of using the locally configured advertisement interval. This command achieves synchronization with the learning timer between the Master and Backup routers.

Configuring the IPv4 VRRP Group Description

Command	**vrrp** *group* **description** *text*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *text*: Indicates the description of a VRRP group.
Command Mode	Interface configuration mode
Usage Guide	A VRRP description helps distinguishing VRRP groups. A description has 80 bytes at most, otherwise wrong configuration is prompted.

Configuring the IPv4 VRRP Delay

Command	**vrrp delay** { **minimum** *min-seconds* | **reload** *reload-seconds* }
Parameter Description	**minimum** *min-seconds*: Indicates the VRRP delay after an interface state changes. **reload** *reload-seconds*: Indicates the VRRP delay after the system starts.
Command	Interface configuration mode

Mode	
Usage Guide	After the delay is configured for a VRRP group on an interface, the VRRP group starts after the delay instead of immediately upon system startup or the interface's resumption, ensuring non-preemption. If the interface receives a VRRP packet during the delay, the delay will be canceled and the VRRP will be started immediately. The two types of delay share a value range of 0 to 60 seconds. This configuration will be effective for both IPv4 and IPv6 VRRP groups of an interface.

Configuring the IPv4 VRRP Version

Command	**vrrp** *group* **version** { **2** | **3** }
Parameter Description	**2**: Indicates VRRPv2. **3**: Indicates VRRPv3.
Command Mode	Interface configuration mode
Usage Guide	Considering the compatibility between VRRPv2 and VRRPv3, specify a standard for IPv4 VRRP based on the actual network condition. VRRPv2 is developed in RFC3768, while VRRPv3 is described in RFC5798. This command is only applicable to IPv4 VRRP.

Specifying a Sub VLAN of a Super VLAN to Receive the IPv4 VRRP Packets

Command	**vrrp detection-vlan {first-subvlan |** *subvlan-id***}**
Parameter Description	**first-subvlan**: Sends IPv4 VRRP packets only to the first UP Sub VLAN interface in a Super VLAN. *subvlan-id*: Sends IPv4 VRRP packets to a specified Sub VLAN.
Command Mode	Interface configuration mode
Usage Guide	This command is used to specify a Sub VLAN of a Super VLAN to receive the IPv4 VRRP packets. IPv4 VRRP packets are sent in a Super VLAN using the following three methods. Packets are sent to the first UP Sub VLAN interface in a Super VLAN, or to a specified Sub VLAN interface in a Super VLAN, or to all the Sub VLAN interfaces in a Super VLAN. If both VRRP and VRRP PLUS are enabled on a Super VLAN interface, VRRP packets are sent to all the UP Sub VLAN interfaces of the Super VLAN interface. This command is configured on a VLAN interface and effective only to Super VLAN interfaces.

Configuring the BFD Support for IPv4 VRRP on an Interface

Command	**vrrp** *group* **bfd** *ip-address*
Parameter Description	*group*: Indicates the VRID of a VRRP group. *ip-address*: Indicates the interface IP address.
Command Mode	Interface configuration mode
Usage Guide	For a Backup router, run this command to correlate an IPv4 VRRP group with BFD without caring the configured IP address. For the Master, as the primary IP address of a Backup router is not known, the router IP address can only be specified by the administrator. If global IPv4 VRRP BFD is configured, this configuration cannot be performed. To enable the BFD support, make sure that IP and BFD session parameters are configured on the target interface.

Configuring Global IPv4 VRRP BFD

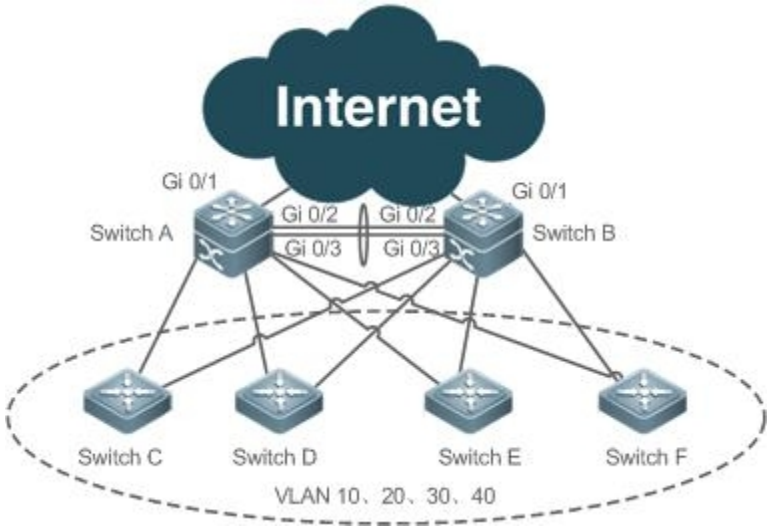
Command	**vrrp bfd** *interface-type interface-number ip-address*
Parameter Description	*interface-type interface-number*: Indicates interface type and ID. *ip-address*: Indicates the interface IP address.
Command Mode	Global configuration mode
Usage Guide	If global IPv4 VRRP BFD is configured, the configured BFD support will be deleted. To enable the BFD support, make sure that IP and BFD session parameters are configured on the target interface. A global IPv4 VRRP BFD session is only applicable to an IPv4 VRRP group consisting of two routers.

Configuration Example

Configuring VRRP+MSTP

Scenario

Figure 4- 8



Configuration Steps

- Enable MSTP on routers (switches A, B, C, D, E and F in this example). Configure VLAN-Instance mapping (mapping VLAN 10 and VLAN 20 to Instance 1, VLAN 30 and VLAN 40 to Instance 2, and the rest VLANs to Instance 0), and configure gateways (Switch A and Switch B in this example) as the root bridges of corresponding instances.
- Add the SVIs of all VLANs to corresponding VRRP backup groups, and configure gateways as the master and backup routers for corresponding backup groups See configuration details in the following table.

Gateway	VLAN ID	SVI	Backup Group	Virtual IP Address	State
Switch A	10	192.168.10.2	VRRP 10	192.168.10.1	Master
Switch B		192.168.10.3			Backup
Switch A	20	192.168.20.2	VRRP 20	192.168.20.1	Master
Switch B		192.168.20.3			Backup
Switch A	30	192.168.30.2	VRRP 30	192.168.30.1	Backup
Switch B		192.168.30.3			Master
Switch A	40	192.168.40.2	VRRP 40	192.168.40.1	Backup
Switch B		192.168.40.3			Master

	● Configure the uplink port (port Gi 0/1 of Switch A and Switch B) of master routers as a monitored interface of master router. ● Step 1: Create VLAN. Create VLAN 10, VLAN 20, VLAN 30 and VLAN 40 respectively on Switch A and Switch B. ● Step 2: Configure MST regions. Map VLAN 10 and VLAN 20 to Instance 1, VLAN 30 and VLAN 40 to Instance 2, and the rest VLANs to Instance 0. ● Step 3: Configure Switch A as the root bridge for MST 0 and MST 1, and Switch B as the root bridge for MST 2. ● Step 4: Enable MSTP. ● Step 5: Configure SVIs of all the VLANs, add the SVIs to corresponding backup groups, and configure virtual IP addresses for the groups. See configuration in the above table. ● Step 6: Configure master routers and backup routers for all the groups. ● Step 7: Configure the uplink ports of master routers as monitored ports of VRRP groups. Caution: Monitored ports should be Layer 3 ports. Step 8: Configure the Internet interfaces of the core routers as AP interfaces.

SwitchA	//Create VLAN 10, VLAN 20, VLAN 30 and VLAN 40 on Switch A.
	SwitchA#configure terminal
	Enter configuration commands, one per line. End with CNTL/Z.
	SwitchA(config)#vlan range 10,20,30,40
	SwitchA(config-vlan-range)#exit
	//Map VLAN 10 and VLAN 20 to Instance 1, VLAN 30 and VLAN 40 to Instance 2, and the rest VLANs to Instance 0.
	SwitchA(config)#spanning-tree mst configuration
	SwitchA(config-mst)#instance 1 vlan 10,20
	%Warning:you must create vlans before configuring instance-vlan relationship
	SwitchA(config-mst)#instance 2 vlan 30,40
	%Warning:you must create vlans before configuring instance-vlan relationship
	SwitchA(config-mst)#exit
	//On Switch A, configure the priority of MST 0 and MST 1 as 4096, and that of MST 2 as 8192.
	SwitchA(config)#spanning-tree mst 0 priority 4096
	SwitchA(config)#spanning-tree mst 1 priority 4096
	SwitchA(config)#spanning-tree mst 2 priority 8192
	//Enabling MSTP
	SwitchA(config)#spanning-tree
	Enable spanning-tree.
	//Configure SVIs of all the VLANs, add the SVIs to corresponding backup groups, and configure virtual IP addresses for the groups.
	SwitchA(config)#interface vlan 10
	SwitchA(config-if-VLAN 10)#ip address 192.168.10.2 255.255.255.0

```
SwitchA(config-if-VLAN 10)#vrrp 10 ip 192.168.10.1
SwitchA(config-if-VLAN 10)#exit
SwitchA(config)#interface vlan 20
SwitchA(config-if-VLAN 20)#ip address 192.168.20.2 255.255.255.0
SwitchA(config-if-VLAN 20)#vrrp 20 ip 192.168.20.1
SwitchA(config-if-VLAN 20)#exit
SwitchA(config)#interface vlan 30
SwitchA(config-if-VLAN 30)#ip address 192.168.30.2 255.255.255.0
SwitchA(config-if-VLAN 30)#vrrp 30 ip 192.168.30.1
SwitchA(config-if-VLAN 30)#exit
SwitchA(config)#interface vlan 40
SwitchA(config-if-VLAN 40)#ip address 192.168.40.2 255.255.255.0
SwitchA(config-if-VLAN 40)#vrrp 40 ip 192.168.40.1
SwitchA(config-if-VLAN 40)#exit
//Increase the priority of the VRRP 10 and VRRP 20 of Switch A to 120.
SwitchA(config)#interface vlan 10
SwitchA(config-if-VLAN 10)#vrrp 10 priority 120
SwitchA(config-if-VLAN 10)#exit
SwitchA(config)#interface vlan 20
SwitchA(config-if-VLAN 20)#vrrp 20 priority 120
SwitchA(config-if-VLAN 20)#exit
//Configure the Gi 0/1 port of Switch A as Route Port and its IP address as10.10.1.1/24.
SwitchA(config)#interface gigabitEthernet 0/1
SwitchA(config-if-GigabitEthernet 0/1)#no switchport
SwitchA(config-if-GigabitEthernet 0/1)#ip address 10.10.1.1 255.255.255.0
SwitchA(config-if-GigabitEthernet 0/1)#exit
//Configure the Gi 0/1 port of Switch A as a monitored port for VRRP 10 and VRRP 20, and a Priority decrement of 30.
SwitchA(config)#interface vlan 10
SwitchA(config-if-VLAN 10)#vrrp 10 track gigabitEthernet 0/1 30
SwitchA(config-if-VLAN 10)#exit
SwitchA(config)#interface vlan 20
SwitchA(config-if-VLAN 20)#vrrp 20 track gigabitEthernet 0/1 30
SwitchA(config-if-VLAN 20)#exit
//Configure ports Gi 0/2 and Gi 0/3 as AP ports, which are Trunk ports.
```

	SwitchA#configure terminal Enter configuration commands, one per line. End with CNTL/Z. SwitchA(config)#interface range gigabitEthernet 0/2-3 SwitchA(config-if-range)#port-group 1 SwitchA(config)#interface aggregateport 1 SwitchA(config-if-AggregatePort 1)#switchport mode trunk

SwitchB	//Create VLAN 10, VLAN 20, VLAN 30 and VLAN 40 on Switch B.
	SwitchB#configure terminal
	Enter configuration commands, one per line. End with CNTL/Z.
	SwitchB(config)#vlan range 10,20,30,40
	SwitchB(config-vlan-range)#exit
	//Map VLAN 10 and VLAN 20 to Instance 1, VLAN 30 and VLAN 40 to Instance 2, and the rest VLANs to Instance 0.
	SwitchB(config)#spanning-tree mst configuration
	SwitchB(config-mst)#instance 1 vlan 10,20
	%Warning:you must create vlans before configuring instance-vlan relationship
	SwitchB(config-mst)#instance 2 vlan 30,40
	%Warning:you must create vlans before configuring instance-vlan relationship
	SwitchB(config-mst)#exit
	//On Switch B, configure the priority of MST 2 as 4096, and that of MST 0 and MST 1 as 8192.
	SwitchB(config)#spanning-tree mst 2 priority 4096
	SwitchB(config)#spanning-tree mst 0 priority 8192
	SwitchB(config)#spanning-tree mst 1 priority 8192
	//Enabling MSTP
	SwitchB(config)#spanning-tree
	Enable spanning-tree.
	//Configure SVIs of all the VLANs, add the SVIs to corresponding backup groups, and configure virtual IP addresses for the groups.
	SwitchB(config)#interface vlan 10
	SwitchB(config-if-VLAN 10)#ip address 192.168.10.3 255.255.255.0
	SwitchB(config-if-VLAN 10)#vrrp 10 ip 192.168.10.1
	SwitchB(config-if-VLAN 10)#exit
	SwitchB(config)#interface vlan 20
	SwitchB(config-if-VLAN 20)#ip address 192.168.20.3 255.255.255.0
	SwitchB(config-if-VLAN 20)#vrrp 20 ip 192.168.20.1

```
SwitchB(config-if-VLAN 20)#exit
SwitchB(config)#interface vlan 30
SwitchB(config-if-VLAN 30)#ip address 192.168.30.3 255.255.255.0
SwitchB(config-if-VLAN 30)#vrrp 30 ip 192.168.30.1
SwitchB(config-if-VLAN 30)#exit
SwitchB(config)#interface vlan 40
SwitchB(config-if-VLAN 40)#ip address 192.168.40.3 255.255.255.0
SwitchB(config-if-VLAN 40)#vrrp 40 ip 192.168.40.1
SwitchB(config-if-VLAN 40)#exit
//Increase the priority of VRRP 30 and VRRP 40 of Switch B to 120.
SwitchB(config)#interface vlan 30
SwitchB(config-if-VLAN 30)#vrrp 30 priority 120
SwitchB(config-if-VLAN 30)#exit
SwitchB(config)#interface vlan 40
SwitchB(config-if-VLAN 40)#vrrp 40 priority 120
SwitchB(config-if-VLAN 40)#exit
//Configure the Gi 0/1 port of Switch B as Route Port and its IP address as 10.10.1.1/24.
SwitchB(config)#interface gigabitEthernet 0/1
SwitchB(config-if-GigabitEthernet 0/1)#no switchport
SwitchB(config-if-GigabitEthernet 0/1)#ip address 10.10.2.1 255.255.255.0
SwitchB(config-if-GigabitEthernet 0/1)#exit
//Configure the Gi 0/1 port of Switch B as a monitored port for VRRP 30 and VRRP 40, and the Interface-Priority as 30.
SwitchB(config)#interface vlan 30
SwitchB(config-if-VLAN 30)#vrrp 30 track gigabitEthernet 0/1 30
SwitchB(config-if-VLAN 30)#exit
SwitchB(config)#interface vlan 40
SwitchB(config-if-VLAN 40)#vrrp 40 track gigabitEthernet 0/1 30
SwitchB(config-if-VLAN 40)#exit
//Configure ports Gi 0/2 and Gi 0/3 as AP ports, which are Trunk ports.
SwitchB #configure terminal
Enter configuration commands, one per line.    End with CNTL/Z.
SwitchB (config)#interface range gigabitEthernet 0/2-3
SwitchB (config-if-range)#port-group 1
SwitchB (config)#interface aggregateport 1
```

	SwitchB (config-if-AggregatePort 1)#switchport mode trunk
Verification	

Switch A	Check the configuration.
	SwitchA#show running-config
	!
	vlan 10
	!
	vlan 20
	!
	vlan 30
	!
	vlan 40
	!
	spanning-tree
	spanning-tree mst configuration
	  instance 0 vlan 1-9, 11-19, 21-29, 31-39, 41-4094
	  instance 1 vlan 10, 20
	  instance 2 vlan 30, 40
	spanning-tree mst 0 priority 4096
	spanning-tree mst 1 priority 4096
	spanning-tree mst 2 priority 8192
	interface GigabitEthernet 0/1
	  no switchport
	  no ip proxy-arp
	  ip address 10.10.1.1 255.255.255.0
	!
	interface GigabitEthernet 0/2
	  port-group 1
	!
	interface GigabitEthernet 0/3
	  port-group 1
	!
	interface AggregatePort 1
	  switchport mode trunk

```
!
interface VLAN 10
 no ip proxy-arp
 ip address 192.168.10.2 255.255.255.0
 vrrp 10 priority 120
 vrrp 10 ip 192.168.10.1
 vrrp 10 track GigabitEthernet 0/1 30
!
interface VLAN 20
 no ip proxy-arp
 ip address 192.168.20.2 255.255.255.0
 vrrp 20 priority 120
 vrrp 20 ip 192.168.20.1
 vrrp 20 track GigabitEthernet 0/1 30
!
interface VLAN 30
 no ip proxy-arp
 ip address 192.168.30.2 255.255.255.0
 vrrp 30 ip 192.168.30.1
!
interface VLAN 40
 no ip proxy-arp
 ip address 192.168.40.2 255.255.255.0
 vrrp 40 ip 192.168.40.1
//Check VRRP status.
SwitchA#show vrrp brief
Interface    Grp    Pri    timer    Own    Pre    State     Master addr     Group addr
VLAN 10      10     120    3        -      P      Master    192.168.10.2    192.168.10.1
VLAN 20      20     120    3        -      P      Master    192.168.20.2    192.168.20.1
VLAN 30      30     100    3        -      P      Backup    192.168.30.3    192.168.30.1
VLAN 40      40     100    3        -      P      Backup    192.168.40.3    192.168.40.1
//Disconnect the uplink of Switch A, and check VRRP status.
SwitchA#show vrrp brief
Interface    Grp    Pri    timer    Own    Pre    State     Master addr     Group addr
```

	VLAN 10 10 90 3 - P Backup 192.168.10.3 192.168.10.1 VLAN 20 20 90 3 - P Backup 192.168.20.3 192.168.20.1 VLAN 30 30 100 3 - P Backup 192.168.30.3 192.168.30.1 VLAN 40 40 100 3 - P Backup 192.168.40.3 192.168.40.1

Switch B	//Check the configuration.
	SwitchB#show running-config
	!
	vlan 10
	!
	vlan 20
	!
	vlan 30
	!
	vlan 40
	!
	spanning-tree
	spanning-tree mst configuration
	  instance 0 vlan 1-9, 11-19, 21-29, 31-39, 41-4094
	  instance 1 vlan 10, 20
	  instance 2 vlan 30, 40
	spanning-tree mst 0 priority 8192
	spanning-tree mst 1 priority 8192
	spanning-tree mst 2 priority 4096
	interface GigabitEthernet 0/1
	  no switchport
	  no ip proxy-arp
	  ip address 10.10.2.1 255.255.255.0
	!
	interface GigabitEthernet 0/2
	  port-group 1!
	interface GigabitEthernet 0/3
	  port-group 1

```
!
interface AggregatePort 1
  switchport mode trunk
!
interface VLAN 10
  no ip proxy-arp
  ip address 192.168.10.3 255.255.255.0
  vrrp 10 ip 192.168.10.1
!
interface VLAN 20
  no ip proxy-arp
  ip address 192.168.20.3 255.255.255.0
  vrrp 20 ip 192.168.20.1
!
interface VLAN 30
  no ip proxy-arp
  ip address 192.168.30.3 255.255.255.0
  vrrp 30 priority 120
  vrrp 30 ip 192.168.30.1
  vrrp 30 track GigabitEthernet 0/1 30
!
interface VLAN 40
  no ip proxy-arp
  ip address 192.168.40.3 255.255.255.0
  vrrp 40 priority 120
  vrrp 40 ip 192.168.40.1
  vrrp 40 track GigabitEthernet 0/1 30
//Check VRRP status.
SwitchB#show vrrp brief
Interface    Grp    Pri    timer    Own    Pre    State      Master addr      Group addr
VLAN 10      10     100    3        -      P      Backup     192.168.10.2     192.168.10.1
VLAN 20      20     100    3        -      P      Backup     192.168.20.2     192.168.20.1
VLAN 30      30     120    3        -      P      Master     192.168.30.3     192.168.30.1
VLAN 40      40     120    3        -      P      Master     192.168.40.3     192.168.40.1
```

```
//Disconnect the uplink of Switch B, and check VRRP status.
SwitchB#show vrrp brief
Interface    Grp   Pri   timer   Own   Pre   State    Master addr    Group addr
VLAN 10      10    100   3       -     P     Master   192.168.10.3   192.168.10.1
VLAN 20      20    100   3       -     P     Master   192.168.20.3   192.168.20.1
VLAN 30      30    120   3       -     P     Master   192.168.30.3   192.168.30.1
VLAN 40      40    120   3       -     P     Master   192.168.40.3   192.168.40.1
```

Common Errors

- Different virtual IP addresses are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- Different VRRP advertisement intervals are configured on the routers in a VRRP group and the learning timer is not configured, resulting in multiple Master routers in the group.
- Different VRRP versions are configured on the routers in a VRRP group, resulting in multiple Master routers in the group.
- For VRRPv2, the Ethernet interfaces of the routers in a VRRP group are all in plain text authentication mode but inconsistent in authentication strings, resulting in multiple Master routers in the group.

4.5 Monitoring

Displaying

Description	Command
Displays the brief or detailed information of IPv4/IPv6 VRRP.	**show** [**ipv6**] **vrrp** [**brief** | *group*]
Displays the information of an IPv4/IPv6 VRRP group on a specified interface.	**show** [**ipv6**] **vrrp interface** *type number* [**brief**]
Displays the statistics of VRRP packets.	**show vrrp packet statistics** [*interface-type interface-number*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs VRRP errors, events, packets and status.	**debug** [**ipv6**] **vrrp**
Debugs VRRP errors.	**debug** [**ipv6**] **vrrp errors**
Debugs VRRP events.	**debug** [**ipv6**] **vrrp events**
Debugs VRRP packets.	**debug vrrp packets** [**acl** *acl-id* | [**icmp** | **protocol**] **interface** *type number* [*group*]] **debug ipv6 vrrp packets** [**acl** *acl-name* | [**icmp** | **protocol**] **interface** *type number* [*group*]]
Debugs VRRP status.	**debug** [**ipv6**] **vrrp state**

5 Configuring VRRP Plus

5.1 Overview

Virtual Router Redundancy Protocol Plus (VRRP Plus) is an extension of VRRP. It uses VRRP to implement gateway backup and load balancing in the IEEE 802.3 local area network (LAN).

A disadvantage of VRRP is that the router in backup state cannot forward packets. To use VRRP to implement load balancing, you need to manually configure multiple VRRP groups and set the gateway addresses of hosts in the LAN to virtual IP addresses of different VRRP groups. This increases the workload of the network administrator. VRRP Plus is designed to address this issue.

With VRRP Plus, load balancing is automatically implemented. That is, traffic of different hosts is automatically distributed to members of the VRRP Plus group, and it is unnecessary to configure multiple VRRP groups or set the gateway addresses of hosts in the LAN to virtual IP addresses of different VRRP groups. This greatly reduces the workload of the network administrator.

5.2 Applications

Application	Description
Enabling Load Balancing Within a VRRP Group	Implement load balancing within a VRRP group without configuring multiple groups or configuring different default gateways for hosts.

5.2.1 Enabling Load Balancing Within a VRRP Group

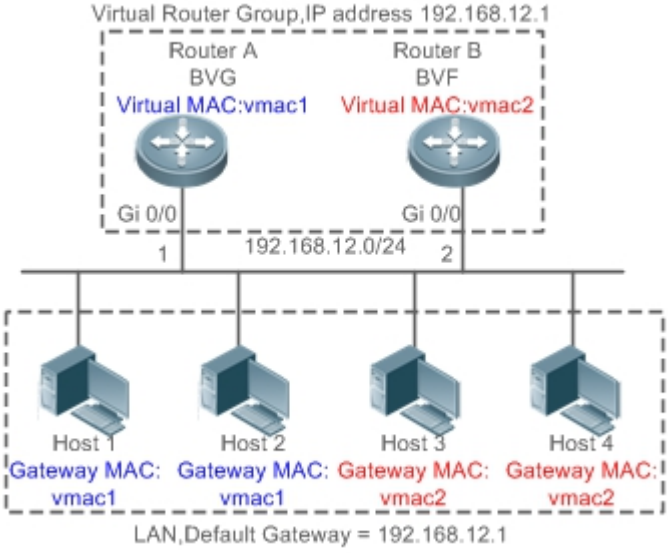
Scenario

Enable load balancing within a VRRP group without configuring without configuring multiple VRRP groups or configuring different default gateways for hosts.

As shown in Figure 5- 1, configure data as follows:

- Configure a VRRP group that consists of Router A and Router B, and enable the VRRP Plus function.
- Configure the default gateway of each host as the master virtual IP address of the VRRP group.

Figure 5- 1 Application topology of IPv4 VRRP Plus



Remarks	4. Two layer-3 (L3) devices, Router A and Router B, form a VRRP Plus group, and the virtual IP address of the group is 192.168.12.1. Router A is the master device of VRRP and functions as a balancing virtual gateway (BVG). Router B is the backup device of VRRP and functions as a balancing virtual forwarder (BVF). 5. Host 1 to Host 4 are hosts in the LAN with the network segment 192.168.12.0/24. Their default gateway addresses are set to the virtual IP address 192.168.12.1 of the VRRP Plus group. 6. The load balancing policy is configured on the device to respond to the ARP requests sent from different hosts. For example, when Host 1 and Host 2 request the gateway ARP, the MAC address 0000.5e00.0101 is returned to Host 1 and Host 2. When Host 3 and Host 4 request the gateway ARP, the MAC address 001A.A916.0201 is returned to Host 3 and Host 4. In this way, packets exchanged between Host 1/Host 2 and the external network are sent to Router A, and packets exchanged between Host 3/Host 4 and the external network are sent to Router B, thereby implementing load balancing.

Deployment

- Deploy VRRP Plus on Router A and Router B to implement load balancing on the local host.

5.3 Features

Basic Concepts

BVG

The BVG allocates virtual MAC addresses to members of the VRRP Plus group. It responds to the gateway ARP/ND requests in the LAN, and forwards packets of hosts in the LAN.

BVF

The BVF forwards packets of hosts in the LAN. If a virtual MAC address is allocated to a BVF, the BVF participates in packet forwarding; otherwise, the BVF does not participate in packet forwarding.

Overview

Feature	Description
VRRP Plus	Extend VRRP and use VRRP to implement gateway backup and load balancing in the IEEE 802.3 LAN.

5.3.1 VRRP Plus

With VRRP Plus, load balancing is automatically implemented. That is, traffic of different hosts is automatically distributed to members of the VRRP Plus group, and it is unnecessary to configure multiple VRRP groups or set the gateway addresses of hosts in the LAN to the virtual IPv4/IPv6 addresses of different VRRP groups.

Basic Principles

Hosts in a LAN use the unified gateway IPv4/IPv6 address (that is, virtual IP address of the VRRP group). When different hosts request the gateway ARP/ND, the BVG responds with different virtual MAC addresses. In this way, traffic of different hosts are distributed to different members of the VRRP Plus group, thereby implementing load balancing.

Relationship Between VRRP Plus and VRRP

VRRP Plus relies on VRRP, and runs in the following way:

A master device in VRRP corresponds to a BVG in VRRP Plus, and a backup device in VRRP corresponds to a BVF in VRRP Plus. Gateway addresses of hosts in the LAN are set to the virtual IPv4/IPv6 address of VRRP.

MAC Address Allocation Rules of the BVG and BVF

The BVG allocates virtual MAC addresses to BVFs. For an IPv4 VRRP Plus group, the BVG directly uses the virtual MAC address of VRRP to ensure compatibility between IPv4 VRRP Plus and VRRP. That is, the virtual MAC address used by the BVG is 00-00-5E-00-01-{VRID}, where VRID is the VRRP group number. The virtual MAC address used by a BVF is 00-1A-A9-16-{MemberID}-{VRID}, where MemberID is the member ID of the BVF in the VRRP Plus group. Currently, a VRRP Plus group can have up to four members. The BVG uses the member ID 01, and the other BVFs use the member IDs 02 to 04.

Load Balancing Policy of VRRP Plus

The BVG responds to the gateway ARP/NS requests sent from hosts in a LAN. Based on the specific load balancing policy, the BVG responds hosts with different virtual MAC addresses. There are three types of load balancing policies:

- Host-dependent policy: A specified virtual MAC address is used to respond to the requests sent by a specified host.
- Round-robin policy: Virtual MAC addresses in the backup group are used in a cyclic manner to respond to the gateway ARP/NS requests sent by hosts.
- Weighted policy: The ARP/NA requests are responded based on the forwarding capability of each device.

If the load balancing mode is changed, load balancing is always implemented in the new load balancing mode. For example, if the polling response mode is previously used, and later the weighted mode is used, load balancing is implemented in weighted mode regardless of the earlier responses of the device. If the weighted policy is used, and the total weight of virtual routers in a VRRP Plus group is 0, the ARP/NS requests are not responded.

Proxy of the Virtual MAC Address

When a device with a virtual MAC address becomes faulty in the backup group, traffic of hosts that use this virtual MAC address as the gateway MAC address will be interrupted.

The BVG in the VRRP Plus backup group can quickly detect the fault, and automatically allocates the virtual MAC address of the faulty BVF to another device in the backup group. The new device acts as the proxy of the faulty device to forward packets of the virtual MAC address. In addition, this proxy device takes over traffic of original hosts to prevent traffic interruption. The virtual MAC address allocated to a device in the backup group can be called master virtual MAC address, and the virtual MAC address used by this device on behalf of another device is called proxy virtual MAC address.

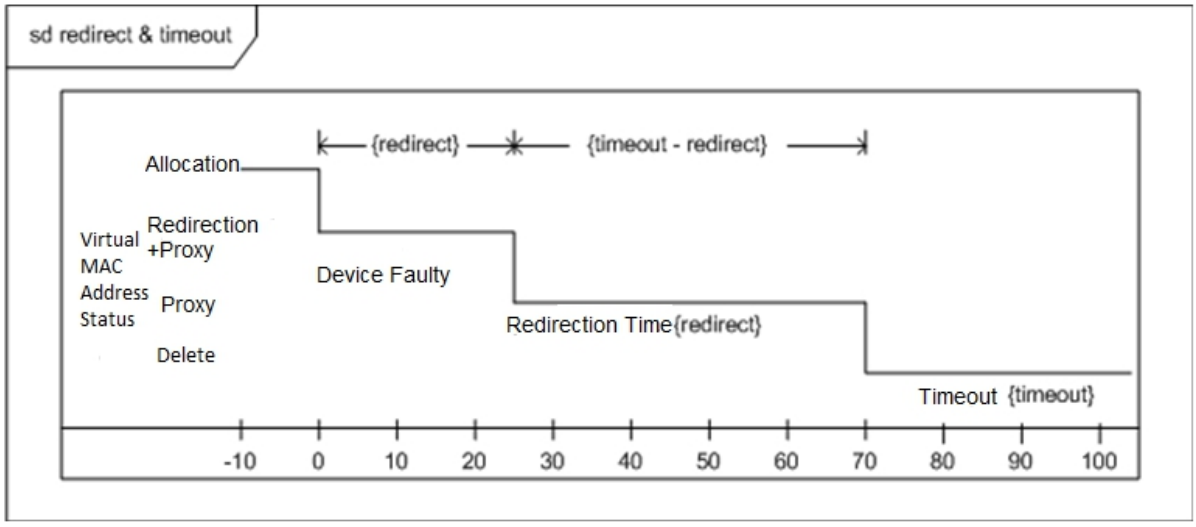
Redirection Time and Timeout of the Proxy Virtual MAC Address

VRRP Plus provides the proxy function for the virtual MAC address so that another device can take the place of a faulty device with a virtual MAC address to forward packets. If the BVF is recovered from the fault, its forwarding role is recovered and the BVF continues to forward packets of the virtual MAC address allocated to this BVF. If the faulty BVF is not recovered, the backup group stops redirecting traffic to this virtual MAC address. That is, when ARP requests are received again, this virtual MAC address is no longer responded. After a sufficient long period of time, it is believed that hosts that use the MAC address as the gateway MAC address already update the ARP/ND table entry of the gateway address, and the traffic is already taken over by other devices. At this time, this virtual MAC address can be deleted, and packets sent to this virtual MAC address are dropped.

VRRP Plus supports configuration of the redirection time and timeout of the backup group. When a device is faulty, the backup group allocates the virtual MAC address of the faulty device to another device. Within the redirection time, the backup group continues to use this virtual MAC address to respond the ARP/NS requests. When the redirection time expires, the backup group no longer uses this virtual

MAC address to respond the requests. When the timeout elapses, the backup group deletes this virtual MAC address and stops using this virtual MAC address for proxy forwarding. Figure 5- 2 shows the changes to the role of the virtual MAC address within the redirection time and timeout.

Figure 5- 2 Changes to the Role of the Virtual MAC Address Within the Redirection Time and Timeout



Weight-based Forwarding

VRRP Plus supports the weight configuration of the backup group. Different weights are configured for different devices. In this way, more traffic is distributed to the device with a greater weight and less traffic is distributed to the device with a smaller weight, thereby fully utilizing the forwarding performance of different devices. When the weight of a BVF in the backup group is smaller than the lower threshold, the BVF automatically exits from the forwarding role. When the weight recovers and is greater than the upper threshold, the BVF automatically applies for the forwarding role. The forwarding role can be recovered when one or more remaining virtual MAC addresses or proxy virtual MAC addresses exist.

Association of VRRP Plus with BFD

VRRP Plus supports association with bidirectional forwarding detection (BFD) to adjust the weight based on the link status. Each device in a backup group can associate its weight with the link status. When a link is abnormal or interrupted, the device automatically decreases its weight. When the weight is too low, the device automatically exits from the forwarding role. If the backup group is currently using the weighted load balancing policy, traffic can be distributed based on the new weight. When the associated link recovers, the device can automatically restore its original weight and the forwarding role. If the backup group is currently using the weighted load balancing policy, traffic can be distributed based on the recovered weight.

Weight-based Forwarding Seizure

VRRP Plus supports the function of seizing the forwarding role. In VRRP Plus, at most four devices can participate in load balancing. That is, a VRRP Plus backup group generates at most four virtual MAC addresses. If more than four devices are added to a VRRP Plus group, only four devices participate in packet forwarding. The remaining devices only listen to the status of other devices and do not participate in packet forwarding. Only when a device participating in packet forwarding is faulty, another device that originally does not participate in packet forwarding will take the place of the faulty device to forward packets. Assume that a VRRP Plus backup group already has four devices and all these devices participate in packet forwarding; a fifth device is added to the VRRP Plus group, and the forwarding capability of this device is strong or the original forwarding role encounters a link failure and consequently degradation of forwarding

performance. In this case, if the seizure mode is enabled, the fifth device can seize the forwarding role from a device with a smaller weight (that is, with lower forwarding capability). A greater weight is configured for a device with stronger forwarding capability. When the weight of a device in listening state is found greater than that of a forwarding device, the device in listening state automatically seizes the forwarding role from the forwarding device. That is, the device with stronger forwarding capability forwards packets, whereas the device with lower forwarding capability is in listening state. This can minimize the waste of resources.

The BVG in a backup group is responsible for allocation of virtual MAC addresses. Therefore, the BVG role cannot be seized, and only the forwarding role of a BVF can be seized. If the BVG device is faulty, VRRP re-elects a new master device, which assumes the BVG role.

Factors Affecting the Forwarding Policy

7. After VRRP Plus is configured, the ARP/NS requests are received from hosts can be responded based on different load balancing policies to implement load balancing among these hosts. However, load balancing cannot be implemented for hosts that have learned the VRRP virtual gateway addresses before configuration of VRRP Plus. Therefore, if VRRP Plus is configured after the VRRP state is changed to Master, real load balancing cannot be implemented before aging of the ARP/NDs learned by hosts. Load balancing is implemented only after the gateway ARP/NDs recorded by the hosts age and the hosts request for new gateway addresses.
8. Periodical sending of gratuitous ARPs on an interface also affect the load balancing function of VRRP Plus. When VRRP Plus is enabled, the function of sending gratuitous ARPs of VRRP virtual IP addresses will be disabled. When an virtual IP address overlaps with an actual IP address, gratuitous ARPs of this address are no longer sent.
9. When an address conflict occurs between a host and the local device, the ARP/NA module will broadcast gratuitous ARP/NA packets of this address. If a conflict of the VRRP Plus virtual address occurs, sending gratuitous ARP/NA packet will result re-learning of the host's gateway MAC address, which negatively affects the load balancing function of VRRP Plus. Therefore, the load balancing function of VRRP Plus is currently not supported in this scenario.

5.4 Configuration

Configuration Item	Description and Command	
Configuring VRRP Plus	(Mandatory) It is used to enable the VRRP Plus function.	
	vrrp balance	Enables the VRRP Plus function of a VRRP backup group with the specified group ID in interface configuration mode.
	(Optional) It is used to configure parameters of a VRRP Plus backup group.	
	vrrp load-balancing	Configures the load balancing policy of VRRP Plus in interface configuration mode.
	vrrp timers redirect	Configures the redirection time and timeout of the proxy virtual MAC address in a VRRP Plus backup group in interface configuration mode.
	vrrp weighting	Configures the weight and upper and lower thresholds of a VRRP Plus backup group in interface configuration mode.
	vrrp forwarder preempt	Configures the forwarding seizure function of a VRRP Plus backup group in interface configuration mode.

5.4.1 Configure VRRP Plus

Configuration Effect

- Enable the VRRP Plus function. (By default, this function is disabled.)
- Configure a weight tracking object of a VRRP Plus backup group.

Notes

- To enable the VRRP Plus function, you must configure the VRRP virtual IP address for the corresponding backup group.

Configuration Steps

Enabling VRRP Plus on an Interface

- By default, VRRP Plus is enabled. Perform this configuration if VRRP Plus is required.

Configuring the Load Balancing Policy of VRRP Plus

- After VRRP Plus is enabled, the host-dependent load balancing policy is used by default.

Configuring the Redirection Time and Timeout of the Proxy Virtual MAC Address in a VRRP Plus Backup Group

- After VRRP Plus is enabled, the redirection time is set to 300s and timeout is set to 14,400s by default.

Configuring the Weight and Upper and Lower Thresholds of a VRRP Plus Backup Group

- After VRRP Plus is enabled, the weight of the backup group is set to 100, the lower threshold to 1, and the upper threshold to 100 by default.

Configuring the Forwarding Seizure Function of a VRRP Plus Backup Group

- After VRRP Plus is enabled, the forwarding seizure function is enabled by default.

Configuring the Weight Tracking Object of a VRRP Plus Backup Group

- The weight tracking object of a VRRP Plus backup group is disabled by default. Perform this configuration if the tracking function is required.

Verification

- Run the **show** *group* **vrrp balance** command to display the VRRP backup group configuration. If the backup group has the packet forwarding tasks, "local" is displayed in the **forwarders** column, and the virtual MAC address allocated to this backup group is also displayed.

Related Commands

Enabling VRRP Plus on an Interface

Command	**vrrp** *group* **balance**
Parameter Description	*group*: Indicates the ID of the VRRP group. The value range of the group ID varies according to the product model.
Command Mode	Interface configuration mode

Usage Guide	VRRP Plus can be enabled only after a VRRP group is configured.

Configuring the Load Balancing Policy of a VRRP Plus Backup Group

Command	**vrrp** *group* **load-balancing**{**host-dependent | round-robin | weighted** }
Parameter Description	*group*: Indicates the ID of the VRRP group. **host-dependent**: Indicates the host-dependent load balancing policy. **round-robin**: Indicates the round-robin load balancing policy. **weighted**: Indicates the weighted load balancing policy.
Command Mode	Interface configuration mode
Usage Guide	After VRRP Plus is enabled, the host-dependent load balancing policy is used by default. The load balancing policy of the entire backup group is determined by the policy configured on the BVG. If you wish to use the same load balancing policy after the role of the BVG device changes, configure the same policy on all devices in the backup group.

Configuring the Redirection Time and Timeout of the Proxy Virtual MAC Address in a VRRP Plus Backup Group

Command	**vrrp** *group* **timers redirect** *redirect timeout*
Parameter Description	*group*: Indicates the ID of the VRRP group. *redirect*: Indicates the redirection time. The value ranges from 0 to 3,600s. The default value is 300s, that is, 5 minutes. *timeout*: Indicates the timeout time. The value ranges from (redirect + 600) to 64,800s. The default value is 14400, that is, 4 hours.
Command Mode	Interface configuration mode
Usage Guide	After VRRP Plus is enabled, the redirection time is set to 300s and timeout is set to 14,400s by default. When a device is faulty, the backup group allocates the virtual MAC address of the faulty device to another device. Within the redirection time, the backup group continues to use this virtual MAC address to respond the ARP/NS requests. When the redirection time expires, the backup group no longer uses this virtual MAC address to respond the requests. When the timeout elapses, the backup group deletes this virtual MAC address.

Configuring the Weight and Upper and Lower Thresholds of a VRRP Plus Backup Group

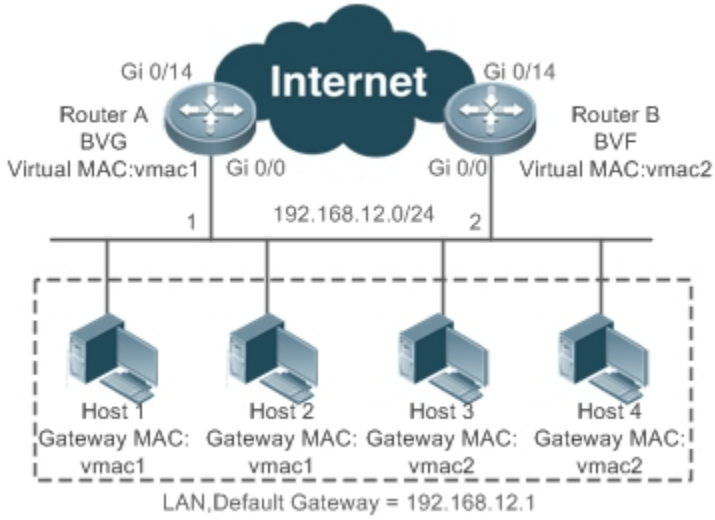
Command	**vrrp** *group* **weighting** *maximum* [**lower** *lower*] [**upper** *upper*]
Parameter Description	*maximum*: Indicates the weight of the backup group. The value ranges from 2 to 254. The default value is 100. **lower** *lower*: Indicates the lower threshold of the backup group. The value ranges from 1 to (**maximum** - 1). The default value is 1. **upper** *upper:* Indicates the upper threshold of the backup group. The value ranges from **lower** to **maximum**. The default value is 100.
Command Mode	Interface configuration mode
Usage Guide	After VRRP Plus is enabled, the weight and upper and lower thresholds of a VRRP Plus backup group are configured by default. You can use this command to configure different weights for different devices so that more traffic is distributed to the device with a greater weight and less traffic is distributed to the device with a smaller weight. When the weight of a BVF in the backup group is lower than the lower threshold, the BVF automatically exits from the forwarding role. When the weight recovers and is higher than the upper threshold, the forwarding role of the BVF is automatically restored.

Configuring the Forwarding Seizure Function of a VRRP Plus Backup Group

Command	**vrrp** *group* **forwarder preempt**
Parameter Description	*group*: Indicates the ID of the VRRP group.
Command Mode	Interface configuration mode
Usage Guide	After VRRP Plus is enabled, the forwarding seizure function is enabled by default. VRRP Plus supports configuration of the forwarding seizure function of a backup group. When the weight of a device in listening state is found greater than that of a forwarding device, the device in listening state automatically seizes the forwarding role from the forwarding device. That is, the device with stronger forwarding capability forwards packets, whereas the device with lower forwarding capability is in listening state.

Configuration Example

Enabling Load Balancing Within an IPv4 VRRP Group

Scenario **Figure 5- 3**	
Configuration Steps	● Configure a VRRP group and enable VRRP Plus respectively on Router A and Router B. Configure the local IP addresses so that Router A becomes a BVG (master) device, and Router B becomes a BVF (backup) device. ● Configure the weighted load balancing policy. Configure the weight tracking object, and set the decrement value to 100. ● Retain default configurations of the weight, upper and lower thresholds, redirection time, timeout, and forwarding seizure of the backup group. ● Set the default gateway addresses of Host 1 to Host 4 in the LAN to the virtual IP address of VRRP, that is, 192.168.12.1.
Router A	FSA#config FSA(config)#interface GigabitEthernet0/0 // 'no switchport' is used on the switch. FSA(config-if-GigabitEthernet 0/0)#no switchport FSA(config-if-GigabitEthernet 0/0)#ip address 192.168.12.3 255.255.255.0

	FSA(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.12.1 FSA(config-if-GigabitEthernet 0/0)#vrrp 1 balance FSA(config-if-GigabitEthernet 0/0)#vrrp 1 load-balancing weighted
Router B	FSB#config FSB(config)#interface GigabitEthernet0/0 FSB(config-if-GigabitEthernet 0/0)#no switchport FSB(config-if-GigabitEthernet 0/0)#ip address 192.168.12.2 255.255.255.0 FSB(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.12.1 FSB(config-if-GigabitEthernet 0/0)#vrrp 1 balance FSB(config-if-GigabitEthernet 0/0)#vrrp 1 load-balancing weighted
Verification	Run the **show vrrp balance** command to display the configuration of the VRRP Plus group. If the backup group has the packet forwarding tasks, "local" is displayed in the **forwarders** column, and the virtual MAC address allocated to this backup group is also displayed.
Router A	FSA# show vrrp balance interface GigabitEthernet0/0 State is BVG Virtual IP address is 192.168.12.1 Hello time 1 sec, hold time 3 sec Load balancing: weighted Redirect time 300 sec, forwarder time-out 14400 sec Weighting 100 (configured 100), thresholds: lower 1, upper 100 There are 2 forwarders Forwarder 1 (local) MAC address: 0000.5e00.0101 Owner ID is 0000.0001.0006 Preemption disabled (BVG cannot be preempted) Forwarder 2 MAC address: 001a.a916.0201 Owner ID is 00d0.f822.33a3 Preemption enabled
Router B	FSB# show vrrp balance interface GigabitEthernet0/0 State is BVF

```
Virtual IP address is 192.168.12.1
Hello time 1 sec, hold time 3 sec
Load balancing: weighted
Redirect time 300 sec, forwarder time-out 14400 sec
Weighting 100 (configured 100), thresholds: lower 1, upper 100
There are 2 forwarders
Forwarder 1
  MAC address:
    0000.5e00.0101
  Owner ID is 0000.0001.0006
  Preemption disabled (BVG cannot be preempted)
Forwarder 2 (local)
  MAC address:
    001a.a916.0201
  Owner ID is 00d0.f822.33a3
  Preemption enabled
```

Common Errors

- VRRP Plus does not take effect because the VRRP virtual IP address is not configured for the related group.

5.5 Monitoring

Displaying

Description	Command
Displays the brief or detailed configuration of VRRP Plus.	**show vrrp balance**
Displays the actions of the VRRP Plus group on a specified interface.	**show vrrp balance interface**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the VRRP Plus function.	**debug vrrp balance**
Debugs errors.	**debug vrrp balance error**
Debugs events of the VRRP Plus group.	**debug vrrp balance event**
Debugs the messages between the VRRP module and the track module.	**debug vrrp balance messages**
Debugs the VRRP Plus packets.	**Debug vrrp balance packets**
Debugs the VRRP Plus group status.	**debug vrrp balance state**

Description	Command
Debugs the timers of the VRRP Plus group.	**debug vrrp balance timer**

6 Configuring BFD

6.1 Overview

Communication failures will interrupt networking and thus affect services. Therefore, it is essential to rapidly locate communication failures on links with adjacent devices to ensure a timely action and service availability. Bidirectional Forwarding Detection (BFD) provides a method of rapidly detecting connectivity of the forwarding path between two adjacent routers in an underloaded way. It can quickly spot faults on the bidirectional forwarding path between two routers for upper-layer protocols such as routing protocols and Multi-Protocol Label Switching (MPLS). As a result, a standby forwarding path is adopted to maintain the performance of the existing network.

Protocols and Standards

- draft-ietf-bfd-base-09: Bidirectional Forwarding Detection
- draft-ietf-bfd-generic-05: Generic Application of BFD
- draft-ietf-bfd-mib-06: Bidirectional Forwarding Detection Management Information Base
- draft-ietf-bfd-v4v6-1hop-09: BFD for IPv4 and IPv6 (Single Hop)
- draft-ietf-bfd-multihop-07: BFD for IPv4 and IPv6 (Multi-hop)
- draft-ietf-bfd-mpls-07: BFD For MPLS LSPs

Currently, draft-ietf-bfd-mib-06 and draft-ietf-bfd-multihop-07 are not supported.

6.2 Applications

Application	Description
BFD Support for OSPF	OSPF utilizes BFD to rapidly detect the neighbor status.
BFD Support for Static Routing	Static routing utilizes BFD to rapidly detect the next-hop reachability of a route.

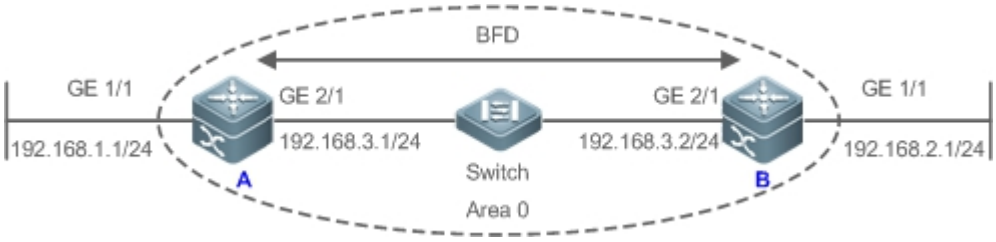
6.2.1 BFD Support for OSPF

Scenario

The Open Shortest Path First (OSPF) protocol dynamically discovers a neighbor by using hello packets. After BFD is enabled, a BFD session is established with the neighbor in the full adjacency to detect the neighbor status. When the neighbor fails, OSFP immediately performs network convergence. The convergence time can be shortened from 120 seconds (by default, on a non-broadcast network, OSPF hello packets are transmitted at an interval of 30 seconds and the neighbor failure time is four times the interval, that is, 120 seconds) to 1 second.

Use the following figure as an example. Router A and Router B are connected through a Layer-2 switch, OSPF is configured on the routers to establish routes, and BFD support for OSPF is enabled on the interfaces of Router A and Router B. When the link between Router B and the Layer-2 switch malfunctions, BFD can rapidly detect the fault and advertise it to OSPF, so as to trigger fast OSPF convergence.

Figure 5- 1



Remarks	A and B are routers. Switch is a Layer-2 switch. A and B are connected through the Layer-2 switch.

Deployment

- Configure IP addresses for interconnected interfaces of Router A and Router B.
- Run OSPF on Router A and Router B.
- Set BFD parameters on interconnected interfaces of Router A and Router B.
- Enable BFD support for OSPF on Router A and Router B.

6.2.2 BFD Support for Static Routing

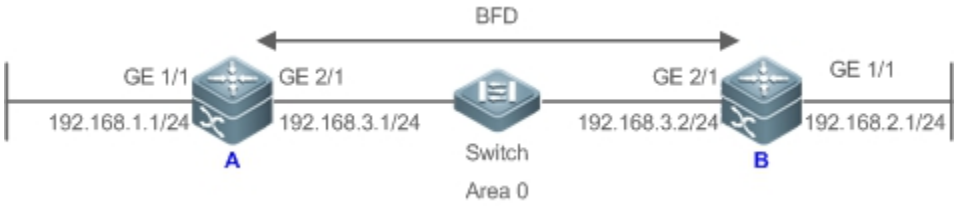
Scenario

BFD support for static routing prevents routers from selecting a faulty static route as the forwarding path and enables rapid routing failover by using an available backup forwarding path.

Different from dynamic routing protocols, static routing does not have the neighbor discovery (ND) mechanism. When BFD support for static routing is configured, the next-hop reachability of a static route relies on the BFD session status. If a BFD session fails, the next hop of a static route is thought unreachable and will not be added to the routing information base (RIB).

Use the following figure as an example. Router A and Router B are connected through a Layer-2 switch, static routing is configured on the routers to establish forwarding paths, and BFD support for static routing is enabled on the interfaces of Router A and Router B. When the link between Router B and the Layer-2 switch malfunctions, BFD can rapidly detect the fault and advertise it to static routing, so as to trigger the system to delete the static route from the RIB, thereby preventing routing errors.

Figure5- 2



Remarks	A and B are routers. Switch is a Layer-2 switch. A and B are connected through the Layer-2 switch.

Deployment

- Configure IP addresses for interconnected interfaces of Router A and Router B.

- Configure static routing on Router A and Router B.
- Set BFD parameters for interconnected interfaces of Router A and Router B.
- Enable BFD support for static routing on Router A and Router B.

6.3 Features

Basic Concepts

Packet Format

Detection packets transmitted by BFD are User Datagram Protocol (UDP) packets, which are classified into control packets and echo packets. Echo packets concern only the local system of a BFD session. Therefore, their formats are not specified. BFD specifies the format of only control packets. Currently, there are two versions (version 0 and version 1) for the format of control packets. Version 1 is adopted by default for establishing a BFD session. If a device receives packets of version 0 from the peer system, the device automatically switches to version 0.

Figure 5- 3

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Vers |  Diag   |Sta|P|F|C|A|D|M|  Detect Mult  |    Length     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       My Discriminator                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Your Discriminator                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Desired Min TX Interval                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Required Min RX Interval                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Required Min Echo RX Interval                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Field	Description
Vers	Indicates the BFD protocol version number, which is 1 currently.
Diag	Indicates the cause for the local system's last change in session state, including: 0 -- No Diagnostic. 1 -- Control Detection Time Expired 2 -- Echo Function Failed 3 -- Neighbor Signaled Session Down 4 -- Forwarding Plane Reset 5 --Path Down 6 -- Concatenated Path Down 7 --Administratively Down
Sta	Indicates the BFD local session state, including: 0 -- AdminDown. 1 -- Down. 2 -- Init. 3 -- Up.

Field	Description
P	Indicates that the transmitter in a BFD session adds this bit in a verification request upon parameter changes, waiting for the peer response.
F	Indicates the bit that must be set in the response packet for responding to the P bit.
C	Indicates the control plane independent. If set, changes of the control plane do not affect BFD detection. For example, if the control plane is OSPF, when OSPF is restarted or experiences graceful restart (GR), BFD can continue to detect the link status.
A	Indicates the authentication present. If set, a session is to be authenticated.
D	Indicates the demand request. If set, the transmitter desires to detect links in Demand mode.
M	Indicates the multipoint bit to be used in point-to-multipoint extensions. It must be set to 0 currently.
Detect Mult	Indicates the detection timeout multiplier. It is used by the detector to calculate the detection timeout time.
Length	Indicates the packet length.
My Discriminator	Indicates the discriminator of the local end connected by a BFD session.
Your Discriminator	Indicates the discriminator of the remote end connected by a BFD session.
Desired Min Tx Interval	Indicates the minimum interval of transmitting BFD packets supported by the local end.
Required Min RX Interval	Indicates the minimum interval of receiving BFD packets supported by the local end.
Required Min Echo RX Interval	Indicates the minimum interval of receiving echo packets supported by the local end. It is set to 0 if the local end does not support the echo function.
Auth Type	(Optional) Indicates the authentication type, including: Simple Password Keyed MD5 Meticulous Keyed MD5 Keyed SHA1 Meticulous Keyed SHA1
Auth Length	Indicates the authentication data length.
Authentication Data	Indicates the authentication data area.

Session Status

A BFD session can be in any of the four basic states: Down, Init, Up, and AdminDown.

1. Down: Indicates that a session is in the Down state or is established just now.
2. Init: Indicates that the local system has communicated with the peer system and desires to bring the session to the Up state.
3. Up: Indicates that a session has been negotiated successfully.
4. AdminDown: Indicates that a session is in the AdminDown state.

BFD migrates the state machine based on the local session state and received BFD packets from the peer end.

A BFD state machine is established and torn down using a three-way handshake mechanism, to ensure that both ends know the status change.

Transmission Interval and Detection Time

Both ends negotiate BFD parameters during the establishment of a BFD session, to determine the transmission interval and detection time.

After a BFD session is established, both ends can dynamically negotiate BFD parameters (for example, minimum transmission interval and minimum receiving interval). After protocols at both ends transmit relevant negotiation packets, they adopt the new transmission interval and detection time, without affecting the current state of the session.

Overview

Feature	Description
BFD Session Establishment	Establishes a BFD session.
BFD Session Detection	Rapidly detects a bidirectional forwarding path.
BFD Support for Applications	Rapidly advertises the BFD detection result.
BFD Protection	Protects BFD from attacks for stability.
BFD Flapping Dampening	Protects stability of associated applications in the case of line instability.

6.3.1 BFD Session Establishment

BFD detection starts from the establishment of a BFD session.

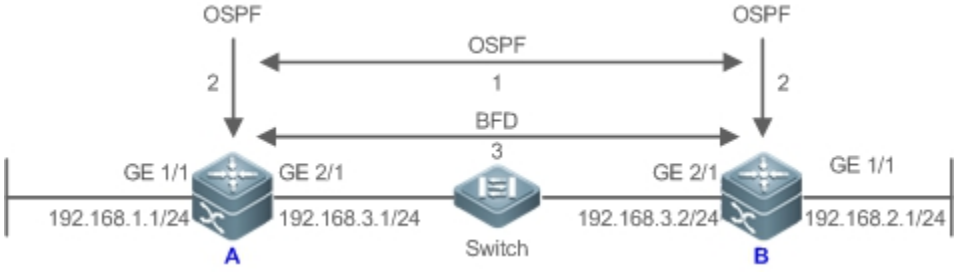
Working Principle

Session Establishment Process

BFD itself is unable to discover neighbors. It needs an upper-layer protocol to specify a neighbor to establish a session.

As shown in the following figure, two routers running OSPF and BFD are connected through a Layer-2 switch.

Figure 5- 4



BFD session establishment process:

1. OSPF discovers a neighbor and establishes a connection with the neighbor.
1. OSPF instructs BFD to establish a session with the neighbor.
2. BFD establishes a session with the neighbor.

BFD Session Establishment Mode

The BFD protocol specifies that a BFD session can be established in two modes:

- Active mode

Before the establishment of a session, BFD actively transmits a control packet for establishing a BFD session regardless of whether it receives a control packet for establishing a BFD session from the peer end.

- Passive mode

BFD does not actively transmit a control packet for establishing a BFD session before a session is established but wait till it receives a control packet for establishing a BFD session from the peer end.

The passive mode is not supported currently.

Negotiation of BFD Session Parameters

Both ends negotiate BFD session parameters during the establishment of a BFD session, to determine the transmission interval and detection time. Pay attention to the following points:

1. BFD session parameters (including **Desired Min Tx Interval**, **Required Min RX Interval**, and **Detect Mult**) must be set for interfaces at both ends. Otherwise, a BFD session cannot be established.

2. Interfaces at both ends negotiate BFD session parameters and detect the session based on the parameters during the establishment of a BFD session.

3. After a BFD session is established, both ends can dynamically negotiate BFD parameters (for example, minimum transmission interval and minimum receiving interval). After protocols at both ends transmit relevant negotiation packets, they adopt the new transmission interval and detection time, without affecting the current state of the session.

6.3.2 BFD Session Detection

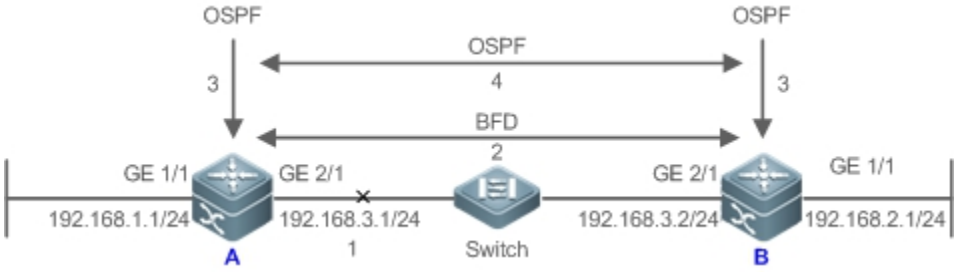
Link detection starts after the establishment of a BFD session. BFD periodically transmits BFD control packets. If it fails to receive BFD packets from the peer end within the detection time, it deems that the session is Down and notifies the associated application to accelerate the convergence.

Working Principle

Detection Process

As shown in the following figure, two routers running OSPF and BFD are connected through a Layer-2 switch.

Figure 5- 5



Handling procedure after a BFD session is Down:

1. The link between Router A and Switch fails.

2. The BFD session between Router A and Router B is Down.

3. BFD notifies the local OSPF that the forwarding path to the neighbor is faulty.

4. OSPF processes the neighbor Down situation. If a backup forwarding path is available, it starts protocol convergence to enable the alternative forwarding path.

Detection Mode

BFD supports the following detection modes:

- Asynchronous mode

In asynchronous mode, systems transmit BFD control packets periodically to each other. If a system fails to receive BFD control packets from the peer end within the detection time, it advertises that the session is Down.

- Query mode

In query mode, it is assumed that each system has an independent method for confirming its connection with other systems. After a BFD session is established, the system stops transmitting BFD control packets unless it needs to explicitly verify the connectivity. In such a case, the system transmits a shot-sequence BFD control packet. If a system fails to receive a returned packet within the detection time, it advertises that the session is Down. If it receives a response from the peer end, the forwarding path is reachable.

- Echo mode

In echo mode, the local system periodically transmits BFD echo packets and a remote system receives and loops back the packets through the forwarding path. If the local system fails to receive several consecutive echo packets within the detection time, it advertises that the session is Down. The echo function can be used together the preceding two detection modes. The echo packet detection function does not require the involvement of the control plane of the remote system. Packets are returned by the forwarding plane of the remote system, which reduces the delay and ensures faster fault detection in comparison with transmission of control packets. The enabling of the echo function in asynchronous mode can greatly reduce transmission of control packets because the detection is accomplished by the echo function. The enabling of the echo function in query mode can thoroughly cancel transmission of control packets after a session is established. The echo function must be enabled at both ends of a BFD session. Otherwise, the echo function does not take effect.

- The query mode is not supported and cannot be configured at present.
- Only BFD session version 1 supports the BFD echo mode.
- The echo mode is not supported for an IPv6 BFD session with the link-local address as the source or destination address.

6.3.3 BFD Support for Applications

By BFD support, the associated applications can utilize the fast fault detection of BFD to improve the protocol convergence performance. In general, the fault detection time can be shortened within 1 second.

Working Principle

After BFD support for a certain application is enabled, a BFD session is established based on the BFD configuration. When a link fault occurs, BFD can rapidly identify the fault and notify the associated application to process, thereby improving its convergence. Currently, BFD supports the following applications:

- BFD support for RIP

After BFD support for the Routing Information Protocol (RIP) is enabled, RIP can utilize the BFD fault detection, which is faster than the ND mechanism of RIP, to improve the protocol convergence. In general, the fault detection time can be shortened within 1 second.

- For more details about BFD support for RIP, see *Configuring RIP*.

- BFD support for OSPF

After BFD support for OSPF is enabled, OSPF can utilize the BFD fault detection, which is faster than the ND mechanism of OSPF, to improve the protocol convergence. In general, the fault detection time can be shortened within 1 second.

- For more details about BFD support for OSPF, see *Configuring OSPF*.

- BFD support for OSPFv3

After BFD support for OSPFv3 is enabled, OSPFv3 can utilize the BFD fault detection, which is faster than the ND mechanism of OSPFv3, to improve the protocol convergence. In general, the fault detection time can be shortened within 1 second.

For more details about BFD support for OSPFv3, see *Configuring OSPFv3*.

- BFD support for BGP

After BFD support for the Border Gateway Protocol (BGP) is enabled, BGP can utilize the BFD fault detection, which is faster than the ND mechanism of BGP, to improve the protocol convergence. In general, the fault detection time can be shortened within 1 second.

For more details about BFD support for BGP, see *Configuring BGP*.

- BFD support for IS-IS

The Intermediate System to Intermediate System (IS-IS) protocol dynamically discovers a neighbor by using hello packets. After BFD is enabled, IS-IS uses BFD to establish a BFD session with a neighbor that is in the Up state and detect the neighbor status. When a BFD neighbor fails, IS-IS immediately performs network convergence. The convergence time can be shortened from 30 seconds (by default, on a point-to-point network, IS-IS hello packets are transmitted at an interval of 10 seconds and the neighbor failure time is triple the interval, that is, 30 seconds) to 1 second.

For more details about BFD support for IS-IS, see *Configuring IS-IS*.

- BFD support for static routing

After BFD support for static routing is enabled, BFD prevents routers from selecting an unavailable static route as the forwarding path during routing and enables routers to rapidly switch to an available backup forwarding path.

Different from dynamic routing protocols, static routing does not have the ND mechanism. Therefore, after BFD support for static routing is configured, the next-hop reachability of a static route relies on the BFD session state. If a BFD session detects a fault, the next hop of a static route is unreachable and the static route is not added to the RIB.

If the remote system deletes a BFD session during the establishment of a BFD session, the BFD session becomes Down. In this case, the system ensures that the forwarding behavior of static routing is not affected.

For more details about BFD support for static routing, see *Configuring NSM*.

- BFD support for PBR

After BFD support for PBR is configured, BFD prevents routers from selecting an unavailable policy route as the forwarding path during routing and enables routers to rapidly switch to an available backup forwarding path.

BFD support for PBR is equivalent to that for static routing. BFD tracks and detects the forwarding path to a specified neighbor. When a BFD session fails, BFD notifies the PBR that the next hop is unreachable. Then, the policy route to the next hop does not take effect.

If the remote system deletes a BFD session during the establishment of a BFD session, the BFD session becomes Down. In this case, the system ensures that the PBR forwarding behavior is not affected.

For more details about BFD support for PBR, see *Configuring PBR*.

- BFD support for VRRP

The BFD support for the Virtual Router Redundancy Protocol (VRRP) can replace the ND mechanism of VRRP to rapidly detect the running status of the active and standby routers. When a fault occurs, it accelerates the active/standby router switching and improves network performance. In general, the fault detection time can be shortened within 1 second.

VRRP can also utilize BFD to track a specified neighbor. If the forwarding path to the neighbor fails during a BFD session, it automatically lowers the VRRP priority to a certain extent to trigger active/standby router switching. This configuration takes effect only when the dynamic routing protocol or other applications notify BFD to establish a session with a neighbor.

For more details about BFD support for VRRP, see *Configuring VRRP*.

- BFD support for VRRP Plus

The BFD support for VRRP Plus can replace the BVF detection conducted by the balancing virtual gateway (BVG) of VRRP Plus to rapidly detect the running status of balancing virtual functions (BVFs). When a fault occurs, it accelerates the forwarding entity switching and improves network performance. In general, the fault detection time can be shortened within 1 second.

VRRP Plus is based on the VRRP protocol. Therefore, no additional configuration is required for BFD support and only VRRP needs to be enabled on devices at both ends and a BFD session is correctly associated.

For more details about BFD support for VRRP Plus, see *Configuring VRRP Plus*.

- BFD support for Layer-3 interfaces

BFD supports changing status of Layer-3 interfaces. In interface configuration mode, use the **bfd bind peer-ip** command to detect the direct address of a specified Layer-3 interface. After this CLI command is executed, a BFD session is created and the status of a Layer-3 interface can be changed based on the detection result of the BFD session, for example, BFD Down or BFD Up. This function is often used in various types of fast reroute (FRR), which uses BFD to detect the interface status to implement fast FRR switching.

Only LDP FRR switching is supported in BFD support for Layer-3 interfaces.

- BFD support for AP member ports

After BFD support for AP member ports is enabled, BFD can rapidly detect a fault occurring on a member port link so that traffic on this link is rapidly distributed to other effective member links. In general, the fault detection time can be shortened within 1 second.

For more details about BFD support for AP member ports, see *Configuring AP*.

6.3.4 BFD Protection

The BFD protection is used to protect BFD against session flapping caused by attacks (for example, a large number of ping packets attack devices).

Working Principle

The BFD protocol is very sensitive. If a BFD-enabled device is attacked (for example, attacked by a large number of ping packets) and BFD sessions flap, the BFD protection can be configured to provide protection. If both BFD and BFD protection are enabled on a device, the device discards the BFD packet from the previous hop, affecting the establishment of a BFD session between the previous-hop device and other devices.

6.3.5 BFD Flapping Dampening

A BFD session may frequently switch over between Down and Up due to link instability. As a result, an associated application (such as static routing) may frequently switch forwarding paths and the running services are affected. The BFD flapping dampening can solve this problem.

Working Principle

A BFD session may frequently switch over between Down and Up. This function allows users to set the delay for status change advertisement. After a BFD session is Up for a certain period of time, BFD notifies an associated application of BFD Up. Otherwise, BFD notifies an associated application of BFD Down.

6.4 Configuration

Configuration	Description and Command	
Configuring BFD Basic Functions	(Mandatory) It is used to establish a BFD session.	
	bfd interval	Sets BFD parameters.
	N/A	Configures the BFD support for applications. The configuration command varies with the associated applications. For details, see their configuration guides.
	(Optional) It is used to configure the BFD detection mode, slow timer, and BDF support forLayer-3 interfaces.	
	bfd echo	Configures the BFD echo mode.
	bfd slow-timer	Configures the BFD slow timer.
	bfd bind peer-ip	Configures the BFD support for Layer-3 interfaces.
Configuring BFD Protection	(Optional) It is used to protect BFD against attacks.	
	bfd cpp	Enables BFD protection.
Configuring BFD Flapping Dampening	(Optional) It is used to protect associated protocols against BFD flapping.	
	bfd up-dampening	Configures BFD flapping dampening.

6.4.1 Configuring BFD Basic Functions

Configuration Effect

- Configure BFD support for applications.
- Establish a BFD session.
- A BFD session detects link faults.

Notes

- Pay attention to the following points when setting BFD session parameters:

1. It is recommended that parameter settings be consistent at both ends of a BFD session, to ensure that application protocols associated with BFD take effect simultaneously and prevent occurrence of one-way forwarding due to different dampening time at both ends.

2. Take into account of transmission bandwidth differences of different interfaces when setting parameters. If the minimum transmission interval and minimum receiving interval are set to very small values, data transmission may be affected due to very large BFD bandwidth occupancy.

- Pay attention to the following points when configuring BFD support for applications:

1. Ensure that it is enabled on neighbors of a BFD session. Otherwise, a BFD session cannot be established. If a dynamic routing protocol or another application requires BFD to establish a session with a neighbor, the BFD session can also be established.

2. If the interface specified by a BFD session is different from the actual BFD packet outbound interface because of IP routing, or if the interface specified during BFD session creation is different from the actual BFD packet inbound interface, a BFD session cannot be established.

- Pay attention to the following points when configuring the BFD detection mode:

1. In the process that the forwarding plane of the peer device returns echo packets transmitted by the local end to the local end, the echo packets may be lost due to congestion of the peer device, causing a session detection failure. In this case, configure Quality of Service (QoS) policies to ensure that echo packets are processed preferentially or disable the echo function.

2. The echo detection function of BFD does not support multi-hop detection. Ensure that the echo function is disabled when configuring multi-hops.

3. The echo mode takes effect only after this mode is enabled at both ends of a BFD session.

4. Before enabling the echo mode of BFD, run the **no ip redirects** command on the neighbors of a BFD session to disable the function of ICMP packet redirection, and run the **no ip deny land** command to disable the Distributed Denial of Service (DDoS) function (prevent the Land-based attack).

Configuration Steps

Setting BFD Parameters

- Mandatory.

- BFD parameters need to be set at BFD session egresses of routers at both ends detected by BFD if no special requirements are raised.

- Take into account of transmission bandwidth differences of different interfaces when setting parameters. If the minimum transmission interval and minimum receiving interval are set to very small values, data transmission may be affected due to very large BFD bandwidth occupancy.

Command	**bfd interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *interval-multiplier*
Parameter Description	**interval** *milliseconds*: Indicates the minimum TX interval, with the unit of milliseconds. **min_rx** *milliseconds*: Indicates the minimum RX interval, with the unit of milliseconds. **multiplier** *interval-multiplier*: Indicates the detection timeout multiplier.
Defaults	No BFD session parameter is configured.
Command Mode	Interface configuration mode
Usage Guide	The fast forwarding function must be enabled before the BFD function is enabled on routers.

Enabling the BFD Echo Mode

- (Optional) Ports run in asynchronous mode by default. If a BFD session needs to run in echo mode, the echo mode needs to be configured.
- Complete the configuration on ports of switches or routers.
- A session runs in asynchronous mode as long as either of routers at both ends is configured to run in asynchronous mode. If routers at both ends are configured to run in echo mode by default, a BFD session finally runs in echo mode.

Command	**bfd echo**
Parameter Description	N/A
Defaults	The BFD echo mode is disabled.
Command Mode	Interface configuration mode
Usage Guide	This command cannot be configured on AP ports. By default, when BFD session parameters are set, the system automatically enables the echo mode. The minimum TX interval and minimum RX interval of echo packets adopt the **Interval milliseconds** and **min_rx milliseconds** parameters of a session. Before enabling the echo mode of BFD, run the **no ip redirects** command on the neighbors of a BFD session to disable the function of ICMP packet redirection, and run the **no ip deny land** command to disable the Distributed Denial of Service (DDoS) function (prevent the Land-based attack).

Configuring the BFD Slow Timer

- (Optional) The default slow timer is 3,000 milliseconds. The value can be changed as required.
- Configure this function in global configuration mode of switches or routers.
- In BFD echo mode or session building, the slow timer is used to control packets. If the value increases, the required time for negotiating and establishing a BFD session becomes longer, and the time required for transmitting slow BFD packets in echo mode is longer.

Command	**bfd slow-timer** [*milliseconds*]
Parameter Description	*milliseconds***:** Indicates the BFD slow timer, with the unit of milliseconds. The value ranges from 1,000 to 30,000 and the default value 2,000 is adopted if it is not set.
Defaults	The transmission interval of slow control packets is 2,000 milliseconds.
Command Mode	Global configuration mode
Usage Guide	This command is used to specify the slow timer in echo mode.

Configuring the BFD Support for Layer-3 Interfaces

- (Optional) Currently, this function is used only when MPLS LDP is used for FRR.
- Configure this function on interfaces of switches or routers.

Command	**bfd bind peer-ip** *src-address* [**source-ip** *dst-address*] **process-pst**
Parameter Description	*src-address*: Indicates the peer IP address of an interface. *dst-address*: Indicates the local IP address of an interface.

Defaults	BFD support for Layer-3 interfaces is not configured by default.
Command Mode	Interface configuration mode
Usage Guide	This command is used to enable BFD support for Layer-3 interfaces so as to rapidly detect connectivity of Layer-3 interfaces.

Configuring the BFD Support for Applications

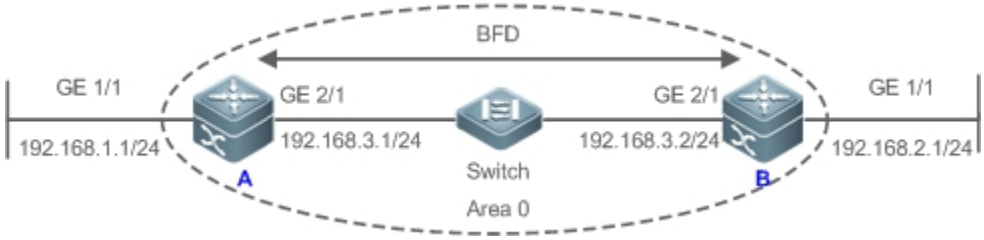
- Mandatory.
- This function is disabled by default.
- The configuration command varies with the associated applications. For details, see their configuration guides.
- This function must be configured at both ends so that a BFD session can be established.
- In RIP routing configuration mode, run the **bfd all interfaces** command to enable BFD support for RIP on all interfaces. For details, see *Configuring RIP*.
- In OSPF routing configuration mode, run the **bfd all interfaces** command to enable BFD support for OSPF on all interfaces. For details, see *Configuring OSPF*.
- In OSPFv3 routing configuration mode, run the **bfd all interfaces** command to enable BFD support for OSPFv3 on all interfaces. For details, see *Configuring OSPFv3*.
- In BGP routing configuration mode, run the **neighbor** *address* **fall-over bfd** command to enable BFD support for BGP. For details, see *Configuring BGP*.
- In IS-IS routing configuration mode, run the **bfd all interfaces** command to enable BFD support for IS-IS on all interfaces. For details, see *Configuring IS-IS*.
- In global configuration mode, run the **ip route static bfd** [**vrf** *vrf-name] interface-type interface-number gateway [* ***source*** *ip-address]* command to enable BFD support for static routing. For details, see *Configuring NSM*.
- In global configuration mode, run the **ipv6 route static bfd** [**vrf** *vrf-name*] *interface-type interface-number gateway* [**source** *ipv6-address*] command to enable BFD support for IPv6 static routing. For details, see *Configuring NSM*.
- Run the **set ip next-hop verify-availability** *next-hop-address* **bfd** [**vrf** *vrf-name*] *interface-type interface-number gateway* command to enable BFD support for PBR. For details, see *Configuring PBR*.
- Run the **set ipv6 next-hop verify-availability** *next-hop-address* **bfd** [**vrf** *vrf-name*] *interface-type interface-number gateway* command to enable BFD support for IPv6 PBR. For details, see *Configuring PBR*.
- Run the **vrrp bfd** *interface-type interface-number ip-address* command to enable BFD support for VRRP. For details, see *Configuring VRRP*.
- VRRP Plus is based on the VRRP protocol. Therefore, no additional configuration is required for BFD support for VRRP Plus. Only VRRP needs to be enabled on devices at both ends and a BFD session is correctly associated.

Verification

- The verification command varies with the associated applications. For details, see their configuration guides.

Configuration Example

Configuring BFD support for OSPF

Scenario **Figure 5- 6**	
Configuration Steps	● Configure IP addresses for interconnected interfaces of Router A and Router B. ● Run OSPF on Router A and Router B. ● Set BFD parameters for interconnected interfaces of Router A and Router B. ● Enable BFD support for OSPF on Router A and Router B.

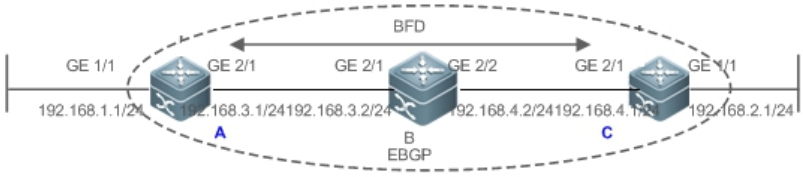
A	A#configure terminal
	A(config)#interface GigabitEthernet2/1
	A(config-if-GigabitEthernet2/1)# no switchport //The configuration is not required on routers.
	A(config-if-GigabitEthernet2/1)#ip address 192.168.3.1 255.255.255.0
	A(config-if-GigabitEthernet2/1)#bfd interval 200 min_rx 200 multiplier 5
	A(config-if-GigabitEthernet2/1)# exit
	A(config)#interface GigabitEthernet1/1
	A(config-if-GigabitEthernet1/1)# no switchport //The configuration is not required on routers.
	A(config-if-GigabitEthernet1/1)#ip address 192.168.1.1 255.255.255.0
	A(config-if-GigabitEthernet1/1)# exit
	A(config)# router ospf 123
	A(config-router)# log-adj-changes detail
	A(config-router)# network 192.168.3.00.0.0.255 area 0
	A(config-router)# network 192.168.1.00.0.0.255 area 0
	A(config-router)# bfd all-interfaces
	A(config-router)# end

B	B#configure terminal
	B(config)#interface GigabitEthernet2/1
	B(config-if-GigabitEthernet2/1)# no switchport //The configuration is not required on routers.
	B(config-if-GigabitEthernet2/1)#ip address 192.168.3.2 255.255.255.0
	B(config-if-GigabitEthernet2/1)#bfd interval 200 min_rx 200 multiplier 5
	B(config-if-GigabitEthernet2/1)# exit
	B(config)#interface GigabitEthernet1/1
	B(config-if-GigabitEthernet1/1)# no switchport //The configuration is not required on routers.

	B(config-if-GigabitEthernet1/1)#ip address 192.168.2.1 255.255.255.0 B(config-if-GigabitEthernet1/1)# exit B(config)# router ospf 123 B(config-router)# log-adj-changes detail B(config-router)# network 192.168.3.00.0.0.255 area 0 B(config-router)# network 192.168.2.00.0.0.255 area 0 B(config-router)# bfd all-interfaces B(config-router)# end
Verification	Display verification.
A	A# show bfd neighbors details OurAddr NeighAddr LD/RD RH/RS Holdown(mult) State Int 192.168.3.1 192.168.3.2 1/2 Up 532 (3) Up Ge2/1 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 200000, MinRxInt: 200000, Multiplier: 5 Received MinRxInt: 50000, Received Multiplier: 3 Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332 Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196 Registered protocols: OSPF Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 3 - Length: 24 My Discr.: 2 - Your Discr.: 1 Min tx interval: 50000 - Min rx interval: 50000 Min Echo interval: 0
B	B# show bfd neighbors details OurAddr NeighAddr LD/RD RH/RS Holdown(mult) State Int 192.168.3.2 192.168.3.1 2/1 Up 532 (5) Up Ge2/1 Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 50000, MinRxInt: 50000, Multiplier: 3 Received MinRxInt: 200000, Received Multiplier: 5

	Holdown (hits): 600(22), Hello (hits): 200(84453) Rx Count: 49824, Rx Interval (ms) min/max/avg: 209/440/332 last: 66 ms ago Tx Count: 84488, Tx Interval (ms) min/max/avg: 153/249/197 last: 190 ms ago Registered protocols: OSPF Uptime: 02:18:49 Last packet: Version: 1 - Diagnostic: 0 I Hear You bit: 1 - Demand bit: 0 Poll bit: 0 - Final bit: 0 Multiplier: 5 - Length: 24 My Discr.: 1 - Your Discr.: 2 Min tx interval: 200000 - Min rx interval: 200000 Min Echo interval: 0

Configuring BFD Support for BGP

Scenario **Figure 5-7**	
Configuration Steps	● Configure IP addresses for ports connecting Router A and Router B. ● Run the OSPF protocol on Router A and Router B. ● Configure IP addresses for ports connecting Router B and Router C. ● Run the OSPF protocol on Router B and Router C. ● Configure EBGP on Router A and Router C, and enable BFD support for BGP.
A	A# configure terminal A(config)# interface GigabitEthernet2/1 A(config)# bfd multi-hop interval 200 min_rx 200 multiplier 5 A(config-if-GigabitEthernet2/1)# no switchport //This command is required on the switch A(config-if-GigabitEthernet2/1)# ip address 192.168.3.1 255.255.255.0 A(config-if-GigabitEthernet2/1)# exit A(config)# interface GigabitEthernet1/1 A(config-if-GigabitEthernet1/1)# no switchport //This command is required on the switch A(config-if-GigabitEthernet1/1)# ip address 192.168.1.1 255.255.255.0 A(config-if-GigabitEthernet1/1)# exit

	A(config)# router ospf 123 A(config-router)# log-adj-changes detail A(config-router)# network 192.168.3.0 0.0.0.255 area 0 A(config-router)# network 192.168.1.0 0.0.0.255 area 0 A(config-router)# exit A(config)#router bgp 100 A(config-router)# neighbor 192.168.4.1 remote-as 200 A(config-router)# neighbor 192.168.4.1 ebgp-multihop 3 A(config-router)# neighbor 192.168.4.1 update-source 192.168.3.1 A(config-router)# neighbor 192.168.4.1 fall-over bfd A(config-router)# end
B	B# configure terminal B(config)# interface GigabitEthernet2/1 B(config-if-GigabitEthernet2/1)# no switchport //This command is required on the switch B(config-if-GigabitEthernet2/1)# ip address 192.168.3.2 255.255.255.0 B(config-if-GigabitEthernet2/1)# exit B(config)# interface GigabitEthernet2/2 B(config-if-GigabitEthernet2/2)# no switchport //This command is required on the switch B(config-if-GigabitEthernet2/2)# ip address 192.168.4.2 255.255.255.0 B(config-if-GigabitEthernet2/2)# exit B(config)# router ospf 123 B(config-router)# log-adj-changes detail B(config-router)# network 192.168.3.0 0.0.0.255 area 0 B(config-router)# network 192.168.4.0 0.0.0.255 area 0 B(config-router)# end
C	C# configure terminal C(config)# interface GigabitEthernet2/1 C(config)# bfd multi-hop interval 200 min_rx 200 multiplier 5 C(config-if-GigabitEthernet2/1)# no switchport //This command is required on the switch C(config-if-GigabitEthernet2/1)# ip address 192.168.4.1 255.255.255.0 C(config-if-GigabitEthernet2/1)# exit C(config)# interface GigabitEthernet1/1 C(config-if-GigabitEthernet1/1)# no switchport //This command is required on the switch C(config-if-GigabitEthernet1/1)# ip address 192.168.2.1 255.255.255.0 C(config-if-GigabitEthernet1/1)# exit C(config)# router ospf 123 C(config-router)# log-adj-changes detail C(config-router)# network 192.168.4.0 0.0.0.255 area 0 C(config-router)# network 192.168.2.0 0.0.0.255 area 0 C(config-router)# exit

	C(config)#router bgp 200 C(config-router)# neighbor 192.168.3.1 remote-as 200 C(config-router)# neighbor 192.168.3.1 ebgp-multihop 3 C(config-router)# neighbor 192.168.3.1 update-source 192.168.4.1 C(config-router)# neighbor 192.168.3.1 fall-over bfd C(config-router)# end
Verification	Check whether the configurations take effect.
A	A# show bfd neighbors details OurAddr NeighAddr LD/RD RH/RS Holdown(mult) State Int 192.168.3.1 192.168.4.1 8192/8192 Up 65 (5) Up Ge2/1 Session state is Up and not using echo function. Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 2000000, MinRxInt: 2000000, Multiplier: 5 Received MinRxInt 2000000, Multiplier: 5 Holdown (hits): 10000(1), Hello (hits): 2000(150) Rx Count: 31, Rx Interval (ms) min/max/avg: 0/0/2000 Tx Count: 206, Tx Interval (ms) min/max/avg: 0/0/2000 Registered protocols: BGP Uptime: 0:00:50 Last packet: Version : 1 - Diagnostic : 3 State bit : Init - Demand bit : 0 Poll bit : 0 - Final bit : 0 Multplier : 5 - Length : 24 My Discr : 8192 - Your Discr : 8192 Min tx interval : 2000000 - Min rx interval: 2000000 Min Echo interval: 0
C	B# show bfd neighbors details OurAddr NeighAddr LD/RD RH/RS Holdown(mult) State Int 192.168.4.1 192.168.3.1 8192/8192 Up 65 (5) Up Ge2/1 Session state is Up and not using echo function. Local Diag: 0, Demand mode: 0, Poll bit: 0 MinTxInt: 2000000, MinRxInt: 2000000, Multiplier: 5 Received MinRxInt 2000000, Multiplier: 5 Holdown (hits): 10000(0), Hello (hits): 2000(794) Rx Count: 5280, Rx Interval (ms) min/max/avg: 0/0/2000 Tx Count: 2470, Tx Interval (ms) min/max/avg: 0/0/2000 Registered protocols: BGP Uptime: 0:01:13 Last packet: Version : 1 - Diagnostic : 0 State bit : Up - Demand bit : 0

	Poll bit : 0 - Final bit : 0 Multplier : 5 - Length : 24 My Discr : 8192 - Your Discr : 8192 Min tx interval : 2000000 - Min rx interval: 2000000 Min Echo interval: 0

Common Errors

- BFD parameters are not set for device interfaces at one end.
- The BFD support for applications is disabled.
- The BFD support for applications is enabled only at one end.

6.4.2 Configuring BFD Protection

Configuration Effect

- If a BFD-enabled device is attacked (for example, attacked by a large number of ping packets) and BFD session flaps accordingly, the BFD protection can be enabled to provide protection.

Notes

- The BFD basic functions must be configured.
- If both BFD and BFD protection are enabled on a device, the device discards the BFD packet from the previous hop, affecting the establishment of a BFD session between the previous-hop device and other devices.
- This function and limitations are applicable only to switches.

Configuration Steps

Enabling BFD Protection

- Optional.
- Configure this function in global configuration mode on switches or routers.
- The BFD protection function raises the processing priority of BFD packets and ensures normal running of BFD services in a scenario in which devices are attacked.

Command	**bfd cpp**
Parameter Description	N/A
Defaults	The BFD protection function is enabled by default.
Command Mode	Global configuration mode
Usage Guide	Enable the BFD protection function to provide protection if a device encounters BFD flapping due to attacks.

Verification

Run the **show running-config** command to verify the configuration on an interface.

Configuration Example

Enabling BFD Protection

Configuration Steps	● Configure this function on a switch on a network where attacks exist. ● Configure the BFD protection function.
	FS#configure terminal FS(config)# bfd cpp FS(config)# end
Verification	N/A

6.4.3 Configuring BFD Flapping Dampening

Configuration Effect

- A BFD session may frequently switch over between Down and Up due to link instability. As a result, a relevant application (such as static routing) may frequently switch forwarding paths and the running services are affected.
- Users can set the delay for status change advertisement, after which BFD notifies an associated application of BFD Up. After a BFD session is Up for a certain period of time, BFD notifies an associated application of BFD Up. Otherwise, BFD notifies it of BFD Down. The purpose is to reduce flapping of associated protocols caused by instable links.

Notes

- The BFD basic functions must be configured.
- If a BFD session does not frequently switch over between Down and Up, the enabling of BFD flapping dampening will delay notifying an associated application of BFD Up.

Configuration Steps

Configuring BFD Flapping Dampening

- (Optional) The BFD flapping dampening is disabled on ports by default. If a BFD session frequently switches over between Down and Up, it is advised to enable this function.
- Configure this function on ports of switches or routers.
- With BFD flapping dampening enabled, it is relieved that associated applications, such as route re-calculation, process quantities of advertisements because of frequent status BFD change. The larger the configured time is, the longer the required BFD stability time is. BFD notifies an application module of BFD Up only after the stability time reaches the configured time.

Command	**bfd up-dampening**[*milliseconds*]
Parameter Description	*milliseconds*: Indicates the delay for status change advertisement, after which BFD notifies an associated application of BFD Up, with the unit of milliseconds. The value ranges from 0 to 300,000. The value 0 indicates that BFD notifies the application layer immediately when a session switches over from Down to Up and the default value is 0.
Defaults	The BFD flapping dampening function is disabled by default.
Command Mode	Interface configuration mode
Usage Guide	This function needs to be enabled only when the link is instable.

	If a BFD session does not frequently switch over between Down and Up, the enabling of BFD flapping dampening will delay notifying an associated application of BFD Up.

Verification

Run the **show running-config** command to verify the configuration on an interface.

Configuration Example

Configuring BFD Flapping Dampening with the Advertisement Delay as 60,000 Milliseconds

Configuration Steps	● Configure this function in an environment where BFD frequent flaps due to link instability. ● Set the delay for status change advertisement to 60,000 milliseconds.
	FS#configure terminal FS(config)# interface fastEthernet 0/2 FS(config)# bfd up-dampening 60000 FS(config)# end
Verification	N/A

6.5 Monitoring

Displaying

Description	Command
Displays BFD session information.	**show bfd neighbors** [**vrf** *vrf-name*] [**client** {**ap** |**bgp**|**ospf**|**rip**|**vrrp**|**static-route** | **pbr**| **vrrp-balance** | **pst** }][**ipv4** *ip-address* | **ipv6** *ip-address*][**details**]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs BFD events.	**debug bfd event**[**interface** *interface-type interface-number* | **ipv4** *ip-address* | **ipv6** *ipv6-address*]
Debugs BFD packets.	**debug bfd packet**[**interface** *interface-type interface-number* | **ipv4** *ip-address* | **ipv6** *ipv6-address*]

7 Configuring IP Event Dampening

7.1 Overview

When the Layer-3 port on a Layer-3 device frequently goes Up and Down due to manual enabling/disabling or other external causes, the routing table on the device will flap repeatedly. If a routing protocol is configured, the protocol may propagate the flap to the entire network, causing repeated updates and recalculation of neighboring routes, which wastes network bandwidths and destabilizes the network. Repeated route updates and recalculation on devices consume many CPU resources, which affects the normal running of customer networks.

IP Event Dampening detects abnormal Up/Down flapping and automatically suppresses frequent port state changes, which prevents the propagation of single-point link failures by a routing protocol. When the port is restored, it will be automatically unsuppressed, thus reducing network flaps and CPU resource consumption while improving network stability.

Protocols and Standards

- RFC2439: BGP Route Flap Dampening

At its core, the suppression algorithm used by IP Event Dampening is the same as that used by BGP Route Flap Dampening.

7.2 Applications

Application	Description
Routed Port Flap Dampening	Monitors the state change of the Layer-3 port on a router, and suppresses frequent port flapping.

7.2.1 Routed Port Flap Dampening

Scenario

In a network that runs a routing protocol, when a port on a router connected to another router frequently goes Up and Down, neighboring routes will be repeatedly updated and recalculated. The routing protocol may propagate the flap to the entire network, causing a network flap. IP Event Dampening can be enabled on the connected routers to monitor port state changes and suppress frequent port flapping, thus reducing network flaps and CPU resource consumption while improving network stability.

Figure 6- 1

Remarks	A and B are routers.

Deployment

Configure IP Event Dampening on portGE0/1 on Router A and portGE0/1 on Router B respectively.

The subinterfaces and the virtual templates of interfaces on routers do not support the dampening feature.

7.3 Features

Basic Concepts

Penalty

A port that goes Up or Down gets a penalty for each state change, but the penalty decays exponentially when the port is stable. In this way, port behaviors can be sensed and controlled intelligently.

Suppress Threshold

When the cumulative penalty of a port exceeds a suppress threshold, the port is considered to flap and will be suppressed.

Half-Life Period

The half-life period is the period required for the penalty to decrease to half of the original value when the port is stable. It defines the speed at which the penalty decays exponentially. The shorter the half-life period, the faster the penalty decays, and the faster the port is detected to be stable, but the flap detection sensitivity is reduced.

Reuse Threshold

When the port no long flaps and its penalty decays to a certain degree (below the suppress threshold), the port is considered to be stable and is unsuppressed.

Maximum Suppress Time

When a port keeps flapping and reaches a very large penalty, the port will not be usable for a long time. To avoid this problem, the maximum suppress time is defined to always maintain the port suppression duration below a certain value no matter how long the port has flapped.

Overview

Feature	Description
Port Flap Suppression	Configure the criteria and parameters of flap suppression on ports to enable switches or routers to identify and suppress frequently flapping ports, which ensures route stability and avoids route flap propagation.

7.3.1 Port Flap Suppression

Working Principle

A port configured with IP Event Dampening is assigned a penalty. The port gets a penalty of 1,000 each time when it goes Down, but the penalty decreases with time. If the port goes Down again, the penalty increases accordingly. When the cumulative penalty exceeds the suppress threshold, the port will be suppressed. For the affected upper-layer protocol, the suppressed port is always Down no matter what the actual port state is. When the penalty decreases to the reuse threshold, the port will be unsuppressed, and the upper-layer protocol can sense the actual port state.

If a Layer-3 port is not configured with IP Event Dampening, or is not suppressed by it, the routing protocol or other protocol concerned about the port status still work normally. When the port is suppressed, the upper-layer protocol considers the port to be Down. Any state change of the port before the port is unsuppressed does not affect the routing table and the route calculation and advertisement performed by the upper-layer routing protocol.

Related Configuration

Configuring IP Event Dampening

- By default, IP Event Dampening is disabled on Layer-3 ports.
- Run the **dampening** [*half-life-period* [*reuse-threshold suppress-threshold max-suppress* [**restart** [*restart-penalty*]]]] command to enable or disable IP Event Dampening on Layer-3 ports.

7.4 Configuration

Configuration	Description and Command	
Enabling IP Event Dampening	(Mandatory)It is used to suppress Layer-3 port flapping.	
	dampening	Configures IP Event Dampening.

7.4.1 Enabling IP Event Dampening

Configuration Effect

When a port configured with IP Event Dampening keeps flapping until the predefined threshold is exceeded, the port is set to Down.

Notes

- When a Layer-3 port on a switch is converted to a Layer-2 port (for example, from a routed port to a switch port), the IP Event Dampening configuration on the port will be deleted.
- Only the main interface on a router can be configured with IP Event Dampening. The configuration takes effect for all subinterfaces of the main interface, but you cannot run the **dampening** command directly on subinterfaces and virtual templates.

Configuration Steps

Configuring IP Event Dampening

- Mandatory.
- Perform the configuration in Layer-3 interface configuration mode.
- You can specify the half-life period, reuse threshold, suppress threshold, maximum suppress time, and initial penalty. If you do not set these parameters, their default values will be used.

Verification

Use any one of the following commands to check whether the configuration takes effect:

- **show running-config**
- **show interfaces** [*interface-id*] **dampening**, which is used to check the IP Event Dampening configuration on a specified port

Related Commands

Enabling IP Event Dampening on a Port

Command	**dampening** [*half-life-period* [*reuse-threshold suppress-threshold max-suppress* [**restart** [*restart-penalty*]]]]
Parameter Description	*half-life-period*: Indicates the half-life period. Value range: <1–30>; default value: 5s. *reuse-threshold:* Indicates the reuse threshold. Value range: <1–20,000>; default value: 1,000. *suppress-threshold:* Indicates the suppress threshold. Value range: <1–20,000>; default value: 2,000.

	max-suppress: Indicates the maximum suppress time. Value range: <1–255>; default value: four times the half-life period. **restart** *restart-penalty:* Indicates the initial penalty. Value range: <1–20,000>; default value: 2,000.
Command Mode	Interface configuration mode
Usage Guide	IP Event Dampening can affect direct routes, host routes, static routes, dynamic routes, and VRRP. When a port is suppressed based on the configured criteria, the affected modules determine that the port is Down and therefore delete corresponding routes. No data packet will be transmitted through the port. When the **dampening** command is rerun on a port configured with IP Event Dampening, the dampening information on the port will be cleared, but the flap count is retained, unless you use the **clear counters** command to clear the counters on the port. If the **max-suppress** parameter is set to a very small value, making the maximum penalty smaller than the suppress threshold, the port will never be suppressed. When such a configuration error occurs, the following message indicating a configuration failure will be printed: % Maximum penalty (10) is less than suppress penalty (2000). Increase maximum suppress time If the available system memory is insufficient to run the **dampening** command, the following message indicating a configuration failure will be printed: % No memory, configure dampening fail!

Configuration Example

Configuring IP Event Dampening on Layer-3 Ports

Scenario **Figure 6- 2**	
Configuration Steps	Enable IP Event Dampening on port GigabitEthernet 0/1 on Router A and on port GigabitEthernet 0/1 on Router B respectively, and set **half-time-period** to 30s, **reuse-threshold** to 1,500, **suppress-threshold** to 10,000, and **max-suppress** to 120s.
A	FS(config)#interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#dampening 30 1500 10000 100
B	FS(config)#interface GigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)#dampening 30 1500 10000 100
Verification	Run the **show interfaces dampening** command to check the IP Event Dampening configuration on the corresponding ports.
	FS#show interfaces dampening GigabitEthernet 0/1 Flaps Penalty Supp ReuseTm HalfL ReuseV SuppV MaxSTm MaxP Restart

	0	0	FALSE	0	30	1500	1000	100	15119	0

Common Errors

- The port on a Layer-3 switch is not converted to a routed port by using the **no swithport** command before IP Event Dampening is configured.

7.5 Monitoring

Clearing

Description	Command
Clears the interface counters.	**clear counters**

For details about the **clear counter** command, see the related chapter for the "Interface" command.

Displaying

Description	Command
Displays the counters on suppressed ports.	**show dampening interface**
Displays the IP Event Dampening configuration on ports.	**show interfaces dampening**

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

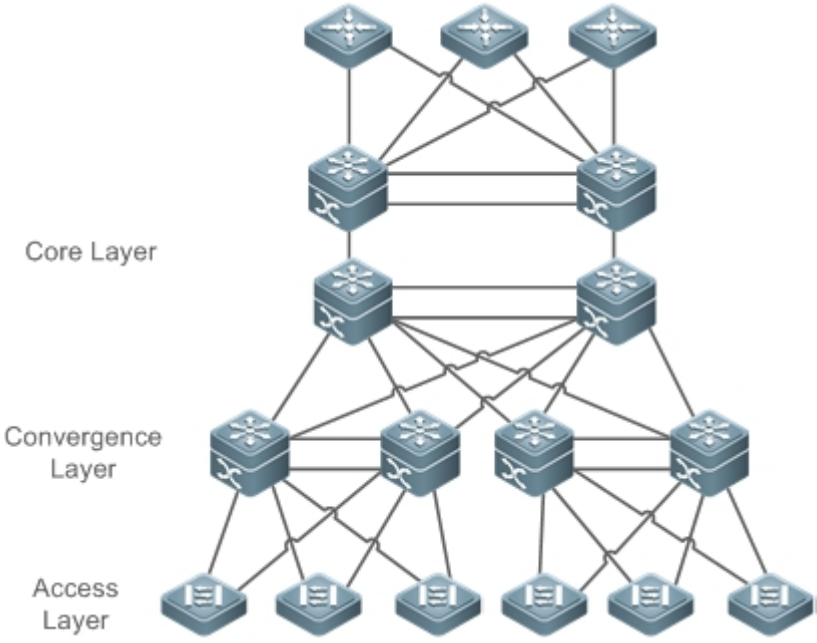
Description	Command
Enables debugging of IP Event Dampening.	**debug dampening interface**

8 Configuring Stacking

8.1 Overview

In order to improve the reliability of networks, the two devices at core layer and convergence layer of traditional networks are configured with two cores to provide redundancy. Access and convergence devices are respectively connected to the cores through two links. The following figure shows a typical traditional network architecture. Redundant network architecture increases the complexity of network design and operation. At the same time, a large number of redundant links reduce the utilization of network resources and return on investment.

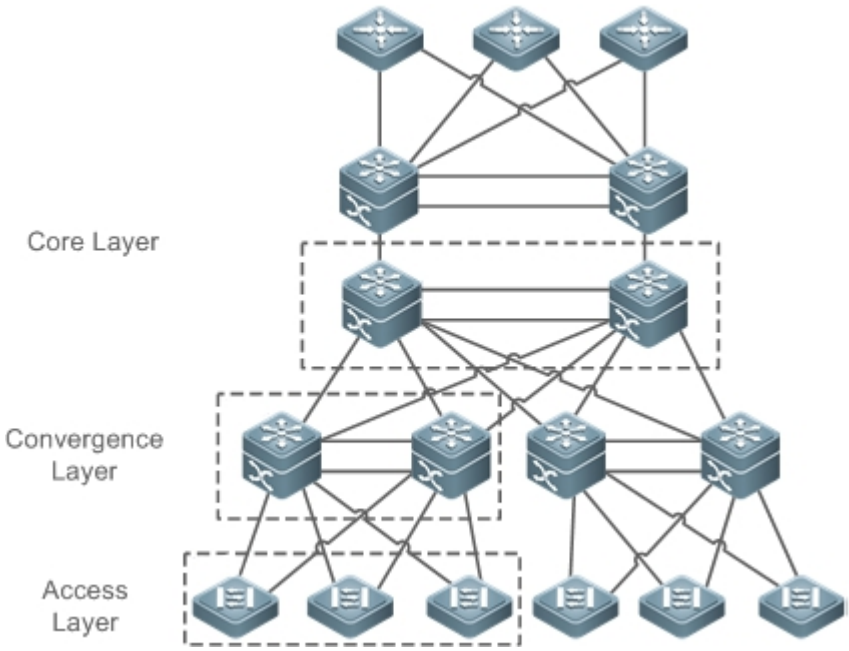
Figure 7- 1 Traditional Network Architecture



Stacking is a kind of network system virtualization technology that supports combining multiple devices into a single virtualized device. As shown in Figure 7-2, access, convergence and core layer devices can respectively form Stackings, and then these Stackings connect to one another to form an end-to-end Stacking network. Compared with traditional network, this networking can:

- Simplify the network topology.
- Reduce the costs of network management and maintenance.
- Shorten application recovery time and service interruption time.
- Enhance the utilization of network resources.

Figure 7-2 End-to-End Stacking Networking



8.2 Applications

Application	Description
Managing Multiple Devices in a Unified Manner	Uses multiple physical devices as a logical device for unified management.
Simplifying Networking Topology	Uses a Stacking as a logical device to simplify the networking topology.

8.2.1 Managing Multiple Devices in a Unified Manner

Scenario

When multiple physical devices form a Stacking system, the physical devices can be viewed as a logical device. All configurations are managed on the global master device.

As shown in Figure 7-3, four devices (numbered as 1, 2, 3, and 4 from left to right) form a Stacking system. Device 1 is the global master device, device 2 is the global slave device, and devices 3 and 4 are the global candidate devices.

- All devices are configured simply on the global master device.

Figure 7- 3

Remarks	The devices from left to right in Figure 7-3 are Device 1, Device 2, Device 3 and Device 4. For details on VSL, see the description in section 1.3.1. Device 1 is the global master device. Device 2 is the global slave device. Devices 3 and 4 are the global candidate devices.

Deployment

- The global master device controls the entire Stacking system, runs control-plane protocols and is involved in data forwarding.
- The global slave device is involved in data forwarding, does not run control-plane protocols, and works as the backup and takes over the work of the global master device when faulty.
- The global candidate devices are involved in data forwarding and do not run control-plane protocols. When the global slave device is faulty, a global candidate device can take over the work of the global slave device. In this case, when the global master and slave devices are faulty, the Stacking system will restart.

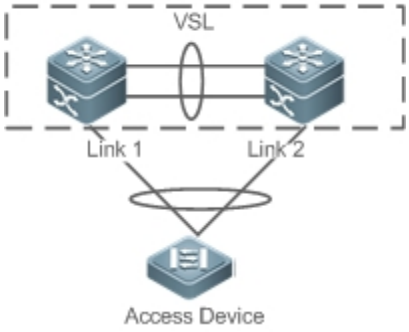
8.2.2 Simplifying Networking Topology

Scenario

In traditional networks as shown in Figure 7-4, redundant devices and lines need to be added to increase the networking reliability; however, many algorithms also need to be introduced to prevent loops, which make the networking more complex. In the Stacking system, all devices are viewed as a logical device. Different devices back up each other, and no loop prevention algorithm needs to be introduced, which can simplify the network.

- Two aggregate switches form a Stacking system. It is unnecessary to configure a loop prevention algorithm. The two switches are redundant mutually.
- The access switch is connected to the aggregate switches through the uplink AP.
- When a switch in the Stacking system is faulty, the other link still works.

Figure 7- 4



Deployment

- The global master device controls the entire Stacking system, runs control-plane protocols and is involved in data forwarding.
- The global slave device is involved in data forwarding, does not run control-plane protocols, and works as the backup and takes over the work of the global master device when the global master device is faulty.
- The access switch is oriented to users and allows access by users' devices.

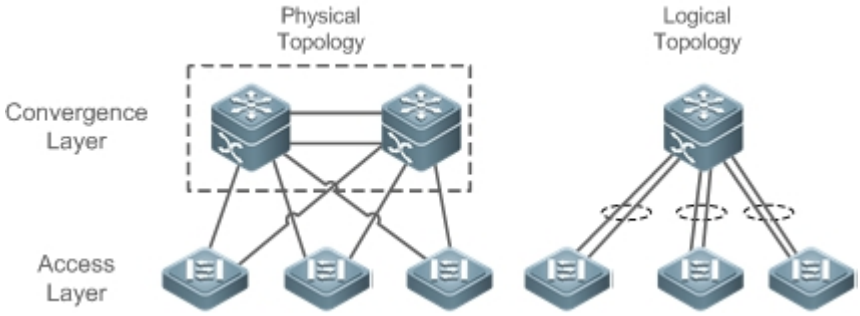
8.3 Features

Basic Concepts

Stacking System

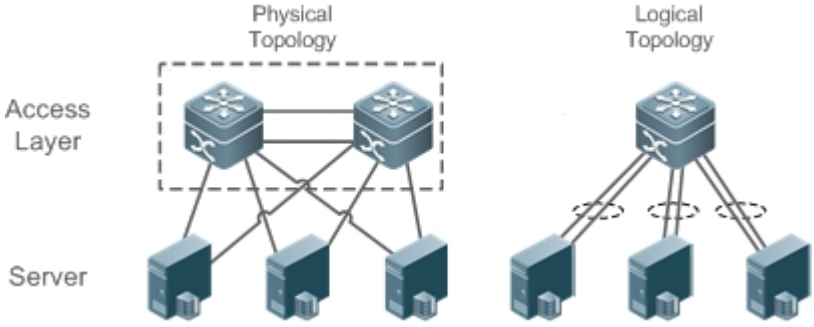
Stacking system is a single logical entity consisting of two or multiple devices in traditional network architecture. For example, the convergence layer Stacking system as shown in the following figure can be seen as a single device that interacts with the core layer and access layer.

Figure 7- 5 Convergence layer Stacking



In the above Stacking network structure, the member devices form a logical entity through internal links and the access layer devices are connected to the Stacking through aggregated links. In this way, there is no layer 2 loop between the access and convergence layers.

Figure 7- 6 Access layer Stacking



Except the core and convergence layer devices, the access layer devices can also form a Stacking system. A server that requires high availability can adopt multiple network cards to form an Aggregate Port (AP) to connect access layer devices. Since AP can only connect to the same access device, the risk of single device fault increases. In this case, Stacking can be used to solve the problem. In the Stacking mode, a server adopts multiple network cards and binds them into an AP to connect different member devices in the same Stacking group. This way can prevent single point failure and network interruption caused by single link failure.

Stacking Domain ID

A Stacking domain has only one ID. Only the devices with the same domain IDs can form a Stacking system.

Member Device ID

Every member device in a Stacking system has a unique ID, namely, Switch ID. Switch IDs can be used in device management or configuring interfaces on member devices. You need to configure an ID for a device when adding the device to a Stacking system and ensure that the ID is unique in the same Stacking system. If an ID conflict occurs, the Stacking system will reserve one device according priority.

Member Device Role

A Stacking system consists of several devices. When establishing a Stacking system, you need to select a global master device and a global slave device. All other devices are global candidate devices. A global master device is elected from multiple devices based on an election protocol. All other devices are global slave devices in the 1: N hot standby mode. When the 1:1 hot standby mode is

supported, one device is the global master device, one device is the global slave device, and all other devices are global candidate devices.

The global master device is responsible for controlling the entire Stacking system, running control plane protocols and participating in data forwarding. Other devices, including the global slave devices and candidate devices, participate in data forwarding but do not run control plane protocols. All received control plane data flows are forwarded to the global master device for processing.

The global slave device also receives the statuses of the global master device in real-time and provide 1:1 or 1:N redundancy with the global master device. If the global master device becomes faulty, the global slave device will take over services from the master device and manage the entire Stacking system.

The following is the method for selecting the master device of a Stacking system:

1. Rules for selecting the master device of a Stacking system include (Continue with the next rule if the previous rule does not help in selecting the master device): a) Select the currently running host as the master device with the highest priority (All devices are not master devices during startup). b) Select the device with the highest priority as the master device. c) Select the device with the lowest device No. as the host. d) Select the device with the smallest MAC address as the master device.
2. In the 1:N hot standby mode, select the device that has the most familiar configurations with the master device as the slave device to prevent dual active devices. The selection order is: the nearest/the highest priority/the smallest MAC address.
3. Stacking system supports hot adding a support device. Even the hot added device has a higher priority than the master device has, the Stacking system does not perform active/standby switch.
4. The startup order of member device may affect the election of master device. A member device may not join in the Stacking system because it starts up too slowly. In this case, the device will be hot added to the Stacking system. Even the device has a higher priority than the master device, the Stacking system does not perform active/standby switchover.

Overview

Feature	Description
Virtual Switching Link (VSL)	In a Stacking system, a virtual link is used to connect all devices.
Topology	Describes the internal topology of a Stacking system.
Dual-Active Detection (DAD)	Avoids that dual master switches coexist in a Stacking domain.
System Management	Describes possible connections between external devices and Stacking devices.
Quick Blinking Location	Manages devices in the Stacking system.

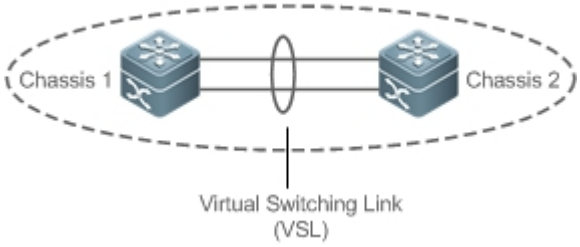
8.3.1 Virtual Switching Link (VSL)

Working Principle

VSL

The Stacking system is a network entity that consists of multiple devices. These devices need to share control information and part of data streams. The VSL is a special link used for transmission of control information and data streams among devices of the Stacking system. For example, the VSL can be established between two devices through 10 Gigabit Ethernet interfaces. Figure 7-7 shows the position of the VSL in the Stacking system.

Figure 7- 7 VSL



The VSL exists in the form of AP groups. The data streams transmitted through the VSL balance load among the aggregation port members according to the traffic balancing algorithm.

VSL Traffic

The control streams transmitted through the VSL between devices include:

5. The protocol packets received by the member devices: These protocol packets need to be forwarded through the VSL to the global master device for processing.
6. The protocol packets processed by the global master device: These protocol packets need to be forwarded through the VSL to the interfaces of other member devices and then sent to the peer devices by these interfaces.

The data streams transmitted through the VSL between devices include:

1. The data stream flooded on the VLAN
2. The data streams that need to be forwarded across devices and transmitted through the VSL

Furthermore, the internal management packets of the Stacking system are also transmitted through the VSL. The management packets include the protocol information switched by the hot backup and configuration information delivered by the host to other member devices.

In terms of the switched port analyzer (SPAN) function, the interface associated with the VSL cannot be regarded as the source port or destination port of the SPAN.

VSL Failure

If a certain member link connected to the VSL AP group fails to work, the Stacking will adjust the configurations of the VSL aggregation port automatically to prevent the traffic from being transmitted through the faulty member link.

If all member links are disconnected to the VSL AP group, the Stacking topology will change. If the original Stacking topology is a ring topology, the ring will convert into a line. For details, see topology ring and line conversion in the section of *Topology Changes*.

Detecting Error Frames on a VSL Interface

When a large number of consecutive error frames are detected on a VSL interface, the interface must be disabled and switched to another VSL interface. The detection method is as follows:

If error frames are found on a VSL interface, perform error frame correction. The system detects the VSL interface every 5 seconds by default. If the number of error frames is greater than the value of *num* as compared with that detected last time, it is assumed that error frames are detected once. If error frames are detected consecutively for the value of *times*, it is assumed that the interface is abnormal. If multiple VSL links are available when error frames are detected, the VSL will be switched. The last VSL will not be switched in order to prevent topology splitting.

Different user scenarios have different requirements for *num* and *times*. The default value of *num* is *3* and that of *times* is *10*. If users have strict requirements on the scenarios, select smaller values for *num* and *times*; if reverse, select greater values.

8.3.2 Topology

The Stacking system supports line topology and ring topology. Devices are connected through a VSL to form a line that is called the line topology.

Working Principle

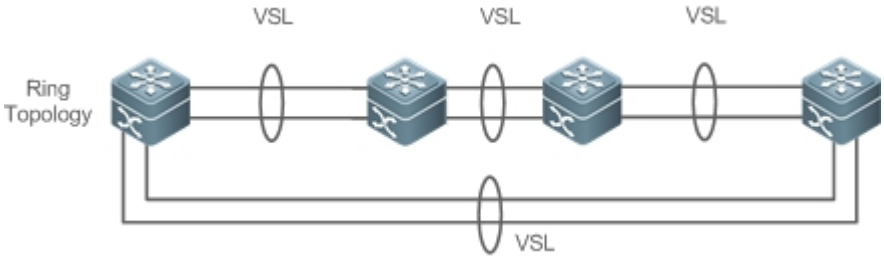
Topology

The line topology is simple. It uses a very few ports and cables. Two devices are connected with a communication link only. Therefore, the VSL has low reliability.

Figure 7- 8 Line topology

Expect for the line topology, devices can also form a ring topology, as shown in Figure 7-9. In the ring topology, the two communication links between devices can back up for each other and perform link redundancy to improve the reliability of the Stacking system.

Figure 7- 9 Ring Topology



You are advised to select the ring topology for the Stacking system, thus the normal operation of the whole Stacking system will not be affected by any single faulty device or VSL.

Besides selecting the ring topology networking, you are advised to configure multiple VSLs for every VSL member to improve the reliability of a single VSL. At least two links are recommended and a maximum of four links can be configured. A reasonable configuration comprises more than two VSLs crossing different cards.

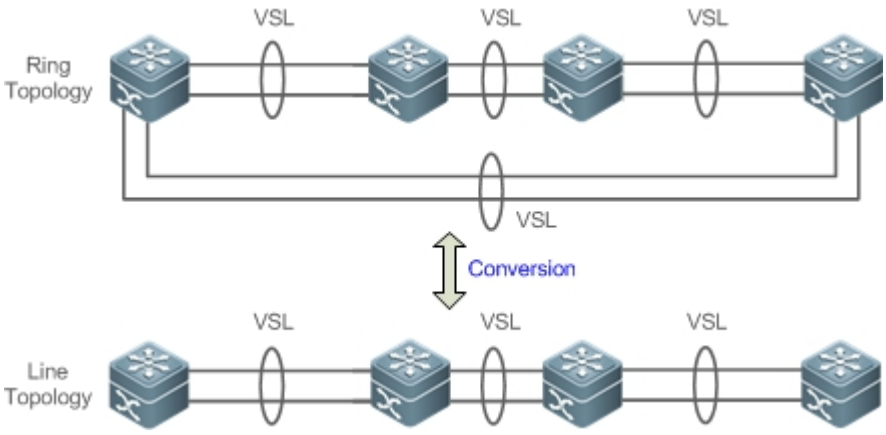
Topology Convergence

Before the establishment of the Stacking, the member devices need to discover neighbors through topology discovery protocols and check devices in the Stacking system to confirm the range of the management domain. Then a global master device is selected to manage the whole Stacking system and a global slave device is selected for backup of the master device. Then the whole Stacking topology is converged. As the start up time differs for different devices, the first convergence time of the topology is also different.

Topology ring and Line Conversion

In a ring topology, if a VSL link is disconnected, the ring topology will convert into a line topology. The whole Stacking system will still run normally without network disconnection. To prevent other VSL links and nodes from being faulty, you are advised to locate the VSL failures and recover the availability of the VSL. After the VSL link is recovered, the line topology will convert into the ring topology.

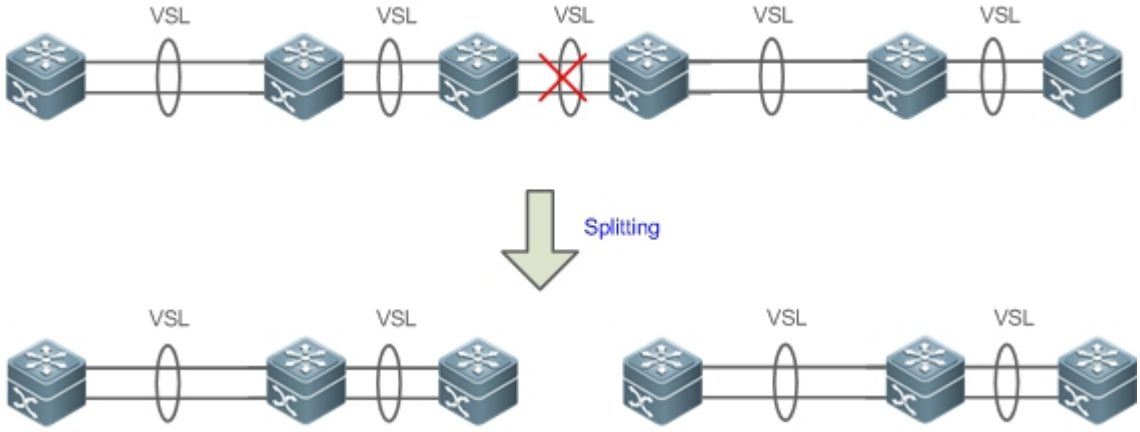
Figure 7- 10 Ring-to-line and line-to-ring



Topology Splitting

In the line topology, if the VSL link is disconnected, the line topology will be split, as shown in Figure 7-11. A Stacking group is split into two groups. In this condition, two devices with the absolutely same configurations may exist on the network, which will cause abnormal operation of the network. Therefore, the multi-active detection (MAD) function (for details, see 1.1.4.6 Multi-Active Detection) needs to be deployed to solve the problem of topology splitting.

Figure 7- 11 Topology splitting



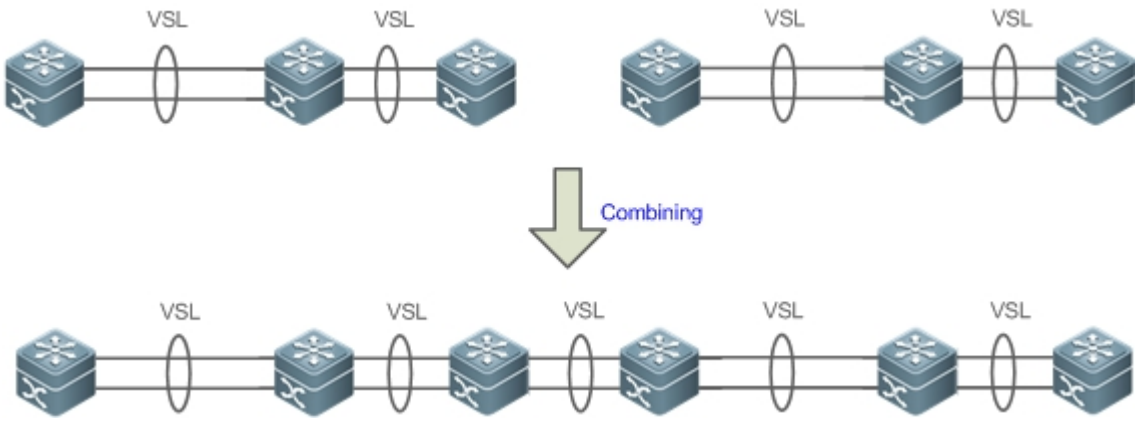
Topology Combining

If the two Stacking groups are connected through the VSL link, the line topology will be combined. During the topology combining, restart one Stacking group and then hot add the other Stacking group.

The principle of topology combining: Minimizing influences on the services during topology combining. The rules are as follows (Judge from the first item. If you cannot select the optimal topology, continue to judge the next item):

- Use the device priority as the first criteria for judging topology combining. Reserve the Stacking group containing a device with the highest priority.
- If the previous item cannot help make a judgment, select the Stacking group with a smaller switch ID (that of the two global master switches).
- If the previous item cannot help make a judgment, reserve the Stacking group with a smaller MAC address (that of the global master switches).

Figure 7- 12 Topology combining



During topology combining of two Stacking groups, the two Stacking groups need to be elected. The Stacking group that fails the election will restart automatically and hot add to the other Stacking group.

8.3.3 Dual-Active Detection (DAD)

Working Principle

When the VSL is disconnected, the slave device switches to the master device. If the original master device is still running, a series of problems including IP address conflict on the LAN will be caused due to there are two master devices and their configurations are the same completely. In this condition, the Stacking system must detect the two devices and take recovery measures. The Stacking system provides two methods to perform MAD as follows:

- Bidirectional forwarding detection (BFD)
- AP-based detection

MAD Rules

1. Select the Stacking group with the highest priority.
2. If the previous item cannot help make a judgment, select the Stacking group with more physical devices.
3. If the previous item cannot help make a judgment, select the Stacking group with a higher health. (Health: total bandwidth of all physical interfaces (except for management and VSL interfaces) in the UP state in the topology.)
4. If the previous item cannot help make a judgment, select the Stacking group with a smaller switch ID (that of the two global master switches).
5. If the previous item cannot help make a judgment, reserve the Stacking group with a smaller MAC address (that of the two global master switches).
6. If the previous item cannot help make a judgment, reserve the Stacking group with a greater startup time (that of the global master switches).

If DAD is not configured, network interruption may be caused after topology splitting.

BFD

The Stacking system supports the BFD to detect multiple master devices. Figure 7-13 shows the topology. A link is added for the two devices on the edges for MAD specially. When the VSL link is disconnected between the global master and slave devices, two master devices exist concurrently. If the BFD function is set, the two master devices will send the BFD packets to each other through the BFD link.

Thereby the same devices are detected on the current system. Finally shut down the Stacking system of a master device according to some rules (for details, see the topology combining rules in the section 1.1.4.4 *Topology Changes*) and enter the recovery state to avoid network abnormality.

Figure 7- 13 BFD

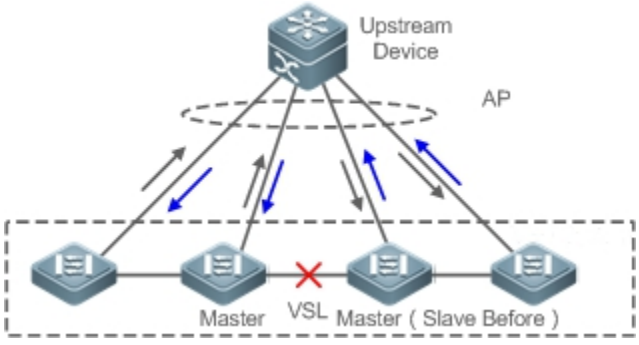
When there is a pair of BFD links, you are advised to deploy the detection links at the two ends of the topology.

You need to adopt the extension BFD and you cannot configure the dual-active detection port by using the existing BFD configurations and commands.

MAD

The Stacking system also supports the MAD dual-active detection mechanism. Figure 7-14 shows the topology. The Stacking system and the upstream device both need to support the MAD function. When the VSL link is disconnected, two master devices exist concurrently. The two master devices respectively send the MAD packets to the member ports of the MAD-APs and then the MAD packets are forwarded to each other through the upstream device. As shown in Figure 7-14, the MAD-AP has four member ports. Each member port is connected to a different device of the Stacking system. When the topology splitting occurs, the four member ports all send and receive the MAD packets. Thereby the same devices are detected on the current system. Finally shut down the Stacking system of a master device according to some rules (for details, see the topology combining rules in the section 7.1.4.4 *Topology Changes*) and enter the recovery state to avoid network abnormality.

Figure 7- 14 MAD based on upstream and downstream devices



In the topology above, the upstream device must be FS device and support the MAD packet forwarding function.

8.3.4 Stacking Traffic Forwarding

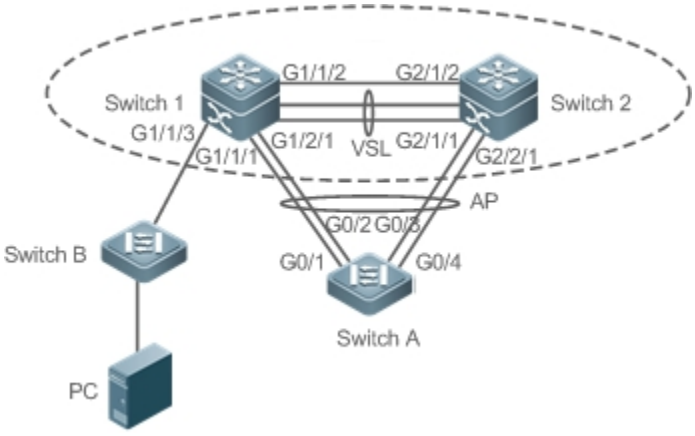
Working Principle

Cross-device AP Group

An AP binds multiple physical links together to form a logical link. The Stacking system supports the AP across the member devices.

As shown in Figure 7- 15, two devices form a Stacking group. The external access device Switch A is connected to the Stacking in the form of the AP. In terms of Switch A, there is no difference between the AP in Figure 7- 15 and the common AP group.

Figure 7- 15 Cross-device aggregation port



Troubleshooting

You are advised to configure the cross-device AP with the physical link between the peripheral device and each Stacking device. On the one hand, the VSL bandwidth can be reserved (prioritize the AP member of the same chassis as the egress to transmitted the cross-chassis AP traffic and prevent unnecessary traffic from being transmitted through the VSL link). On the other hand, the network reliability can be improved (if a certain chassis is faulty, the member ports of normal devices can work normally).

The follows sections describe the possible faults of the cross-device AP and the consequences.

- Single link failure

If a single link of the cross-device AP is faulty but other links still work normally, the cross-device AP will reallocate the traffic for the remaining normal links.

- Link failure of all cross-device AP member ports on the global master device

If the links of all cross-device AP member ports on the global master device fail to work, only the member ports of other member devices continue working normally. In terms of the data stream transmitted through the AP to the Stacking system, if the data stream forwarding egress is on the global master device, the system will forward the data stream to the corresponding egress on the global master device through the VSL link.

The control plane protocols are still running on the global master device. Therefore, the protocol packets that enter the Stacking system need to be forwarded to the global master device through the VSL link for protocol computing.

- Failure of all links of other member devices

If all links of the cross-device AP and a single device A fail to work, only the member ports of other member devices continue working normally. In terms of the data stream transmitted through the AP to the Stacking system, if the data stream forwarding egress is on the member device A, the system will forward the data stream to the corresponding egress on the member device A through the VSL.

- Failure of all links

If all links of the cross-device AP fail to work, the interface status will be Link-Down.

- Global master device fault

If the global master device is faulty, the hot backup switching is performed to switch the original slave device to the master device. Meanwhile, the member ports on other member devices continue working. The link failure is detected on the peer device connected to

the Stacking through this AP. Therefore, the traffic balancing algorithm needs to be adjusted to allocate the data stream to normal links.

- Member device fault

If a member device is faulty, the AP member link connected to this member device is disconnected. However, other member links still work normally. The link failure is detected on the peer device connected to the Stacking through this AP. Therefore, the traffic balancing algorithm needs to be adjusted to allocate the data stream forwarding paths to normal links.

Traffic Balancing

In a Stacking system, traffic may have multiple egresses. The AP and ECMP have their own traffic balancing algorithms, for example, using destination or source MAC addresses. For details, see the *Configuring Aggregate Port*. The local forwarding first (LFF) can be configured detailed in this configuration manual. Packets received by a device are forwarded on this device first. In this way, packets can be forwarded to other devices without using a VSL.

8.3.5 System Management

Working Principle

Access to the Console

The master device console of Stacking system manages multiple devices on the system simultaneously. The consoles of the slave and candidate devices do not support command line input. However, you can configure the Stacking system on the master device for a specified member device and log in to the master device console through the serial port of the slave device. A session can be used to redirect to the master console of a device.

Slot Naming

In terms of the chassis device, in the Stacking mode, the slot is named with the device number (Switch ID). Therefore, the slot number turns from one-dimensional into two-dimensional. For example, cable clip 1/1 indicates the slot numbered 1 of the slot 1 on a member device.

Interface Naming

In the Stacking working mode, a slot number may occur in multiple devices. Therefore, the interface is named with the device number (Switch ID).

For example, interface gigabitEthernet 1/0/1 indicates the Gigabit port 1 on the slot 0 of the device whose ID is 1; interface gigabitEthernet 2/0/2 indicates the Gigabit port 2 on the slot 0 of the device whose ID is 2.

Access to the File System

In the Stacking working mode, you can access to the file system on other member devices from the master device. The detailed access method is the same to that of the local file system. The unique difference is that different URL prefixes are used.

System Upgrade

Generally the Stacking system requires version consistency of the main program version numbers of the member devices. However, there are so many member devices that it takes too much time and energy to perform upgrade one by one in the standalone mode and it is also easy to make mistakes. FS switches provide consummate system upgrade solution to help you with system upgrade by adopting the two methods as follows:

- When the Stacking system is being established: the system will automatically align the main program version numbers of all

member devices. Once the main program versions are discovered inconsistency, the main program of the master device will be selected to be synchronized to all member devices.

- After the Stacking system is established: the main program version will be synchronized to all member devices automatically by using the file that is downloaded by the TFTP.

SYSLOG

All member devices of the Stacking system can display the SYSLOG. The SYSLOG generated by the master device is displayed on the master device console with the same format to that in the standalone mode. The SYSLOG generated by other member devices is also displayed on the master device console, but the message format is different from that in the standalone mode because the device number information is added.

For example, the SYSLOG information generated in the standalone state is "%VSU-5-DTM_TOPO_CVG:Node discovery done. Topology converged." The SYSLOG information generated by the member device numbered 3 is "%VSU-5-DTM_TOPO_CVG:(3) Node discovery done. Topology converged."

8.3.6 Quick Blinking Location

In a network cabling environment, the equipment room where switches are located and the operation console are often at different places. If there are many devices in the environment, network administrators cannot easily locate the locations of specific devices.

Quick blinking location provides network administrators with a method for locating devices by means of quick blinking. By enable this function for a device on the console, you can easily find the corresponding device in the equipment room.

When quick blinking location is enabled, the status LED cannot show original status until the quick blinking location is disabled.

8.4 Configuration

Configuration		Configuration and Command	
Configuring Stacking in the Standalone Mode		(Mandatory) It is used to configure Stacking in the standalone mode.	
		switch virtual domain	Configures the domain ID.
		switch	Configures the switch ID.
		switch priority	Configures the switch priority.
		vsl-port	Enters the VSL interface configuration mode.
		port-member interface	Configures the VSL member interface.
		switch convert mode virtual	Changes the standalone mode to the Stacking mode.
		(Optional) It is used to configure the device attributes in the Stacking mode.	
		switch description	Configures the device description.
		switch crc	Configures error frame check.
Configuring Stacking in the Stacking Mode	Configuring Stacking Attributes	(Optional) It is used to configure the device attributes in the Stacking mode.	
		switch domain	Changes the domain ID.
		switch renumber	Changes the switch ID.
		switch description	Configures the device description.
		switch crc	Configures error frame check.

Configuration		Configuration and Command	
	Configuring the VSL	(Optional) It is used to configure a VSL.	
		vsl-port	Enters the VSL interface configuration mode.
		port-member interface	Configures a VSL member interface.
	Configuring Dual-Active Detection	(Mandatory) It is used to configure DAD.	
		dual-active detection	Configures DAD.
		dual-active bfd interface	Configures the BFD DAD interface.
		dual-active interface	Configures an AP as a DAD interface.
		dual-active exclude interface	Configures an excluded interface.
	Configuring Traffic Balancing	(Optional) It is used to configure traffic balancing in the Stacking mode.	
		switch virtual aggregateport-lff enable	Configures the AP LFF mode.
		switch virtual ecmp-lff enable	Configures the ECMP LFF mode.
	Changing the Stacking Mode to the Standalone Mode	(Optional) It is used to change the Stacking mode to the standalone mode.	
		switch convert mode standalone	Changes the Stacking mode to the standalone mode.
		(Optional) It is used to quickly locate a device.	
		led-blink	Enables quick blinking location.

8.4.1 Configuring Stacking in the Standalone Mode

Configuration Effect

Start up the switch in the standalone mode to set relevant Stacking parameters to establish the Stacking system.

Configuration Steps

Configuring Stacking Attributes

- A switch starts in the standalone mode by default. You need to set the same domain ID on the two chassis of the established Stacking system. The domain ID must be unique within the local area network (LAN). Furthermore, you need to set the ID of each chassis in the Stacking.
- Run the **switch virtual domain** *domain_id* command to configure the domain ID. This command is mandatory.
- Run the **switch** *switch_id* command to configure the device ID in the Stacking. This command is mandatory. For devices with the same priorities in the Stacking system, a device with the smallest device ID is selected as the global master device.
- Run the **switch** *switch_id* **priority** *priority_num* command to configure the device priority. This command is mandatory.
- The value ranges from 1 to 255. A larger value means a higher priority.
- Run the **switch** *switch_id* **description** *switch1* command to configure the device alias. This command is optional. The default name is FS. For easy identification of devices in the network environment, this item can be selected to set the device alias.
- A maximum of 32 characters are allowed.

Command	**switch virtual domain** *number*
Parameter Description	*number:* Indicates domain ID of the Stacking
Defaults	The default domain ID is 100.
Command Mode	Global configuration mode
Usage Guide	Only two devices with the same domain ID can form a Stacking. The domain ID must be unique within the LAN.

Command	**switch** *switch_id*
Parameter Description	*switch_id:* indicates the switch ID in the Stacking system. The value varies with products.
Defaults	The default device ID is 1.
Command Mode	Domain configuration mode
Usage Guide	The device ID identifies each virtual device member. In Stacking mode, the interface name format changes to "switch/slot/port" from "slot/port", in which "switch" is the device ID. If either chassis are active or if the role of the just started chassis is uncertain and both have the same priority, the chassis with a smaller ID is elected as the active one. This command can be only used to modify the device ID in standalone mode. In Stacking mode, run the **switch renumber** command to modify the device ID. The modified device ID takes effect only after you restart the device, regardless of in standalone mode or in Stacking mode.

Command	**switch** *switch_id* **priority** *priority_num*
Parameter Description	*switch_id:* Indicates a switch ID for which a priority needs to be configured. *priority_num:* Indicates the switch priority, ranging from 1 to 255.
Defaults	The default device priority is 100.
Command Mode	Domain configuration mode
Usage Guide	A larger value means a higher priority. A device with the highest priority is chosen as the master device. You can run this command in the standalone or Stacking mode. The modified priority takes effect only after you restart the device. This command is not used to modify the value of *switch_id*. In the standalone mode, if *switch_id* is set to *1*, running the **switch** *2* **priority** *200* command does not work. You can first set *switch_id* to **2** and then run the **switch** *2* **priority** *200* command. In the Stacking mode, *switch_id* indicates the ID of the currently running switch. If the ID does not exist, the configuration does not take effect.

Command	**switch** *switch_id* **description** *dev-name*
Parameter Description	*switch_id:* Indicates the device ID. *dev-name:* Indicates the device description, no greater than 32 characters.
Defaults	N/A
Command Mode	Domain configuration mode

Usage Guide	This command is configured on a device in whether standalone or Stacking mode and takes effect immediately after configuration.

The command used for configuring a priority can modify the priority only rather than modify a switch ID. Therefore, you must enter the current switch ID correctly for the configuration. For example, you have set the switch ID to 1. If you enter **switch** 2 **priority** 100, the priority configuration cannot take effect.

Configuring the VSL

- To establish the Stacking system, you need to decide which ports are configured as the VSL member ports.
- Run the **vsl-port** command to enter the VSL interface configuration mode. This command is mandatory.
- Run the **port-member interface** *interface-name* [**copper** | **fiber**] command to add a VSL interface. This command is mandatory.
- When the device enters the VSL interface configuration mode, the VSL interface can be configured or deleted.

Command	**vsl-port**
Parameter Description	N/A
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	You can run this command in the standalone or Stacking mode.

Command	**port-member interface** *interface-name* [**copper** | **fiber**]
Parameter Description	*interface-name*: Indicates a two-dimensional interface name, such as Tengigabitethernet 1/1 and Tengigabitethernet 1/3. **copper**: Indicates the copper interface attribute. **fiber:** Indicates the optical interface attribute.
Defaults	N/A
Command Mode	VSL interface configuration mode
Usage Guide	Add a member interface of the VSL link. *interface-name* indicates the two-dimensional interface name in the standalone mode. The two-dimensional interface can be the 10 Gigabit interface or Gigabit interface. (The Gigabit interface can be an opto-copper interface. If the media type is not specified, the Gigabit copper interface is adopted by default.) For an opto-copper interface, you must specify its optical or copper interface attribute. A VSL interface for a chassis device must be a 10 Gigabit interface. You can run this command in the Stacking mode or standalone mode. The command can take effect after the command configuration is saved and the device where the VSL member interface resides is restarted.

In the standalone mode, the VSL configurations cannot take effect immediately unless the device shifts into the Stacking mode and restart.

Configuring Error Frame Check

- Run the **switch crc** command to configure error frame check. This command is optional. Run this command to modify the default method for checking error frames.
- If error frames are found on a VSL interface, perform error frame correction. The system detects the VSL interfaces every 5 seconds by default. If the number of error frames is greater than 3 as compared with that detected last time, it is assumed that error frames are detected once. If error frames are detected consecutively for 10 times, it is assumed that the interface is abnormal. If multiple VSL links are available when error frames are detected, the VSL will be switched. The last VSL will not be switched in order to prevent topology splitting.

Command	**switch crc errors** *error_num* **times** *time_num*
Parameter Description	*error_num:* Configures the increase of error frames between two detections. When the number of error frames is greater than the increase, it is assumed that error frames are detected once. *time_num:* Configures the number of times after which an action needs to be taken (the action can be displaying a prompt or disabling the interface).
Defaults	The default value of **errors** is 3; the default value of **times** is 10.
Command Mode	Domain configuration mode
Usage Guide	The system detects the VSL interfaces every 5 seconds by default. If the number of error frames is greater than 3 as compared with that detected last time, it is assumed that error frames are detected once. If error frames are detected consecutively for 10 times, it is assumed that the interface is abnormal. The default action for an abnormal interface is displaying a log prompt. You can set the action to disabling the interface. If the interface is disabled, you must recover it by unplugging and plugging it.

Different products have different requirements for error frame check and different processing for VSL interfaces. In version 11.0, error frame check is configurable.

Changing the Standalone Mode to the Stacking Mode

- Use the **switch convert mode virtual** command to change the standalone mode to the Stacking Mode.
- In the standalone mode, the software will take the following actions after you run the **switch convert mode virtual** command.

  Back up the global configuration file *config.text* in the standalone mode as *standalone.text* for subsequent use.

  Clear the contents of the configuration file *config.text*.

  Write the relevant Stacking configurations to the special configuration file *config_VSU.dat*.
- If there is a *virtual_switch.text* file on the switch, the system will prompt you whether to overwrite the contents of the file *virtual_switch.text* to the file *config.text* (the file *virtual_switch.text* is a backup file for the file *config.text* when the switch shifts from the Stacking mode to the standalone mode). Then you can click **Yes** or **No**. Finally the switch restarts in the Stacking mode and reads Stacking parameters in the file *config_VSU.dat*.

Command	**switch convert mode virtual**
Parameter Description	N/A
Defaults	The switch is in the standalone mode by default.
Command Mode	Privileged EXEC mode
Usage Guide	Change the standalone mode to the Stacking mode.

Verification

Run the **show switch virtual config** [*switch_id*] command to check the Stacking configuration of the current switch in the standalone mode.

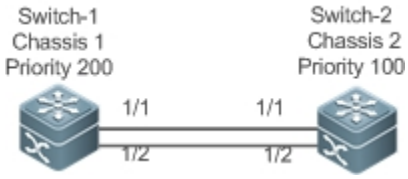
Command	**show switch virtual config** [*switch_id*]
Parameter Description	*switch_id*: Indicates the switch ID. After this parameter is specified, only the Stacking configuration of the specified device is displayed.
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to display the Stacking configuration in the standalone or Stacking mode.

The relevant Stacking configurations are set for a single physical switch and the configurations are stored in the special configuration file config_VSU.dat. Therefore, you can view the current Stacking configurations by running the **show switch virtual config** command rather than the **show running config** command.

In the standalone mode, the Stacking running information is null. When you enter commands such as show switch virtual, the system will prompt you that the switch is in the standalone mode and there is no Stacking running information.

Configuration Example

Configuring **Stacking** in the Standalone Mode

Scenario **Figure 7- 16**	 Switch 1 and Switch 2 form a Stacking system. The domain ID is 100. The chassis on the left side is configured as Chassis 1, with the priority of 200, alias of Switch 1, and the VSL interfaces of 1/1 and 1/2. The chassis on the right side is configured as Chassis 2, with the priority of 100, alias of Switch 2, and the VSL interfaces of 1/1 and 1/2.
Configuration Steps	5. Perform the following configuration on the Switch 1: ● Configure Stacking attributes and VSL interfaces. ● Change the standalone mode to the Stacking mode. 6. Perform the following configuration on the Switch 2: ● Configure Stacking attributes and VSL interfaces. ● Change the standalone mode to the Stacking mode.
Switch-1	FS# configure terminal FS(config)# switch virtual domain 100 FS(config-vs-domain)#switch 1 FS(config-vs-domain)#switch 1 priority 200 FS(config-vs-domain)#witch 1 description switch-1 FS(config-vs-domain)# switch crc errors 10 times 20 FS(config-vs-domain))#exit

	FS(config)#vsl-port FS(config-vsl-port)#port-member interface Tengigabitethernet 1/1 FS(config-vsl-port)#port-member interface Tengigabitethernet 1/2 FS(config)#exit FS#switch convert mode virtual
Switch-2	FS# configure terminal FS(config)# switch virtual domain 100 FS(config-vs-domain)# switch 2 FS(config-vs-domain)# switch 2 priority 200 FS(config-vs-domain)# switch 2 description switch-2 FS(config-vs-domain)# switch crc errors 10 times 20 FS(config-vs-domain))#exit FS(config)#vsl-port FS(config-vsl-port)#port-member interface Tengigabitethernet 1/1 FS(config-vsl-port)#port-member interface Tengigabitethernet 1/2 FS(config-vsl-port)#exit FS#switch convert mode virtual
Verification	● Run the **show switch virtual config** command to view the Stacking attributes of Switch 1 and Switch 2.
Switch-1	FS#show switch virtual config switch_id: 1 (mac: 0x1201aeda0M) ! switch virtual domain 100 ! switch 1 switch 1 priority 100 ! switch convert mode virtual ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/2 ! switch crc errors 10 times 20

	!
Switch-2	FS#show switch virtual config switch_id: 2 (mac: 0x1201aeda0E) ! switch virtual domain 100 ! switch 2 switch 2 priority 100 ! switch convert mode virtual ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/2 ! switch crc errors 10 times 20 !

Common Errors

A VSL interface of a chassis device must be 10 Gigabit or higher.

8.4.2 Configuring Stacking in the Stacking Mode

Configuring Stacking Attributes

Configuration Effect

During the Stacking system running, you can modify the parameters, such as domain ID, switch ID, and priority of the master device or the slave device. However, you can only log in to the Stacking master device console to modify these parameters, but cannot enter the global configuration mode from the slave device console.

Notes

- Among the commands above, the all configuration commands take effect only after the switch restarts except the **switch** *sw_id* **description** *switch1* command that can take effect immediately.

Configuration Steps

Entering the Domain Configuration Mode

- Optional.
- Run this command in the Stacking mode to enter the domain configuration mode. Switches with the same domain ID form a Stacking system. You can modify or configure the domain ID, switch priority, and switch ID only after entering the domain configuration mode in the Stacking mode.

Command	**switch virtual domain** *domain_id*
Parameter Description	*domain_id*: Indicates the virtual domain ID of the Stacking system.
Defaults	The default domain ID is 100.
Command Mode	Global configuration mode
Usage Guide	Only two devices with the same domain ID can form a Stacking system. The domain ID must be unique on a LAN.

Changing the Domain ID

- Optional.
- To modify the value of *domain_id* for a device, you can configure this item on the master device console of the Stacking system.

Command	**switch** *switch_id* **domain** *new_domain_id*
Parameter Description	*switch_id*: Indicates the ID of the currently running switch in the Stacking mode, ranging from 1 to 8. *new_domain_id*: Indicates the modified domain ID, ranging from 1 to 255.
Defaults	The default domain ID is 100.
Command Mode	Domain configuration mode
Usage Guide	Run this command only in the Stacking mode. In addition, the setting can take effect only after the device is restarted.

Changing the Switch ID

- Optional.
- To modify the value of *switch_id* for a device, you can configure this item on the master device console of the Stacking system.

Command	**switch** *switch_id* **renumber** *new_switch_id*
Parameter Description	*switch_id*: Indicates the ID of a switch. In a Stacking system, the switch ID ranges from 1 to 16 for cassette switches, and from 1 to 4 for chassis switches. *new_switch_id*: Indicates the modified switch ID.
Defaults	N/A
Command Mode	Domain configuration mode
Usage Guide	Run this command only in the Stacking mode. In addition, the setting can take effect only after the device is restarted.

Changing the Switch Priority

- Optional.
- To modify the priority of a device, you can configure this item on the master device console of the Stacking system.
- A larger value means a higher priority. Select the device with the highest priority as the master device.

Command	**switch** *switch_id* **priority** *priority_num*
Parameter	*switch_id:* Indicates a switch ID for which a priority needs to be configured.

Description	*priority_num:* Indicates the switch priority, ranging from 1 to -255 for cassette switches.
Defaults	The default priority is 100.
Command Mode	Domain configuration mode
Usage Guide	A larger value means a higher priority. Select the device with the highest priority as the master device. You can run this command in the standalone or Stacking mode. The modified priority takes effect only after you restart the device. This command is not used to modify the value of **switch_id**. In the standalone mode, if **switch_id** is set to **1**, running the **switch 2 priority 200** command does not work. You can first set **switch_id** to **2** and then run the **switch 2 priority 200** command. In the Stacking mode, **switch_id** indicates the ID of the currently running switch. If the ID does not exist, the configuration does not take effect.

Configuring the Device Description

- Optional.
- To configure the description for a device, you can configure this item on the master device console of the Stacking system.
- Run the **switch** *switch_id* **description** *switch1* command to configure the device description. A maximum of 32 characters are allowed.

Command	**switch** *switch_id* **description** *dev-name*
Parameter Description	*switch_id:* Indicates a switch ID for which a priority needs to be configured. *dev_name*: Indicates the device name.
Defaults	N/A
Command Mode	Domain configuration mode
Usage Guide	You can run this command in the standalone or Stacking mode. The configuration takes effect immediately in the Stacking mode.

Configuring Error Frame Check

- Optional.
- Run the **switch crc errors** *error_num* **times** *time_num* command to configure the conditions for triggering error frame check.

Command	**switch crc errors** *error_num* **times** *time_num*
Parameter Description	*error_num:* Configures the increase of error frames between two detections. When the number of error frames is greater than the increase, it is assumed that error frames are detected once. time_num: Configures the number of times after which an action needs to be taken (the action can be displaying a prompt or disabling the interface).
Defaults	The default value of **errors** is **3**; the default value of **times** is **10**.
Command Mode	Domain configuration mode
Default Level	14
Usage Guide	N/A

Saving the Configuration File

Run the **exit** command to exit from the virtual device configuration mode and run the **write** command to save the configurations to the

config_VSU.dat file.

Verification

Use the **show switch virtual** [**topology | config**] command to display the current Stacking running information, topology or configuration parameters.

Command	**show switch virtual** [**topology | config**]
Parameter Description	**Topology**: Indicates topology information. **Config**: Indicates the Stacking configurations.
Command Mode	Privileged EXEC mode
Usage Guide	View the domain ID, and the device ID, status and role of each device.

Configuration Example

Configuring **Stacking** Attributes

Scenario **Figure 7-17**	 Switch 1 and Switch 2 form a Stacking system. Modify the chassis ID of Switch 2 to 3 and its priority to 150. Assume that Switch 1 is the global master switch and perform the configuration on the global master switch.
Configuration Steps	● Modify the configurations of Switch 2.
Switch-1	FS#config FS(config)# switch virtual domain 100 FS(config-vs-domain)# switch 2 renumber 3 FS(config-vs-domain)# switch 2 priority 150 FS(config-vs-domain)# switch 2 description switch-3
Verification	● Run the **show switch virtual config** command for verification.
Switch-1	FS#show switch virtual config switch_id: 1 (mac: 0x1201aeda0M) ! switch virtual domain 100 ! switch 1 switch 1 priority 100 !

	switch convert mode virtual ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/2 !
	switch_id: 3 (mac: 0x1201aeda0E) ! switch virtual domain 100 ! switch 3 switch 3 priority 150 ! switch convert mode virtual ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/2 ! switch 3 description switch-3 !

Configuring the VSL

Configuration Effect

When switches form a Stacking system or when the Stacking system is running, you can shift between common interfaces and VSL interfaces. However, you can only log in to the master device console of the Stacking system for modification, but cannot enter the global configuration mode from the slave device console.

Notes

- You can log in to the console of the Stacking system by using a serial port or telnet, in order to add or delete the configurations of VSL member interfaces.
- To prevent incorrect connections in actual scenarios, the VSL AP uses dynamic negotiation. You need to configure the VSL interface pool first, and then add the VSL interface pool to the same AP after successful negotiation. Interfaces connecting to the same device are within the same AP.

Configuration Steps

Entering the VSL Interface Configuration mode

- Run the **vsl-port** command to enter the VSL-PORT configuration mode. This command is optional.
- When the device enters the VSL-PORT configuration mode, the VSL interface can be configured or deleted.

Command	**vsl-port**
Parameter Description	N/A
Defaults	N/A
Command Mode	Global configuration mode
Usage Guide	You can run this command in the standalone or Stacking mode.

Configuring a VSL Member Interface

- Run the **port-member interface** *interface-name* [**copper** | **fiber**] command to add a VSL interface. This command is optional.
- Run the **port-member interface** command to configure a VSL member interface.

Command	**port-member interface** *interface-name* [**copper** | **fiber**]
Parameter Description	*interface-name*: Indicates a two-dimensional interface name, such as GigabitEthernet 0/1 and GigabitEthernet 0/3. **copper**: Indicates the copper interface attribute. **fiber:** Indicates the optical interface attribute.
Defaults	N/A
Command Mode	VSL interface configuration mode
Usage Guide	You can run this command in the Stacking mode or standalone mode. The command can take effect after the command configuration is saved and the device where the VSL member interface resides is restarted.

During the Stacking system running, the configured VSL member links take effect immediately. VSL interfaces need to be configured for all devices.

For chassis devices, VSL interfaces must be optical interfaces of 10 Gigabit or higher; for cassette devices, VSL interfaces can be optical and copper interfaces of Gigabit or higher.

Modules on chassis devices must be modules of 10 Gigabit or higher.

40G one-to-four interfaces cannot be configured as VSL interfaces.

For a 40G port (no matter whether splitting is performed for the interface), its member interfaces (namely, four 10G interfaces) cannot be shifted to VSL member interfaces.

If an interface has been configured as an NLB reflex interface, this interface can be shifted to a VSL member interface only after the NLB reflex interface configuration is deleted.

To prevent a loop that may occur when a VSL member interface exits from the VSL AP, the system automatically sets the member interface to the shutdown state when the command is executed to make the VSL member interface exit from the VSL AP. After the VSL member interface exits from the VSL AP, you can reconnect the link and run the **no shutdown** command to enable this interface again. When you configure a VSL interface, the system will shut it down first. If the configuration fails and you want to use it as a common interface, you can run the **no shutdown** command to enable this interface again. Add a member interface number that must be a three-dimensional interface number. For example, in the VSL-PORT configuration mode, if you run the **port-member interface** Tengigabitethernet 1/1/1 command, it indicates that you configure the global three-dimensional interface

1/1/1 as a VSL interface.

If Stacking topology splitting occurs when you change a VSL interface to a common interface, the VSL interface cannot be deleted. You can disconnect the physical interface first and then delete the VSL interface.

Verification

- Use the show switch virtual link [port] to display the current VSL link running information in the Stacking mode.

Command	**show switch virtual link** [**port**]
Parameter Description	**port**: Displays the status information of the VSL member interfaces.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuration Example

Configuring the VSL

Scenario Figure 7-18	
Configuration Steps	- Add interface 1/1/3 as the VSL interface for Switch 1 and delete interface 1/1/2 from the VSL interface.
Switch-1	FS#config FS(config)# vsl-port FS(config-vsl-port)# port-member interface Tengigabitethernet 1/1/3 FS(config-vsl-port)# no port-member interface Tengigabitethernet 1/1/2
Verification	- Run the **show switch virtual config** command to view the VSL. Assume that Switch 1 is the global master switch and run the command on the global master switch.
Switch-1	FS#show switch virtual config switch_id: 1 (mac: 0x1201aeda0M) ! switch virtual domain 100 ! switch 1 switch 1 priority 100 !

	switch convert mode virtual ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/3 !
	switch_id: 3 (mac: 0x1201aeda0E) ! switch virtual domain 100 ! switch 3 switch 3 priority 150 ! switch convert mode virtual ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/2 ! switch 3 description switch-3 !

Configuring Dual-Active Detection

Configuration Effect

Configure the relevant detection mechanism to prevent the dual-active is being generated.

Notes

- The DAD can be configured only in the Stacking mode. You are not allowed to configure the DAD mechanism in the standalone mode.
- All DAD configurations will take effect immediately after being configured on the master or slave devices in global configuration mode by running the **show running-config** command.
- The BFD-detected configuration information can be displayed only by running the dual-active detection display command rather that the BFD display command.

Configuration Steps

Configuring the BFD DAD

- The BFD DAD requires establishing a directly connected link between two switches. The interfaces on the two ends must be physical routing interfaces. The following configuration must be performed on both chassis.
- Enter the interface configuration mode of the DAD interface and configure the DAD interface as a routing interface.
- After exiting from the interface configuration mode, run the **switch virtual domain** *domain_id* command to enter the domain configuration mode.
- In the domain mode, run the **dual-active detection bfd** command to enable BFD. This command is optional and can be used when BFD DAD needs to be configured.
- In the domain configuration mode, run the **dual-active bfd interface** *interface-name* command to configure the BFD DAD interface. This command is optional and can be used to configure the BFD DAD interface when BFD DAD is configured.
- Delete the BFD DAD interface. If no BFD DAD interface is available, BFD detection cannot be used.

Command	**switch virtual domain** *domain_id*
Parameter Description	*domain_id*: Indicates the domain ID.
Defaults	The default domain ID is 100.
Command Mode	Global configuration mode
Usage Guide	Only two devices with the same domain ID can form a Stacking system. The domain ID must be unique on a LAN.

Command	**dual-active detection** { **aggregateport | bfd** }
Parameter Description	**Aggregateport**: Specifies the AP detection mode. **bfd**: Specifies the BFD detection mode.
Defaults	The DAD is disabled.
Command Mode	Domain configuration mode
Usage Guide	Configure this command only in the Stacking mode.

Command	**dual-active bfd interface** *interface-name*
Parameter Description	*interface-name*: Indicates the interface type and ID.
Defaults	N/A
Command Mode	Domain configuration mode
Usage Guide	A BFD DAD interface must be a routing interface and on different switches.

The BFD detection interfaces must be directly connected physical routing ports. The two ports must be on different devices.

The interface type is not limited. The dual-active detection link is only used to transmit BFD packets with a small amount of traffic. Therefore, you are advised to adopt the Gigabit interface or 100 M interface as the dual-active detection interface.

After the layer 3 routing interface that is configured with two master devices is converted into a layer 2 switch interface (run the switchport command under this interface), the BFD dual-active detection will be cleared automatically.

You are advised to directly connect BFD detection interfaces only to the master and slave devices.

When the Stacking system detects dual-active conflict and brings another Stacking group to the recovery state, you can resolve the problem only by rectifying the VSL fault, but not directly restoring the Stacking group in the recovery state; otherwise, dual-active conflict may be caused on the network.

Configuring the AP-based DAD

- To configure the AP-based DAD, you must configure an aggregate port (AP) first and then specify the AP port as the DAD interface.
- Run the **port-group** *ap-num* command to add a physical member interface to the AP.
- After entering the domain configuration mode, run the **dual-active detection aggregateport** command to enable AP detection mode. This command is optional. You can run this command when AP detection needs to be configured.
- Run the **dual-active interface** *interface-name* command to configure the AP as the DAD interface. This command is optional. Yu can run this command to configure the AP as the DAD interface when AP detection needs to be configured.
- Run the **dad relay enable** command to enable dual-active detection packet relay for upstream and downstream interfaces. This command is optional. You can run this command to relay DAD packets (dual-active detection packets) when AP-based DAD is configured.
- Disabling AP-based DAD will inactivate DAD.
- Delete the detected interface. If no AP-based DAD interface is available, AP-based DAD cannot be used.
- The AP-based DAD packet relay is disabled by default.

Command	**dual-active detection** { **aggregateport** | **bfd** }
Parameter Description	**aggregateport**: Specifies the AP detection mode. **bfd**: Specifies the BFD detection mode.
Defaults	The DAD is disabled.
Command Mode	Domain configuration mode
Usage Guide	Configure this command only in the Stacking mode.

Command	**dual-active interface** *interface-name*
Parameter Description	*interface-name*: Indicates the interface type and interface ID. An AP-based DAD interface must be specified.
Defaults	N/A
Command Mode	Domain configuration mode
Usage Guide	Only one AP-based DAD interface can be configured. This interface must be created before you configure an AP as a DAD interface. Subsequently configured DAD interfaces will overwrite the previous ones.

Command	**dad relay enable**
Parameter Description	N/A
Defaults	The AP-based DAD packet relay is disabled by default.
Command Mode	Interface configuration mode

Usage Guide	This command can only be executed on the AP.

You are advised to distribute the physical interfaces that are added to the AP-based detection interface to different devices.

Configuring the excluded interface in the recovery mode

- When two master devices are detected, one of them must enter the recovery mode. In the recovery mode, you need to disable all service interfaces. For some special usages (for example, configuring a management switch from which you can log in to a remote interface), you can set some ports to excluded interfaces that are not disabled in the recovery mode.
- In the domain configuration mode, run the **dual-active exclude interface** *interface-name* command to specify an excluded interface that will not be disabled in the recovery mode. This command is optional.

Command	**dual-active exclude interface** *interface-name*
Parameter Description	*interface-name*: Indicates the interface type and interface ID.
Defaults	N/A
Command Mode	Domain configuration mode
Usage Guide	Configure this command only in the Stacking mode. An excluded interface must be a routing interface instead of a VSL interface. You can configure multiple excluded interfaces.

The excluded interface must be routing rather than VSL.

After the excluded interface is converted from a routing one into a switch interface (run the **switchport** command under this interface), the configurations of the excluded interface that is associated with this interface will be cleared automatically.

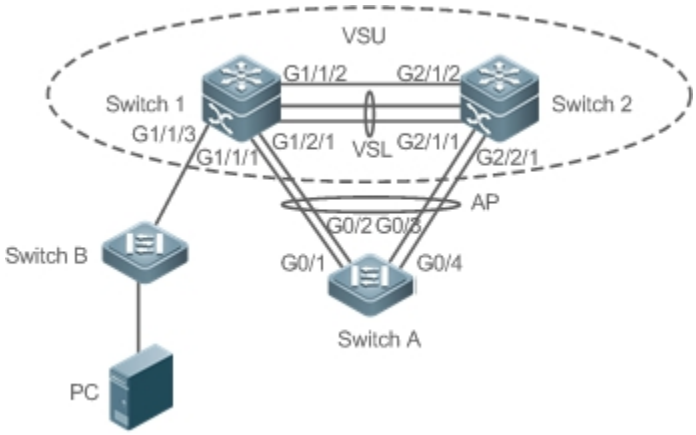
Verification

Use the **show switch virtual dual-active** { **aggregateport** | **bfd** | **summary** } to display the current DAD configuration.

Command	**show switch virtual dual-active** { **aggregateport** | **bfd** | **summary**}
Parameter Description	**aggregateport**: Displays DAD information on the AP. **bfd**: Displays BFD-based DAD information. **summary**: Displays DAD summary.
Command Mode	Privileged EXEC mode
Usage Guide	N/A

Configuration Example

Configuring the BFD DAD

Scenario **Figure 7-19**	 ● Switch 1 and Switch 2 form a Stacking (The domain ID is 1) system. The priorities of Switch 1 and Switch 2 are 200 and 150 respectively. The links between Te1/3/1 and Te1/3/2 of Switch 1 and Te2/3/1 and Te2/3/2 of Switch 2 are established respectively to form a VSL between Switch 1 and Switch 2. The G0/1, G0/2, G0/3 and G0/4 interfaces of Switch A are connected to G1/1/1 and G1/2/1 of Switch 1 and G2/1/1 and G2/2/1 of Switch 2 to form an AP group including four member links. The ID of the AP group is 1. All members of AP group 1 are Gigabit optical interfaces. G1/1/2 and G2/1/2 are routing interfaces. ● G1/1/2 and G2/1/2 are a pair of BFD DAD interfaces.
Configuration Steps	● Configure G1/1/2 and G2/1/2 as routing interfaces. ● Enable the BFD DAD. ● Configure G1/1/2 and G2/1/2 as BFD DAD interfaces. Since Switch 1 and Switch 2 are in a Stacking system, the preceding configuration can be performed on either Switch 1 or Switch 2. on the following example configures the functions on Switch 1.
Switch 1	FS(config)# interface GigabitEthernet 1/1/2 FS(config-if-GigabitEthernet 1/1/2)# no switchport FS(config)# interface GigabitEthernet 2/1/2 FS(config-if-GigabitEthernet 2/1/2)# no switchport FS(config-if)# switch virtual domain 1 FS(c config-vs-domain)# dual-active detection bfd FS(config-vs-domain)# dual-active bfd interface GigabitEthernet 1/1/2 FS(config-vs-domain)# dual-active bfd interface GigabitEthernet 2/1/2
Switch A	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# interface aggretegateport 1 FS(config-if-aggretegateport 1)# interface range GigabitEthernet 0/1-4 FS(config-if-aggretegateport 1)# port-group 1 FS(config)# interface vlan 1

	FS(config-if-vlan 1)#ip address 1.1.1.2 255.255.255.0 FS(config-if-vlan 1)#exit FS(config)#interface aggregateport 1 FS(config-if-AggregatePort 1)# dad relay enable FS(config-if-AggregatePort 1)# exit
Verification	● View the DAD configuration. ● View the BFD DAD configuration.
Switch 1	FS# show switch virtual dual-active summary BFD dual-active detection enabled: No Aggregateport dual-active detection enabled: Yes Interfaces excluded from shutdown in recovery mode: In dual-active recovery mode: NO FS# show switch virtual dual-active bfd BFD dual-active detection enabled: Yes BFD dual-active interface configured: GigabitEthernet 1/1/2: UP GigabitEthernet 2/1/2: UP

Common Errors

- A BFD DAD interface is not a routing interface.
- Neither BFD DAD nor AP-based DAD are enabled and activated.

Configuring Traffic Balancing

Configuration Effect

In the Stacking system, if egresses are distributed on multiple devices, the Local Forward First (LFF) can be configured.

Notes

The default configuration is LFF.

Configuration Steps

Configuring the AP LFF mode

- In the domain configuration mode, run the **switch virtual aggregateport-lff enable** command to enable the AP LFF mode. This command is optional.
- The member ports of AP can be distributed on two chassis of the Stacking system. You can configure whether the AP egress traffic is forwarded through local member ports first based on actual traffic conditions.
- If this function is disabled, traffic is forwarded based on the AP configuration rules. For details, see the *Configuring Aggregate Port*.

Command	**switch virtual aggregateport-lff enable**
Parameter Description	N/A
Defaults	This function is enabled by default.
Command Mode	Domain configuration mode
Usage Guide	Enable the AP LFF in the Stacking mode.

Configuring the ECMP LFF mode

- In the domain configuration mode, run the **switch virtual ecmp-lff enable** command to enable the ECMP LFF mode. This command is optional.
- The Equal-Cost MultiPath (ECMP) routing egress can be distributed on two chassis of the Stacking system. You can configure whether the ECMP egress traffic is forwarded through local member ports first based on actual traffic conditions.
- If this function is disabled, traffic is forwarded based on the ECMP configuration rules. For details, see the *Configuring Aggregate Port*.

Command	**switch virtual ecmp-lff enable**
Parameter Description	N/A
Defaults	This function is enabled by default.
Command Mode	Domain configuration mode
Usage Guide	Enable the ECMP LFF in the Stacking mode.

In the Stacking mode, the across-chassis AP LFF mode and the ECMP LFF mode are disabled by default.

To deploy a Stacking statussystem for layer-3 switches, you are advised to configure the IP-based AP load balancing (src-ip, dst-ip abd src-dst-ip).

Verification

Use the **show switch virtual balance** command to display the current traffic balancing mode of the Stacking system.

Command	**show switch virtual balance**
Parameter Description	N/A
Command Mode	Privileged EXEC mode
Usage Guide	Use this command to display the configuration of the traffic balancing mode in the Stacking mode.

Configuration Example

Configuring the LFF

Scenario **Figure 7-20**	
	In Figure 7- 20, Switch 1 and Switch 2 form a Stacking system. It is assumed that Switch 1 is the global master switch and configuration is performed on Switch 1.
Configuration Steps	● Configure the AP LFF.
Switch-1	FS#config FS(config)# switch virtual domain 100 FS(config-vs-domain)# switch virtual aggregateport-lff enable
Verification	● Run the **show switch virtual balance** command for verification.
Switch-1	FS#show switch virtual balance Aggregate port LFF: enable Ecmp lff enable

Changing the Stacking Mode to the Standalone Mode

Configuration Effect

Dismiss the Stacking system into individual devices that can operate in the standalone mode.

Configuration Steps

- Run the **switch convert mode standalone** [*switch_id*] command to change the Stacking mode to the standalone mode. This command is optional.
- After you run this command, the system will prompt you as follows: Whether to restore the configuration file to standalone text? If **yes**, the configuration file will be restored; if **no**, the configuration of virtual device mode will be cleared.

Command	**switch convert mode standalone** [*switch_id*]
Parameter Description	*switch_id*: Indicates the switch ID.
Defaults	The switch is in the standalone mode by default.
Command Mode	Privileged EXEC mode
Usage Guide	After you run the **switch convert mode standalone** command, the master switch backs up the global configuration files of all VSDs in the Stacking mode as *vsd.virtual_switch.text.vsd* ID. Then, the master switch clears the global configuration files *config.text* of all VSDs in the Stacking mode, and asks you whether to overwrite the global configuration files *config.text* with *vsd.standalone.text.vsd* ID. If you select **yes**, the content of *vsd.standalone.text.vsd* ID will overwrite the global configuration file *config.text* of all VSDs; otherwise, the master switch does not recover *config.text*. Finally, restart the switch. This command can be used in the standalone mode or Stacking mode. If the command is executed in the standalone

	mode, the mode switching is performed on the current switch. If the command contains the *sw_id* parameter and is executed in the Stacking mode, the mode switching is performed on the switch with the ID specified by *sw_id*. If the command does not contain the *sw_id* parameter, the mode switching is performed on the master switch. You are advised to switch the mode of the slave switch and then that of the master switch.

Configuration Example

Changing the **Stacking** Mode to the Standalone Mode

Scenario **Figure 7- 21**	 In Figure 7- 21, it is assumed that Switch 1 and Switch 2 form a Stacking system and Switch 1 is the global master switch.
Configuration Steps	● Change the mode of Switch 1 to the standalone mode. ● Change the mode of Switch 2 to the standalone mode.
Switch-1	FS# switch convert mode standalone 1 FS# switch convert mode standalone 2
Verification	Run the **show switch virtual config** command to display the switch status.
Switch-1	FS#show switch virtual config switch_id: 1 (mac: 0x1201aeda0M) ! switch virtual domain 100 ! switch 1 switch 1 priority 100 ! switch convert mode standalone ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/3 !
	switch_id: 2 (mac: 0x1201aeda0E) !

	switch virtual domain 100 ! switch 2 switch 2 priority 150 ! switch convert mode standalone ! port-member interface Tengigabitethernet 1/1 ! port-member interface Tengigabitethernet 1/2 ! switch 2 description switch-2 !

8.4.3 Configuring Quick Blinking Location

Configuration Effect

- Enable quick blinking location of a switch to make the status LED of the switch quickly blink.

Notes

If you do not disable quick blinking location, the system automatically disables the function 30 minutes after it is enabled.

Configuration Steps

Enabling/Disabling quick blinking location

- Mandatory. Use this function on a switch that needs to be located.
- In the privileged EXEC mode, run the **led-blink** command to enable quick blinking location.

Command	**led-blink** { **enable** | **disable** } [**device** *device_id*]
Parameter Description	**enable**: Enables quick blinking location. **disable**: Disables quick blinking location. *device_id*: Indicates the device ID.
Defaults	Quick blinking location is disabled by default.
Command Mode	Privileged EXEC mode
Usage Guide	Run this command without the *device_id* parameter to enable or disable the quick blinking search in the standalone mode. In the Stacking mode, you can set the *device_id* parameter to enable or disable this function for a specified device. If you ignore the **device_id** parameter, you can enable or disable this function for all devices in the Stacking system. If you do not disable this function, the system automatically disables the function 30 minutes after it is enabled. This configuration cannot be saved. Quick blinking location will be disabled upon restart or failover.

Verification

- Check whether the status LED of a switch quickly blinks.

Configuration Example

Enabling quick blinking location for the two Stacking devices

Scenario	Assume that Switch 1 and Switch 2 form a Stacking system and Switch 1 is the global master device.
Configuration Steps	• Enter the **led-blink enable device** *2* command on the Switch 1 console to enable quick blinking location. • Enter the **led-blink enable device** *2* command on the Switch 1 console to disable quick blinking location.
Verification	When quick blinking location is enabled, check whether the status LED of Switch 2 quickly blinks.

8.5 Monitoring

Displaying

Description	Command
Displays the current Stacking operation, topology or configuration.	**show switch virtual** [**topology | config | role**]
Displays the current dual-active configuration.	**show switch virtual dual-active** { **bfd** | **aggregateport** | **summary** }
Displays the current VSL running information in the Stacking mode.	**show switch virtual link** [**port**]
Redirects to the console of the master switch or any switch.	**session** { **device** *switch_id* | **master** }
Displays the current switch ID.	**show switch id**

9 Configuring RNS

9.1 Overview

The reliable network service (RNS) tests specific services provided by a peer device to monitor the service availability, integrity of the end-to-end connection, and service quality. Using the RNS test results, you can:

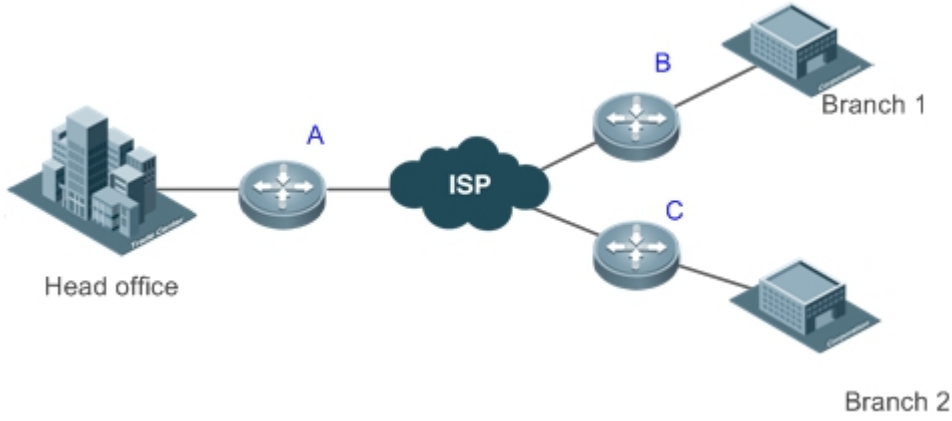
- Learn the network performance in time and take measures accordingly to handle related network performance problems.
- Diagnose and locate network faults.

9.2 Applications

9.2.1 Testing and Evaluating Service Performance

Scenario

As shown in the following figure, a company is going to deploy a video conference system between the headquarters and branches, and has completed the related quality of service (QoS) configurations. Before formal deployment, it must be checked whether the services can be provisioned normally under the existing service pressure of the company. The video conference system is sensitive to the User Datagram Protocol (UDP) delay and UDP transmission jitter of the network. The traditional ping tool can test the Internet Control Message Protocol (ICMP) performance, but cannot effectively evaluate the UDP transmission performance and cannot meet the requirement for jitter measurement.



Remarks	A, B, and C are switches.

Deployment

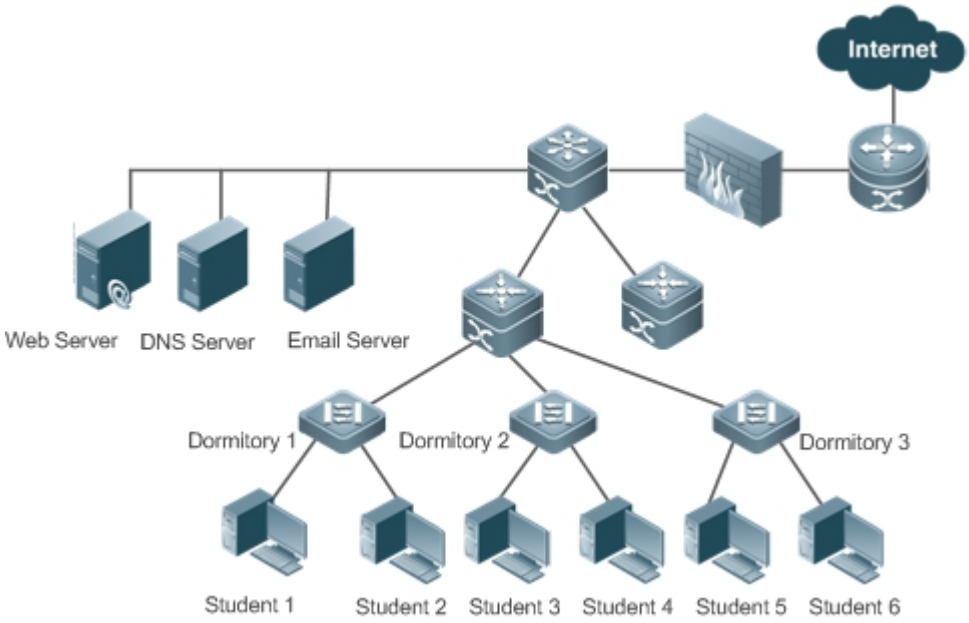
- Configure RNS on the egress switching device or switch of each branch to test the UDP jitter and delay.
- On Switch A, specify the IP address and UDP port of the egress switching device or switch in the headquarters, and then UDP packets can be automatically sent. Based on the configurations, the egress switching device or switch in the headquarters can automatically respond to the UDP packets. The egress switching device or switch of the branch processes the sent and received packets, and calculates the UDP jitter. To learn the performance in different periods of time, you also need to configure scheduling functions, such as periodically start/stop and repeated running, for the RNS.

9.2.2 Locating Network Faults

Scenario

On the campus network as shown in Figure 8- 1, Student 1 reports a Web server access failure, Student 3 reports an Internet access failure, and Student 6 reports an email sending/receiving failure.

Figure 8- 1



Deployment

- The administrator directly enables the DNS function on the access switch of the dormitories to test whether the domain name service (DNS) server is faulty. If DNS fails, an ICMP echo packet is automatically triggered to test whether the Web server is reachable.
- When a fault occurs, the administrator only needs to start a test, and the subsequent tests can be automatically triggered. Then, the administrator can check the test results to locate the fault, which greatly reduces the workload for the administrator.

9.3 Features

Basic Concepts

↘ RNS Instance

An RNS instance can be treated as an RNS process. Before performing the RNS, you must create an RNS instance. In the RNS instance, you need to configure the RNS parameters, such as the test type, test destination address, and test frequency. The instance ID is globally unique.

Feature	Description
RNS Instance	Monitor the network connectivity, service availability, integrity of end-to-end connection, and service quality.
Track Support for the RNS	Track the test results and notify the related module of the results.

9.3.1 RNS Test

Monitor the network connectivity, service availability, integrity of end-to-end connection, and service quality. For example, test whether the DNS function of the device is normal. Currently, the RNS supports the following types of tests: ICMP echo, DNS, and TCP.

Working Principle

ICMP Echo Test

ICMP echo is a basic function of the RNS, and is implemented in compliance with the RFC 2925. An ICMP packet is sent to check whether the destination is reachable and to calculate the network response time and packet loss rate.

An ICMP echo request packet is sent to the destination IP address based on the preset test time and frequency. Upon receipt of the ICMP echo request packet, an ICMP echo reply packet is returned from the destination IP address. Through the ICMP echo test, the response time and packet loss rate is calculated based on the information relating to the received ICMP echo reply packet, for example, the receipt time and number of packets. In this way, the current network performance and status are reflected. The ICMP echo test results and historical records will be recorded, and you can use the command line to display them.

The prerequisite for a successful ICMP echo test is that destination devices can correctly respond to ICMP echo request packets.

TCP Test

The TCP test is used to test the availability of a TCP connection. A TCP connection can be established based on the configured destination IP address and port ID. If the TCP connection is established successfully, the test is successful; otherwise, the test fails.

DNS Test

In a DNS test, a DNS client is simulated to send a domain name resolution request to a specified DNS server. You can determine whether the DNS server is available and the domain name resolution speed by checking the domain name resolution result and the time required for domain name resolution. In the DNS test, the domain name resolution process is simulated, and the mapping between the resolved domain name and the IP address is not saved. The DNS test results and historical records will be recorded in the test group. You can use the command line to check the test results and historical records.

Procedure for Configuring an RNS Instance Test

1. Create an instance and configure the test based on the test type.
2. Start the instance.
3. Use the RNS instance to construct a packet of the specific test type and send the packet to the peer end.
4. Upon receipt of the test packet, the peer end returns a reply packet of a corresponding type.
5. The RNS instance calculates the packet loss rate and round trip time based on whether a reply packet is received and the time of reply packet receipt.
6. Use the show or debug command to check the test result.

The preceding describes general procedures for RNS instance tests. For details about configuration, see the following sections.

Related Configuration

Configuring the Test Repeat Interval

By default, the test repeat interval is 60s.

In RNS configuration mode, run the **frequency** *millisecond* command to configure the test repeat interval.

Configure the frequency based on the following formula to ensure correct test calculation.

(**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configuring the Test Timeout

The default timeout varies according to the test type. You can run the **show ip rns configuration** command to display the timeout of a specific test type.

In RNS configuration mode, run the **timeout** *milliseconds* command to configure the timeout of an instance.

Configure the timeout based on a formula. For details, see the "Usage Guide" of the **frequency** command.

Configure the test time threshold.

Configuring the Test Threshold

By default, the test threshold is 5,000 ms.

In RNS configuration mode, run the **threshold** *milliseconds* command to configure the instance test threshold.

Configure the threshold based on a formula. For details, see the "Usage Guide" of the **frequency** command.

Configuring a Tag for the Test

No default configuration is available.

In RNS configuration mode, run the **tag** *text* command to configure a test tag.

You can run the **tag** command to specify a tag to identify the test.

Configuring the Protocol Payload Size

The default protocol payload size varies with the test type. By default, the protocol payload size is the minimum or appropriate size for protocol packets of the corresponding test type.

In RNS configuration mode, run the **request-data-size** *bytes* command to configure the protocol payload size.

Perform this configuration in IP RNS configuration mode.

Configuring the TOS Field of the Test Packet

By default, the TOS is 0.

In RNS configuration mode, run the **tos** number command to configure the TOS field in the IPv4 header of RNS test packets.

Configuring the VRF

No default configuration is available.

In RNS configuration mode, run the **vrf** *vrf-name* command to virtual routing and forwarding (VRF) for the RNS instance.

9.3.2 Track Support for the RNS

Objects that can be tracked include: test result of an RNS instance, RNS list status, link status on an interface, and track list status. When the track status changes, an action of other modules is triggered.

Working Principle

The test result of an RNS instance is tracked as follows:

- Configure a track object for tracking the test result of an RNS instance.
- When the test result of the RNS instance changes, the RNS module sends a status change message to the track module.
- The track module receives the test result. After the preset delay, if the test result remains unchanged, the status of the track object is modified, and the module of the track object is notified of the modification. If the test result recovers within the period, the status of the track object is not modified and the corresponding module is not notified.

Related Configuration

Configuring a Track Object for Tracking the Link Status of An Interface

By default, the function of tracking the link status of an interface is disabled.

Run the **track interface line-protocol** command to configure a track object, which is used to track the link status of an interface.

If the link status of the interface is UP, the status of the track object is UP. If the link status of the interface is DOWN, the status of the track object is also DOWN.

Configuring a Track Object for Tracking the Test Result of an RNS Instance

By default, the function of tracking the test result of an RNS instance is disabled.

Run the **track rns** command to configure a track object, which is used to track the test result of an RNS instance. The RNS instance ID ranges from 1 to 500.

If the RNS test succeeded, the track object is in Up state. If the RNS test failed, the track object is in Down state.

Configuring a Track Object for Tracking the Test Result of an RNS List

By default, the function of tracking the test result of an RNS list is disabled.

Run the **track rns-list** command to configure a track object, which is used to track the test result of an RNS list. The RNS instance ID ranges from 1 to 500.

The result can be the AND or OR operation result of all member status. If the result of this track object is set to the OR operation result of all member status, and the OR result of the status of all the tracked RNS objects is UP, the status of this track object is UP. If the OR result of the status of all the tracked RNS objects is DOWN, the status of this track object is also DOWN. If the result of this track object is set to the AND operation result of all member status, and the AND result of the status of all the tracked RNS objects is UP, the status of this track object is UP. If the AND result of the status of all the tracked RNS objects is DOWN, the status of this track object is also DOWN.

Configuring a Track Object for Tracking the Status of a Track List

By default, the function of tracking the status of a track list is disabled.

Run the **track list** command to configure a track object, which is used to track the status of a track list. The result can be the AND or OR operation result of all member status.

If the result of this track object is set to the OR operation result of all member status, when all RNS tests succeeded, the track object is in Up state. If one RNS test failed, the track object is in Down state. If the result of this track object is set to the OR operation result of all member status, when all RNS tests failed, the track object is in Down state. If one RNS test succeeded, the track object is in Up state.

Configuring a Track List Member

By default, no member is configured for the track list.

Run the **object** command to configure a track list member. The status of the member can be the same as or contrary to that of the corresponding track object.

Adjusting the Delay for Notifying the Status Change of a Track Object

By default, the delay for notifying the status change of a track object is 0.

Run the **delay** command to adjust the delay for track notification, including the delay for notifying the status change of a track object from UP to DOWN and the delay for notifying the status change of a track object from DOWN to UP. The delay ranges from 0 to 180. The unit is second.

A longer delay indicates that it takes more time before the module that is concerned with the track object is notified of the status. A shorter delay indicates that it takes less time before the module that is concerned with the track object is notified of the status.

9.4 Configuration

Configuration Item	Description and Command	
Configuring RNS Basic Functions	(Mandatory) It is used to configure basic function parameters of the RNS.	
	ip rns	Supports detailed configuration and brief configuration. ● Detailed configuration: An RNS operation object is defined, and used as the configuration ID for subsequent tests and parameters. ● Brief configuration: Subsequent configuration is not required, and tests can be started in one step. Currently, ICMP echo, DNS, and TCP tests can be started in one step.
	ip rns reaction-configuration	Configures the proactive threshold monitoring and triggering mechanism of the RNS test.
	ip rns reaction-trigger	Triggers another type of the RNS test in pending state when the monitoring threshold exceeds the expectation during an RNS test.
	ip rns schedule	Configures the scheduling method, start time, and life time of an RNS test.
	ip rns restart	Restarts an RNS test.
	ip rns reset	Clears all the IP RNS configurations.
Configuring the ICMP Echo Test	(Optional) It is used to implement the ICMP echo test.	
	icmp-echo	Creates an ICMP echo test instance.
	request-data-size	Configures the protocol payload size.
	frequency	Configures the test repeat interval.
	tag	Configures a tag.
	threshold	Configures the test time threshold.
	timeout	Configures the test timeout.
	tos	Configures the TOS field in the IPv4 header of test packets.
	vrf	Configure the VRF of a test.

Configuration Item	Description and Command	
Configuring the DNS Test	(Optional) It is used to implement the DNS test.	
	dns	Creates a DNS test instance.
	frequency	Configures the test repeat interval.
	tag	Configures a tag.
	threshold	Configures the test time threshold.
	timeout	Configures the test timeout.
	tos	Configures the TOS field in the IPv4 header of test packets.
	vrf	Configures the VRF of a test.
Configuring the TCP Connect Test	(Optional) It is used to implement the TCP connect test.	
	tcp-connect	Creates a TCP test instance.
	request-data-size	Configures the protocol payload size.
	frequency	Configures the test repeat interval.
	tag	Configures a tag.
	threshold	Configures the test time threshold.
	timeout	Configures the test timeout.
	tos	Configures the TOS field in the IPv4 header of test packets.
	vrf	Configure the VRF of a test.
Configuring the Track Support for the RNS	(Optional) It is used to configure the track support for other test modules.	
	track rns	Configures a track object for tracking the test result of an RNS instance.
	track rns-list	Configures a track object for tracking the status of an RNS list.
	track interface line-protocol	Configures a track object for tracking the link status of an interface.
	track list	Configures a track object for tracking the status of a track list.
	object	Configures a member object for a track list object.
	delay	Configures the delay for notifying the status change of a track object.

9.4.1 Configuring RNS Basic Functions

Configuration Effect

- Detailed configuration: Configures an RNS instance to complete basic configuration of the RNS instance.
- Brief configuration: Configure and start an RNS instance at a time. (Optional)

Notes

- In detailed configuration mode, if you do not configure the test type after entering the IP RNS mode by running the command, the RNS instance will not be created.

- In detailed configuration mode, after configuring an RNS instance, you need to run the **ip rns schedule** command to configure the startup policy; otherwise, the test will not be implemented.

Configuration Steps

Defining an RNS Operation Object

- Mandatory.
- Unless otherwise required, define an RNS operation object on each switch.
- Brief configuration is optional.

Configuring the Proactive Threshold Monitoring and Triggering Mechanism for an RNS Test

- Perform this configuration if it is required to configure the proactive threshold monitoring and triggering mechanism for the test.
- Perform this configuration on every switching device unless otherwise required.

Enabling an RNS Instance to Trigger Another RNS Instance

- Perform this configuration if it is required to trigger another RNS test in pending state when the monitoring threshold exceeds the expectation during an RNS test.
- If schedule parameters are not configured for the triggered RNS instance, the default schedule parameters are applied.
- Unless otherwise required, apply this configuration to each switch.

Configuring Schedule Parameters of an RNS Instance

- Perform this configuration on every switching device unless otherwise required.
- In the case of brief configuration, this command is already configured using the default values, and manual configuration is not required.

Restarting an RNS Instance

- Perform this configuration, or directly run the **ip rns schedule X start-time now** command if it is required to restart an IP RNS instance in pending state.

Clearing Configurations of All RNS Instances

- Perform this configuration if it is required to clear configurations of all the IP RNS instances, for example, when a lot of instances are configured but configurations are found incorrect.

Verification

- Run the **show ip rns configuration** command to display configurations of RNS instances.

Related Commands

Defining an IP RNS Operation Object

Command	**ip rns** *operation-number* [{**dns** *destination-hostname* **name-server** *ip-address* **|** **icmp-echo** *destination-ip-address* **|** **tcp-connect** *destination-ip-address port-number*} [**frequency** *seconds*] [**timeout** *milliseconds*] [**threshold** *milliseconds*]]]

Parameter Description	*operation-number*: Indicates the RNS instance ID. The value ranges from 1 to 500. For details about configuration of **frequency**, **timeout**, and **threshold**, see the configuration of the specific test type.
Command Mode	Global configuration mode
Usage Guide	Currently, the RNS supports only IPv4-related tests, but not IPv6-related tests. At most 500 tests can be configured, depending on the performance of devices. The test function is only a value-added function. When a large number of tests are configured and consume a lot of system resources, the test function may be temporary disabled to ensure normal operation of core services, such as route forwarding. Detailed configuration (executing mandatory items of **ip rns** *operation-number*): Run this command and enter the IP-RNS configuration mode. In this mode, you can define various test types. If the test type is not configured, the RNS test is not created. After configuring an RNS test, you must run the **ip rns schedule** command to configure its schedule parameters; otherwise, the test cannot be conducted. After configuring the type of an RNS test, you can run the **ip rns** command to enter the mode of the test type. To modify the type of an RNS instance, you need to first delete the RNS instance by running the **no ip rns** command in global configuration mode. Brief configuration (executing the optional test items that proceeds **ip rns**):After optional items are executed, it is equivalent that **ip rns operation-number**, **ip rns schedule**, detailed test configuration (such as the ICMP echo test), **frequency**, **timeout**, and **threshold** are executed according to the logical sequence. Among these commands, the **ip rns schedule** command is executed to start a test by using the **start-time now life forever** parameter. For details about restrictions of other configuration items, see the related description in the detailed configuration. Similarly, to modify a briefly configured test, you need to first delete this RNS instance by running the **no ip rns** command in global configuration mode.

Configuring the Proactive Threshold Monitoring and Triggering Mechanism for the Test

Command	**ip rns reaction-configuration** *operation-number* **react** *monitored-element*[**action-type** *option*][**threshold-type** {**average** [*number-of-measurements*] | **consecutive** [*occurrences*] | **immediate** | **never**| **xofy** [*x-value y-value*] }] [**threshold-value** *upper-threshold lower-threshold*]
Parameter Description	*operation-number*: Indicates the RNS instance ID. The value ranges from 1 to 500. *monitored-element*: Specifies the monitored element. **action-type** *option*: Indicates the action taken after the test is triggered. **average** [*number-of-measurements*]: Indicates that the subsequent associated actions are triggered if the average of *number-of-measurements* of the monitored element exceeds the threshold. **consecutive** [*occurrences*]: Indicates that the test is triggered if the consecutive number of *occurrences* of the monitored element exceeds the threshold. The default value of *occurrences* is 5. The value ranges from 1 to 16. **immediate**: Indicates that the test is triggered immediately after the monitored element exceeds the threshold. **never**: Indicates that the test is never triggered. **xofy** [*x-value y-value*]: Indicates that results of X tests exceed the threshold in the last Y tests. The default values of X and Y are 5. The value of X or Y ranges from 1 to 16. **threshold-value** *upper-threshold lower-threshold*: Indicate the upper and lower thresholds. ● When ***monitored-element*** *is* **rtt**, the thresholds are the time. For default values, see "Usage Guide". The value ranges from 0 to 60,000 ms. ● Note that you do not need to configure **threshold-value** when **react** is set to **timeout**.

Command Mode	Global configuration mode
Usage Guide	You can configure multiple thresholds for one RNS test to monitor different elements. The following table provides the mapping between test types and monitored elements.

monitored-element	icmp-echo	dns
timeout		
rtt		

The following table lists the default thresholds of each monitored element.

Monitored Element	Upper Threshold	Lower Threshold
timeout	-	-
rtt	5000ms	0ms

Enabling an RNS Instance to Trigger Another RNS Instance

Command	**ip rns reaction-trigger** *operation-number target-operation*
Parameter Description	*operation-number*: Indicates the number of the source RNS instance that triggers the action. The value ranges from 1 to 500. *target-operation*: Indicates the number of the target RNS instance that is triggered. The value ranges from 1 to 500.
Command Mode	Global configuration mode
Usage Guide	The trigger function is generally used in network fault diagnosis scenario. In a common scenario, you do not need to configure the trigger function.

Configuring Schedule Parameters of an RNS Instance

Command	**ip rns schedule** *operation-number*[**life**{ **forever**|*seconds*}] [**start-time**{ *hh:mm*[*:ss*] [*month day*|*daymonth*] |**pending**|**now**|**after***hh:mm:ss}*] [**recurring**]
Parameter Description	*operation-number*: Indicates the number of the RNS operation. The value ranges from 1 to 500. **lifeforever**: Indicates that the RNS operation life time is valid forever. **life** *seconds*: Indicates the running time of the RNS instance in seconds. *hh:mm[:ss]*: Indicates the start time of the RNS instance, in 24-hour format. *month*: Indicates the start month of the RNS instance. The default value is the current month. *day*: Indicates the start date of the RNS instance. The default value is the current date. **pending**: Indicates that the start time of the RNS instance is not defined, which is the default. **now**: Indicates that the operation start time is now, that is, the operation starts now. **after***hh:mm:ss*: Indicates that the RNS instance starts after a delay of hh:mm:ss. **recurring**: Indicates whether the RNS instance starts at the same time every day.

Command Mode	Global configuration mode
Usage Guide	If the schedule parameters of an RNS instance have been configured by running the **ip rns schedule** command, parameters cannot be modified during running. To modify the configuration, you need to run the **no ip rns schedule** command to delete the schedule parameters. life { *seconds* } indicates the running time of the RNS instance. That is, the test stops after a period of time in seconds.

Restarting an RNS Test by Running the ip rns restart Command

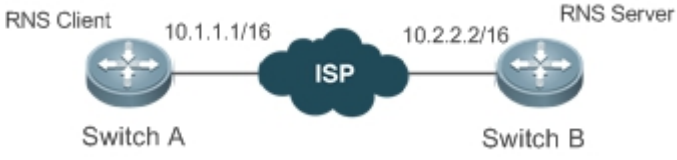
Command	**ip rns restart** *operation-number*
Parameter Description	*operation-number*: Indicates the number of the RNS instance. The value ranges from 1 to 500.
Command Mode	Global configuration mode
Usage Guide	This command restarts an RNS test for which the scheduling policy is configured and is in pending state. This command is invalid for an RNS test for which the scheduling policy is not configured.

Clearing Configurations of All the IP RNS Instances by Running the ip rns reset Command

Command	**ip rns reset**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	This command clears configurations of all the IP RNS instances. It is used only in extreme cases, for example, when a lot of RNS tests are configured but the configurations are found incorrect.

Configuration Example

Configuring RNS Basic Functions

Scenario **Figure 8- 2**	
Configuration Steps	● Configure instance 1 on Switch A. ● Configure the scheduling method, start time, and life time of instance 1. ● Configure the proactive threshold monitoring and triggering mechanism of instance 1. ● Trigger instance 2 in pending state when the monitoring threshold of instance 1 exceeds the expectation.
Switch A	A# configure terminal

	A(config)# ip rns 1 A(config-ip-rns)#icmp-echo 10.1.1.1 A(config-ip-rns-icmp-echo)#exit A(config)ip rns schedule 1 start-time now life forever A(config)ip rns reaction-configuration 1 react timeout threshold-type immediate action-type trigger A(config)ip rns reaction-trigger 1 2
Verification	Run the **show ip rns configuration** command to display the instance configurations. Router#show ip rns configuration 1 Entry number: 1 Tag: fs555 Type of operation to perform: icmp-echo Operation timeout (milliseconds): 5000 Operation frequency (milliseconds): 60000 Threshold (milliseconds): 5000 Recurring (Starting Everyday): FALSE Life (seconds): 3500 Next Scheduled Start Time:Start Time already passed Target address/Source address: 2.2.2.3/0.0.0.0 Request size (ARR data portion): 36

9.4.2 Configuring the ICMP Echo Test

Configuration Effect

Create an ICMP echo test instance.

Notes

- The RNS basic functions must be configured.

Configuration Steps

Creating an ICMP Echo Test Instance

- Mandatory.
- Unless otherwise required, create ICMP echo test instances on each switch.

Configuring Common Optional Parameters of the Test

- Mandatory if common optional parameters of the test, for example, the repeat interval, tag, time threshold, timeout, and TOS, are required to be changed..

- Perform this configuration on every switching device unless otherwise required.

Configuring the Protocol Payload Size

- Perform this configuration if it is required to change the protocol payload size of the test.
- Perform this configuration on every switching device unless otherwise required.

Verification

- Run the **show ip rns configuration** command to display the instance configurations.

Related Commands

Creating an ICMP Echo Test Instance

Command	**icmp-echo** { **oob** { *destination-ip-address* | *destination-hostname* [**name-server** *ip-address*] } [**source-ipaddr** *ip-address*] **via** *type num* **next-hop** *ip-address* } | { { *destination-ip-address* | *destination-hostname* [**name-server** *ip-address*] } [**source-ipaddr** *ip-address* | **source-interface** *interface-type interface-number*] [**out-interface** *type num* [**next-hop** *ip-address*]] }
Parameter Description	**oob**: Indicates the test on the MGMT interface. *destination-ip-address*: Indicates the destination IP address. *destination-hostname*: Indicates the destination host name. **name-server** *ip-address*: Specifies the DNS server when the destination host name is configured. By default, the DNS server configured by using the **ip name-server** command is used for address resolution. **source-ipaddr** *ip-address*: Indicates the source IP address. **source-interface** *interface-type interface-number*: Indicates the source interface. **out-interface** *type num*: Specifies the outgoing interface (non-MGMT interface) of the test packet. **via** *type num*: Specifies the MGMT interface as the outgoing interface of the test packet. **next-hop** *A.B.C.D*: Indicates the IP address of the next hop.
Command Mode	IP RNS configuration mode (config-ip-rns)
Usage Guide	After an ICMP echo test is started, the system sends an ICMP echo request packet to test whether the device is connected to the target host. After an ICMP-Echo test instance is created, the system enters the IP RNS ICMP echo mode. By default, the protocol payload size of an ICMP echo request packet is 36 bytes. You can run the **request-data-size** command to change the packet size. You need to configure the RNS test type (for example, ICMP echo and DNS) before configuring parameters. To modify the type of an RNS instance, you need to delete the RNS instance by running the **no ip rns** command in global configuration mode.

Configuring the Protocol Payload Size of an RNS Instance

Command	**request-data-size** *bytes*
Parameter Description	*bytes*: Indicates the bytes of a test packet. The minimum and maximum bytes vary with the test type. You need to configure this parameter based on the command prompt in corresponding test mode.
Command	IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo)

Mode	
Usage Guide	This command is used to stuff some bytes in the test packet so that large packets can be used for the test.

Configuring the Test Repeat Interval

Command	**frequency** *milliseconds*
Parameter Description	*milliseconds*: Indicates the packet sending interval in ms. The default value is 60,000 ms. The value ranges from 10 to 604,800,000. The maximum value is one week.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	After an RNS instance is started, tests are conducted periodically. You can run the **frequency** command to specify the repeat interval. You need to configure the frequency based on the following formula to ensure correct test calculation. (**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configuring a Tag for an RNS Instance

Command	**tag** *text*
Parameter Description	*text*: Sets the test tag. The value is a string of up to 79 characters.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	This command specifies a tag for a test, which is often used to indicate the function of the test.

Configuring the Time Threshold for an RNS Instance

Command	**threshold** *milliseconds*
Parameter Description	*milliseconds*: Indicates the time threshold for the test. The value ranges from 0 to 60,000, in the unit of milliseconds. The default value is 5,000.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	Configure the threshold based on the following formula to ensure correct test calculation. (**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configuring the Timeout for an RNS Instance

Command	**timeout** *millisecond*
Parameter	*millisecond*: Indicates the test timeout. The value ranges from 10 to 604,800,000. The unit is ms. The default timeout

Description	varies according to the test type.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	Configure the timeout based on the following formula to ensure correct test calculation. (**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configuring the TOS Field in the IPv4 Packet Header of an IP RNS Test

Command	**tos** *number*
Parameter Description	*number*: Sets the TOS field in the IPv4 header of test packets. The value ranges from 0 to 255. The default value is 0.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	TOS is a 8-bit field in the IPv4 packet header. By setting the TOS, you can control the priority of the test packet. For different TOS fields, the processing priorities are different on the intermediate routers.

Configuring the VRF of an RNS Test

Command	**vrf** *vrf-name*
Parameter Description	*vrf-name*: Specifies the VRF name.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	This command specifies the VRF of the test packet.

Configuration Example

Figure 8- 3	
	Configure RNS instance 1 and related parameters on Switch A.
Switch A	A# configure terminal A(config)# ip rns 1 A(config-ip-rns)#icmp-echo 10.2.2.2 A(config-ip-rns-icmp-echo)#exit A(config)#ip rns schedule 1 start-time now life forever
	Run the **show ip rns configuration** command to display the instance configurations.

Switch A	A#show ip rns configuration 1 Entry number: 1 Tag: Type of operation to perform: icmp-echo Operation timeout (milliseconds): 5000 Operation frequency (milliseconds): 60000 Threshold (milliseconds): 5000 Recurring (Starting Everyday): FALSE Life (seconds): foerver Next Scheduled Start Time:Start Time already passed Target address/Source address: 10.2.2.2/0.0.0.0 Request size (ARR data portion): 36

9.4.3 Configuring the DNS Test

Configuration Effect

Create a DNS test instance.

Notes

- The RNS basic functions must be configured.

Configuration Steps

Creating a DNS Test Instance

- Mandatory.
- Unless otherwise required, create DNS test instances on each switch.

Configuring Common Optional Parameters of the Test

- Mandatory if common optional parameters of the test, for example, the repeat interval, tag, time threshold, timeout, and TOS, are required to be changed.
- Perform this configuration on every switching device unless otherwise required.

Verification

- Run the **show ip rns configuration** command to display the instance configurations.

Related Commands

Creating a DNS Test Instance

Command	**dns** { **oob** *destination-hostname* **name-server** *ip-address* [**source-ipaddr** *ip-address*] **via** *type num* **next-hop** *ip-address* } | { *destination-hostname* **name-server** *ip-address* [**source-ipaddr** *ip-address*] [**out-interface** *type num* [**next-hop** *ip-address*]] }
Parameter	**oob**: Indicates the test on the MGMT interface. *destination-hostname*: Indicates the destination host name.

Description	**name-server** *ip-address*: Indicates the DNS IP address. **source-ipaddr** *ip-address:* Indicates the source IP address. **out-interface type** *num*: Specifies the outgoing interface (non-MGMT interface) of the test packet. **via** *type num*: Specifies the MGMT interface as the outgoing interface of the test packet. **next-hop** *ip-address*: Indicates the IP address of the next hop when the outgoing interface is specified.
Command Mode	IP RNS configuration mode (config-ip-rns)
Usage Guide	After a DNS test is started, the system sends a DNS parsing request packet to test whether the device is connected to the target host. After a DNS test instance is created, the system enters the IP RNS DNS mode. You need to configure the RNS test type before configuring parameters. To modify the type of an RNS instance, you need to delete the RNS instance by running the **no ip rns** command in global configuration mode.

Configuring the Test Repeat Interval

Command	**frequency** *milliseconds*
Parameter Description	*milliseconds*: Indicates the packet sending interval in ms. The default value is 60,000 ms. The value ranges from 10 to 604,800,000. The maximum value is one week.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	After an RNS instance is started, tests are conducted periodically. You can run the **frequency** command to specify the repeat interval. You need to configure the frequency based on the following formula to ensure correct test calculation. (**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configuring a Tag for an RNS Instance

Command	**tag** *text*
Parameter Description	*text*: Sets the test tag. The value is a string of up to 79 characters.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	This command specifies a tag for a test, which is often used to indicate the function of the test.

Configuring the Time Threshold for an RNS Instance

Command	**threshold** *milliseconds*
Parameter Description	*milliseconds: Indicates the time threshold for the test. The value ranges from 0 to 60,000, in the unit of milliseconds. The default value is 5,000.*
Command	IP RNS DNS configuration mode (config-ip-rns-dns)

Mode	IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	Configure the threshold based on the following formula to ensure correct test calculation. (**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configuring the Time Threshold for an RNS Instance

Command	**timeout** *millisecond*
Parameter Description	*millisecond*: Indicates the test timeout. The value ranges from 10 to 604,800,000. The unit is ms. The default timeout varies according to the test type.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	Configure the timeout based on the following formula to ensure correct test calculation. (**frequency** *milliseconds*) > (**timeout** *milliseconds*) >= (**threshold** *milliseconds*)

Configures the TOS Field in the IPv4 Header of Test Packets

Command	**tos** *number*
Parameter Description	*number*: Sets the TOS field in the IPv4 header of test packets. The value ranges from 0 to 255. The default value is 0.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	TOS is a 8-bit field in the IPv4 packet header. By setting the TOS, you can control the priority of the test packet. For different TOS fields, the processing priorities are different on the intermediate routers.

Configuring the VRF of an RNS Test

Command	**vrf** *vrf-name*
Parameter Description	*vrf-name*: Specifies the VRF name.
Command Mode	IP RNS DNS configuration mode (config-ip-rns-dns) IP RNS ICMP echo configuration mode (config-ip-rns-icmp-echo) IP RNS TCP configuration mode (config-ip-rns-tcp)
Usage Guide	This command specifies the VRF of the test packet.

Configuration Example

Scenario **Figure 8- 4**	
Configuration Steps	Configure RNS instance 1 and related parameters on Switch A.
Switch A	A# configure terminal A(config)# ip rns 1 A(config-ip-rns)# dns www.fs.com name-server 10.2.2.2 A(config-ip-rns-dns)#exit A(config)ip rns schedule 1 start-time now life forever
Verification	Run the **show ip rns configuration** command to display the instance configurations.
Switch A	A#show ip rns configuration 1 Entry number: 1 Tag: Type of operation to perform: dns Operation timeout (milliseconds): 5000 Operation frequency (milliseconds): 60000 Threshold (milliseconds): 5000 Recurring (Starting Everyday): FALSE Life (seconds): forever Next Scheduled Start Time:Start Time already passed Target host name: www.fs.com Name Server: 10.2.2.2

Common Errors

- The DNS IP address is incorrect.

9.4.4 Configuring the TCP Connect Test

Configuration Effect

Create a TCP test instance to implement a TCP connect test.

Notes

- The RNS basic functions must be configured.
- The target host must be able to respond to the TCP connection request.

Configuration Steps

Creating a TCP Test Instance

- (Mandatory) Unless otherwise required, create TCP test instances on each switch.

Configuring Common Optional Parameters of the Test

- Mandatory if common optional parameters of the test, for example, the repeat interval, tag, time threshold, timeout, and TOS, are required to be changed.
- Perform this configuration on every switch unless otherwise required.

Verification

- Run the **show ip rns configuration** command to display the instance configurations.

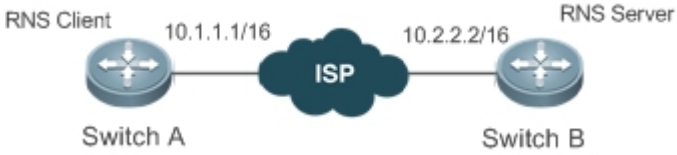
Related Commands

Creating a TCP Test Instance

Command	**tcp-connect** { *destination-ip-address* | *destination-hostname* [**name-server** *ip-address*] } *port-number*
Parameter Description	*destination-ip-address*: Destination IP address *destination-hostname*: Indicates the destination host name. **name-server***ip-address*: Indicates the IP address of the DNS server. *port-number*: Indicates the TCP port to be tested.
Command Mode	IP RNS configuration mode (config-ip-rns)
Usage Guide	After a TCP test is started, the system tries to establish a TCP connection to the specified port of a specified host to test whether the specified port is available. After a TCP IP RNS instance is created, the system enters the IP RNS TCP mode.

Commands for configuring common optional parameters of a test, including **frequency**, **tag**, **threshold**, **timeout**, and **tos**, are provided in the description about configuring an ICMP echo test, and therefore omitted here.

Configuration Example

Scenario	
Configuration Steps	Configure RNS instance 1 and related parameters on Switch A.
Switch A	A# configure terminal A(config)# ip rns 1 A(config-ip-rns)# tcp-connect 10.2.2.2 8000 A(config-ip-rns-tcp)#exit A(config)ip rns schedule 1 start-time now life forever
Verification	Run the **show ip rns configuration** command to display the instance configurations.

Switch A	A#show ip rns configuration 1 Entry number: 1 Tag: Type of operation to perform: tcp-connect Operation timeout (milliseconds): 5000 Operation frequency (seconds): 60 Threshold (milliseconds): 5000 Recurring (Starting Everyday): FALSE Life (seconds): forever Next Scheduled Start Time:Start Time already passed Target Address: 10.2.2.2 Target Port: 8000

Common Errors

- The target host does not respond to the TCP connection request.
- The TCP port for the RNS test is incorrectly configured.

9.4.5 Configuring the Track Support for the RNS

Configuration Effect

- Configure the track function to track the test result of an RNS instance.
- Configure the track function to track the link status of an interface.
- Configure the track function to track the status of a track list.
- Configure the track function to track the status of an RNS list.

Notes

- To configure the track function to track the test result of an RNS instance, you need to configure the related RNS instance.
- To configure the track function to track the link status of an interface, you need to configure the related interface.
- To configure the track function to track the status of a track list, you need to configure the members for the related track list.
- To configure the track function to track the status of an RNS list, you need to configure the members for the related RNS list.

Configuration Steps

Configuring a Track Object

- Perform this operation if it is required to create a track object.
- The following four methods are available to create a track object:

- Create a track object for tracking the test result of an RNS instance: Perform this configuration on every switching device unless otherwise required.

- Create a track object for tracking the link status of an interface: Perform this configuration on every switching device unless otherwise required.
- Create a track object for tracking the status of a track list: Perform this configuration on every switching device unless otherwise required.
- Create a track object for tracking the status of an RNS list: Perform this configuration on every switching device unless otherwise required.

Configuring the Notification Delay of a Track Object

- Perform this configuration if it is required to delay notification of the status change of a track object.
- Delay for notifying the status change of a track object includes the delay for notifying the status change of a track object from UP to DOWN and the delay for notifying the status change of a track object from DOWN to UP. You can configure either delay or both of delays.
- Perform this configuration on every switching device unless otherwise required.

Configuring a Track Member

- Perform this configuration if it is required to configure a track object for tracking the status of a track list.
- When configuring a track member, you can set the status of a member meeting conditions to UP or DOWN.
- Perform this configuration on every switching device unless otherwise required.

Verification

Observe the status of a track object when the status of the track object (such as test results of an RNS instance, link status of an interface, or status of a track list) changes.

- After the preset delay, run the **show track** command to check whether the current track status changes.

Related Commands

Configuring a Track Object for Tracking the Link Status of An Interface

Command	**track** *object-number* **interface** *interface-type interface-number* **line-protocol**
Parameter Description	*object-number*: Indicates the number of a track object. The value ranges from 1 to 700. *Interface-type interface-number*: Indicates the interface type and interface number.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure a track object for tracking the link status of an interface. When the link status of the interface is UP, the status of the corresponding track object is UP.

Configuring a Track Object for Tracking the Test Result of an RNS Test

Command	**track** *object-number* **rns** *entry-number*
Parameter Description	*object-number*: Indicates the number of a track object. The value ranges from 1 to 700. *entry-number*: Indicates the number of an RNS instance. The value ranges from 1 to 500.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure a track object for tracking the result of an RNS test. If the test succeeded, the track object

	is in Up state.

Configuring a Track Object for Tracking the Status of a Track List

Command	**track** *object-number* **list boolean** { **and** | **or** }
Parameter Description	*object-number*: Indicates the number of a track object. The value ranges from 1 to 700.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure a track object for tracking the status of a track list. The result can be the AND or OR operation result of all member status.

Configuring a Track Member

Command	**object** *object-number*[**not**]
Parameter Description	*object-number*: Indicates the number of a track object. The value ranges from 1 to 700.
Command Mode	Track configuration mode
Usage Guide	Run this command to configure a member for a track list. The number of track list members that can be configured is restricted only by the capacity of track objects.

Configuring a Track Object for Tracking the Status of an RNS List

Command	**track** *object-number* **rns-list** *men-list*{ **and | or** }
Parameter Description	*object-number*: Indicates the number of a track object. The value ranges from 1 to 700. men-list: Indicates the RNS list that is tracked. **mem-list** can be an RNS instance or a series of RNS instances. If **mem-list** is a series of RNS instances, the format is as follows: Smallest RNS ID–Greatest RNS ID, for example, 10–20. The RNS ID ranges from 1 to 500.
Command Mode	Global configuration mode
Usage Guide	Run this command to configure a track object for tracking the status of an RNS list. The result can be the AND or OR operation result of all member status.

Configuring the Notification Delay of a Track Object

Command	**delay** { **up** *seconds* [**down** *seconds*] **|** [**up** *seconds*] **down** *seconds* }
Parameter Description	**up *seconds***: Specifies the delay for notifying the status change of a track object from DOWN to UP. The value ranges from 0 to 180. The unit is second. The default value is 0. **down** *seconds*: Specifies the delay for notifying the status change of a track object from UP to DOWN. The value ranges from 0 to 180. The unit is second. The default value is 0.
Command Mode	Track configuration mode
Usage Guide	When the status of a track object frequently changes, the status of the client that use this track object will frequently change as well. Using this command can delay notification of the status change of a track object. For example, if the status of a track

	object changes from UP to DOWN, and **delay down 10** is configured, the DOWN status of the track object is notified 10s later. If the status of the track objects changes to UP again within this period of time, no notification is sent. For the client that uses this track object, the status of the track object is always UP.

Displaying the Track Object Statistics

Command	**show track** [*object-number*]
Parameter Description	*object-number*: Indicates the number of a track object. The value ranges from 1 to 700. The default is all track objects.
Command Mode	Privileged EXEC mode
Usage Guide	Run this command to display statistics of track objects.

Configuration Example

Configuring Track Object3 for Tracking the Link Status of the Interface FastEthernet 1/0

Configuration Steps	● Configure a track object for tracking the link status of an interface. ● Configure the delay for notifying the status change from UP to DOWN.
	FS# configure terminal FS(config)# track 3 interface FastEthernet 1/0 line-protocol FS(config-track)# delay down 10 FS(config-track)# exit
Verification	Change the link status of the interface FastEthernet 1/0 to DOWN. ● Immediately check the status of the track object, and verify that the status is still UP. ● Check the status of the track object 10s later, and verify that the status changes to DOWN.
	FS# show track 3 Track 3 Interface FastEthernet 1/0 The state is Up, delayed Down (5 secs remaining) 1 change, current state last: 300 secs Delay up 0 secs, down 10 secs

Configuring Track Object 3 (When the status of track object 1 is UP, and the status of track object 2 is DOWN, the status of track object 3 is UP.)

Configuration Steps	● Configuring track object 1 and track object 2. ● Configure track object 3, and its members include track object 1 and track object 2.

	FS # config FS(config)#track 1 interface gigabitEthernet 0/0 line-protocol FS(config-track)#delay up 20 down 40 FS(config-track)#exit FS(config)# FS(config)#track 2 interface gigabitEthernet 0/1 line-protocol FS(config-track)#delay down 30 FS(config-track)#exit FS(config)# track 3 list Boolean and FS(config-track)#object 1 FS(config-track)#object 2 not FS(config-track)# exit
Verification	When the status of track objects 1 and 2 change, check the status of track object 3. ● When the status of track object 1 changes from DOWN to UP, and the status of track object 2 remains DOWN, verify that the status of track object 3 changes from DOWN to UP. ● When the status of track object 1 remains UP, and the status of track object 2 changes from DOWN to UP, verify that the status of track object 3 changes from UP to DOWN.
	FS# show track 3 Track 3 List boolean and Object 1 Object 2 not The state is Down 1 change,current state last:10 secs Delay up 0 secs,down 0 secs

Configuring Track Object 5 for Tracking the Test Result of RNS Instance 7

Configuration Steps	● Configure an RNS test. ● Configure a track object to track the result of the RNS test. ● Configure the delay for notifying the test result change from successful to unsuccessful, and the delay for notifying the test result change from unsuccessful to successful.
	FS# configure terminal FS (config)#ip rns 7

	FS (config-ip-rns)#icmp-echo 2.2.2.2 FS (config-ip-rns-icmp-echo)#exit FS (config)#ip rns schedule 7 start-time now life forever FS(config)# track 5 rns 7 FS (config-track)# delay up 20 down 30 FS (config-track)# exit
Verification	Let the test result of RNS instance 7 change from successful to unsuccessful. ● When the test result changes to unsuccessful, immediately check the status of track object 7, and verify that the status is still UP. ● Check the status of the track object 30s later, and verify that the status changes to DOWN.
	FS# show track 5 Track 5 Reliable Network Service 7 The state is Down 2 change, current state last: 10 secs Delay up 20 secs, down 30 secs

Configuring Track Object 5 for Tracking the Test Results of an RNS List (consisting of RNS Instances 1, 2–5, and 8)

Configuration Steps	● Configure and start an RNS test (see "RNS Configuration"). ● Configure a track object for tracking the test result of an RNS list. ● Configure the delay for notifying the test result change from UP to DOWN, and the delay for notifying the test result change from DOWN to UP.
	FS(config)# track 5 rns-list 1,2-5,8 and FS (config-track)# delay up 20 down 30 FS (config-track)# exit
Verification	Let the test result of one of the RNS instances 1, 2-5, and 8 changes from successful to unsuccessful. ● When the test result changes to unsuccessful, immediately check the status of track object 7, and verify that the status is still UP. ● Check the status of the track object 30s later, and verify that the status changes to DOWN.

	FS# show track 5 Track 5 rns-list 1,2-5,8 and The state is Down 2 change, current state last: 10 secs Delay up 20 secs, down 30 secs

Common Errors

- The track object for tracking an RNS test is configured, but the RNS test is not configured.
- The track object for tracking the link status of an interface is configured, but the corresponding interface is not configured.
- The track object for tracking the status of a track list, but no member of the RNS list is configured.
- The track object configured for tracking an RNS list, but the RNS test is not configured.

9.5 Monitoring

Displaying

Description	Command
Displays configurations of one or more RNS instances.	**show ip rns configuration** [*operation-number*]
Displays detailed statistics of one or more RNS instances.	**show ip rns collection-statistics** [*operation-number*]
Displays the current RNS status.	**show ip rns operational-state** [*operation-number*]
Displays the proactive threshold monitoring information of one or more RNS instances.	**show ip rns reaction-configuration** [*operation-number*]
Displays information about the test triggered by one or more RNS instances.	**show ip rns reaction-trigger** [*operation-number*]
Displays the brief statistics of one or more RNS instances.	**show ip rns statistics** [*operation-number*]
Displays the brief statistics of one or more track objects.	**show track** [*object-number*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs the track module.	**debug track** { **all | proc-event | rdnd-event | client** }
Debugs the RNS module.	**debug rns** { **all | interface | lib | rdnd-event | restart | rns_id** [**0, 500**] **| server** }

Chapter10 Network Management & Monitoring Configuration

1. Configuring SNMP
2. Configuring RMON
3. Configuring NTP
4. Configuring SNTP
5. Configuring SPAN-RSPAN
6. Configuring ERSPAN
7. Configuring sFlow

1 Configuring SNMP

1.1 Overview

Simple Network Management Protocol (SNMP) became a network management standard RFC1157 in August 1988. At present, because many vendors support SNMP, SNMP has in fact become a network management standard and is applicable to the environment where systems of multiple vendors are interconnected. By using SNMP, the network administrator can implement basic functions such as information query for network nodes, network configuration, fault locating, capacity planning, and network monitoring and management.

SNMP Versions

Currently, the following SNMP versions are supported:

- SNMPv1: The first official version of SNMP, which is defined in RFC1157.
- SNMPv2C: Community-based SNMPv2 management architecture, which is defined in RFC1901.
- SNMPv3: SNMPv3 provides the following security features by identifying and encrypting data.

7. Ensuring that data is not tampered during transmission.
8. Ensuring that data is transmitted from legal data sources.
9. Encrypting packets and ensuring data confidentiality.

Protocols and Standards

- RFC 1157, Simple Network Management Protocol (SNMP)
- RFC 1901, Introduction to Community-based SNMPv2
- RFC 2578, Structure of Management Information Version 2 (SMIv2)
- RFC 2579, Textual Conventions for SMIv2
- RFC 3411, An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks
- RFC 3412, Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)
- RFC 3413, Simple Network Management Protocol (SNMP) Applications
- RFC 3414, User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)
- RFC 3415, View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)
- RFC 3416, Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)
- RFC 3417, Transport Mappings for the Simple Network Management Protocol (SNMP)
- RFC 3418, Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)
- RFC 3419, Textual Conventions for Transport Addresses

1.2 Applications

Application	Description
Managing Network Devices Based on SNMP	Network devices are managed and monitored based on SNMP.

1.2.1 Managing Network Devices Based on SNMP

Scenario

Take the following figure as an example. Network device A is managed and monitored based on SNMP network manager.

Figure 1- 1

Remarks	A is a network device that needs to be managed. PC is a network management station.

Deployment

The network management station is connected to the managed network devices. On the network management station, users access the Management Information Base (MIB) on the network devices through the SNMP network manager and receive messages actively sent by the network devices to manage and monitor the network devices.

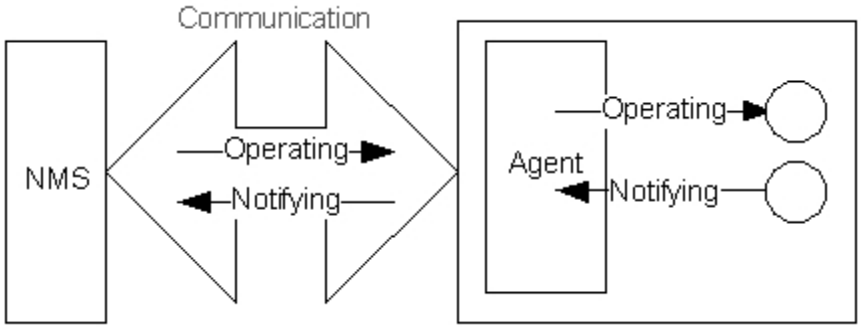
1.3 Features

Basic Concepts

SNMP is an application layer protocol that works in C/S mode. It consists of three parts:

- SNMP network manager
- SNMP agent
- MIB

Figure 1- 2 shows the relationship between the network management system (NMS) and the network management agent.



SNMP Network Manager

The SNMP network manager is a system that controls and monitors the network based on SNMP and is also called the NMS.

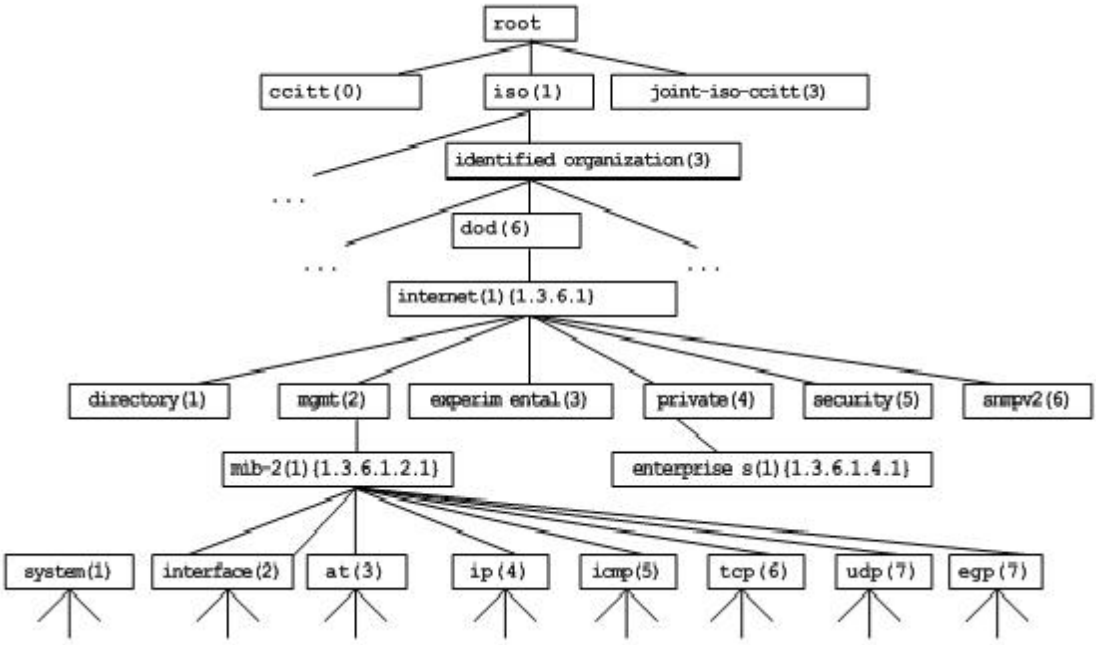
SNMP Agent

The SNMP agent (hereinafter referred to as the agent) is software running on the managed devices. It is responsible for receiving, processing, and responding to monitoring and control packets from the NMS. The agent may also actively send messages to the NMS.

MIB

The MIB is a virtual network management information base. The managed network devices contain lots of information. To uniquely identify a specific management unit among SNMP packets, the MIB adopts the tree hierarchical structure. Nodes in the tree indicate specific management units. A string of digits may be used to uniquely identify a management unit system among network devices. The MIB is a collection of unit identifiers of network devices.

Figure 1- 3 Tree Hierarchical Structure



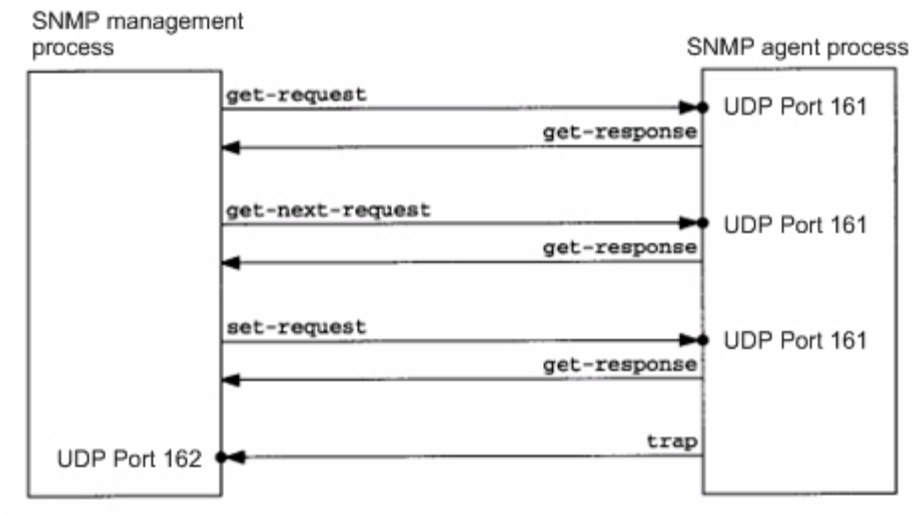
Operation Types

Six operation types are defined for information exchange between the NMS and the agent based on SNMP:

- Get-request: The NMS extracts one or more parameter values from the agent.
- Get-next-request: The NMS extracts the parameter value next to one or more parameters from the agent.
- Get-bulk: The NMS extracts a batch of parameter values from the agent.
- Set-request: The NMS sets one or more parameter values of the agent.
- Get-response: The agent returns one or more parameter values, which are the operations in response to the three operations performed by the agent on the NMS.
- Trap: The agent actively sends a message to notify the NMS of something that happens.

The first four packets are sent by the NMS to the agent and the last two packets are sent by the agent to the NMS. (Note: SNMPv1 does not support the Get-bulk operation.) Figure 1- 4 describes the operations.

Figure 1- 4 SNMP Packet Types



The three operations performed by the NMS on the agent and the response operations of the agent are based on UDP port 161. The trap operation performed by the agent is based on UDP port 162.

Overview

Feature	Description
Basic SNMP Functions	The SNMP agent is configured on network devices to implement basic functions such as information query for network nodes, network configuration, fault locating, and capacity planning.
SNMPv1 and SNMPv2C	SNMPv1 and SNMPv2C adopt the community-based security architecture, including authentication name and access permission.
SNMPv3	SNMPv3 redefines the SNMP architecture, namely, it enhances security functions, including the security model based on users and access control model based on views. The SNMPv3 architecture already includes all functions of SNMPv1 and SNMPv2C.

1.3.1 Basic SNMP Functions

Working Principle

Working Process

SNMP protocol interaction is response interaction (for exchange of packets, see Figure 1- 4). The NMS actively sends requests to the agent, including Get-request, Get-next-request, Get-bulk, and Set-request. The agent receives the requests, completes operations, and returns a Get-response. Sometimes, the agent actively sends a trap message and an Inform message to the NMS. The NMS does not need to respond to the trap message but needs to return an Inform-response to the agent. Otherwise, the agent re-sends the Inform message.

Related Configuration

Shielding or Disabling the SNMP Agent

By default, the SNMP function is enabled.

The **no snmp-server** command is used to disable the SNMP agent.

The **no enable service snmp-agent** command is used to directly disable all SNMP services.

Setting Basic SNMP Parameters

By default, the system contact mode, system location, and device Network Element (NE) information are empty. The default serial number is 60FF60, the default maximum packet length is 1,572 bytes, and the default UDP port ID of the SNMP service is 161.

The **snmp-server contact** command is used to configure or delete the system contact mode.

The **snmp-server location** command is used to configure or delete the system location.

The **snmp-server chassis-id** command is used to configure the system serial number or restore the default value.

The **snmp-server packetsize** command is used to configure the maximum packet length of the agent or restore the default value.

The **snmp-server net-id** command is used to configure or delete the device NE information.

The **snmp-server udp-port** command is used to set the UDP port ID of the SNMP service or restore the default value.

Configuring the SNMP Host Address

By default, no SNMP host is configured.

The **snmp-server host** command is used to configure the NMS host address to which the agent actively sends messages or to delete the specified SNMP host address. In the messages sent to the host, the SNMP version, receiving port, authentication name, or user can be bound. This command is used with the **snmp-server enable traps** command to actively send trap messages to the NMS.

Setting Trap Message Parameters

By default, SNMP is not allowed to actively send a trap message to the NMS, the function of sending a Link Trap message on an interface is enabled, the function of sending a system reboot trap message is disabled, and a trap message does not carry any private field.

By default, the IP address of the interface where SNMP packets are sent is used as the source address.

By default, the length of a trap message queue is 10 and the interval for sending a trap message is 30s.

The **snmp-server enable traps** command is used to enable or disable the agent to actively send a trap message to the NMS.

The **snmp trap link-status** command is used to enable or disable the function of sending a Link Trap message on an interface.

The **snmp-server trap-source** command is used to specify the source address for sending messages or to restore the default value.

The **snmp-server queue-length** command is used to set the length of a trap message queue or to restore the default value.

The **snmp-server trap-timeout** command is used to set the interval for sending a trap message or to restore the default value.

The **snmp-server trap-format private** command is used to set or disable the function of carrying private fields in a trap message when the message is sent.

The **snmp-server system-shutdown** command is used to enable or disable the function of sending a system reboot trap message.

Setting the SNMP Attack Protection and Detection Function

By default, the SNMP attack protection and detection function is disabled.

The **snmp-server authentication attempt** *times* **exceed** { **lock** | **lock-time** *minutes* | **unlock** } command is used to set and enable the attack protection and detection function.

1.3.2 SNMPv1 and SNMPv2C

SNMPv1 and SNMPv2C adopt the community-based security architecture. The administrator who can perform operations on the MIB of the agent is limited by defining the host address and authentication name (community string).

Working Principle

SNMPv1 and SNMPv2 determine whether the administrator has the right to use MIB objects by using the authentication name. The authentication name of the NMS must be the same as an authentication name defined in devices.

SNMPv2C adds the Get-bulk operation mechanism and can return more detailed error message types to the management workstation. The Get-bulk operation is performed to obtain all information from a table or obtain lots of data at a time, so as to reduce the number of request responses. The enhanced error handling capabilities of SNMPv2C include extension of error codes to differentiate error types. In SNMPv1, however, only one error code is provided for errors. Now, errors can be differentiated based on error codes. Because management workstations supporting SNMPv1 and SNMPv2C may exist on the network, the SNMP agent must be able to identify SNMPv1 and SNMPv2C packets and return packets of the corresponding versions.

Security

One authentication name has the following attributes:

- Read-only: Provides the read permission of all MIB variables for authorized management workstations.
- Read-write: Provide the read/write permission of all MIB variables for authorized management workstations.

Related Configuration

Setting Authentication Names and Access Permissions

The default access permission of all authentication names is read-only.

The **snmp-server community** command is used to configure or delete an authentication name and access permission.

This command is the first important command for enabling the SNMP agent function. It specifies community attributes and NMS scope where access to the MIB is allowed.

1.3.3 SNMPv3

SNMPv3 redefines the SNMP architecture and includes functions of SNMPv1 and SNMPv2 into the SNMPv3 system.

Working Principle

The NMS and SNMP agent are SNMP entities. In the SNMPv3 architecture, SNMP entities consist of the SNMP engine and SNMP applications. The SNMP engine is used to send and receive messages, identify and encrypt information, and control access to managed objects. SNMP applications refer to internal applications of SNMP, which work by using the services provided by the SNMP engine.

SNMPv3v determines whether a user has the right to use MIB objects by using the User-based Security Model (USM). The security level of the NMS user must be the same as that of an SNMP user defined in devices so as to manage devices.

SNMPv3 requires the NMS to obtain the SNMP agent engine IDs on devices when the NMS manages devices. SNMPv3 defines the discover and report operation mechanisms. When the NMS does not know agent engine IDs, the NMS may first send a discover message to the agent and the agent returns a report message carrying an engine ID. Later, management operations between the NMS and the agent must carry the engine ID.

Security

- SNMPv3 determines the data security mechanism based on the security model and security level. At present, security models include: SNMPv1, SNMPv2C, and SNMPv3. SNMPv3 includes SNMPv1 and SNMPv2C into the security model.

SNMPv1 and SNMPv2C Security Models and Security Levels

Security Model	Security Level	Authentication	Encryption	Description
SNMPv1	noAuthNoPriv	Authentication name	N/A	Data validity is confirmed through authentication name.
SNMPv2c	noAuthNoPriv	Authentication name	N/A	Data validity is confirmed through authentication name.

SNMPv3 Security Model and Security Level

Security Model	Security Level	Authentication	Encryption	Description
SNMPv3	noAuthNoPriv	User name.	N/A	Data validity is confirmed through user name.
SNMPv3	authNoPriv	MD5 or SHA	N/A	The data authentication mechanism based on HMAC-MD5 or HMAC-SHA is provided.
SNMPv3	authPriv	MD5 or SHA	DES	The data authentication mechanism based on HMAC-MD5 or HMAC-SHA and data encryption mechanism based on CBC-DES are provided.

Engine ID

An engine ID is used to uniquely identify an SNMP engine. Because each SNMP entity includes only one SNMP engine, one SNMP engine uniquely identifies an SNMP entity in a management domain. Therefore, the SNMPv3 agent as an entity must has a unique engine ID, that is, SnmpEngineID.

An engine ID is an octet string that consists of 5 to 32 bytes. RFC3411 defines the format of an engine ID:

- The first four bytes indicate the private enterprise ID (allocated by IANA) of a vendor, which is expressed in hexadecimal.
- The fifth byte indicates remaining bytes:
- 0: Reserved.
- 1: The later four bytes indicate an IPv4 address.
- 2: The later 16 bytes indicate an IPv6 address.
- 3: The later six bytes indicate a MAC address.
- 4: Text consisting of 27 bytes, which is defined by the vendor.
- 5: Hexadecimal value consisting of 27 bytes, which is defined by the vendor.
- 6-127: Reserved.
- 128-255: Formats specified by the vendor.

Related Configuration

Configuring an MIB View and a Group

By default, one view is configured and all MIB objects can be accessed.

By default, no user group is configured.

The **snmp-server view** command is used to configure or delete a view and the **snmp-server group** command is used to configure or delete a user group.

One or more instructions can be configured to specify different community names so that network devices can be managed by NMSs of different permissions.

Configuring an SNMP User

By default, no user is configured.

The **snmp-server user** command is used to configure or delete a user.

The NMS can communicate with the agent by using only legal users.

An SNMPv3 user can specify the security level (whether authentication and encryption are required), authentication algorithm (MD5 or SHA), authentication password, encryption password (only DES is available currently), and encryption password.

1.4 Configuration

Configuration	Description and Command	
Configuring Basic SNMP Functions	(Mandatory) It is used to enable users to access the agent through the NMS.	
	enable service snmp-agent	Enables the agent function.
	snmp-server community	Sets an authentication name and access permission.
	snmp-server user	Configures an SNMP user.
	snmp-server view	Configures an SNMP view.
	snmp-server group	Configures an SNMP user group.
	snmp-server authentication	Configures the SNMP attack protection and detection function.
Enabling the Trap Function	(Optional) It is used to enable the agent to actively send a trap message to the NMS.	
	snmp-server host	Configures the NMS host address.
	snmp-server enable traps	Enables the agent to actively send a trap message to the NMS.
	snmp trap link-status	Enables the function of sending a Link Trap message on an interface.
	snmp-server system-shutdown	Enables the function of sending a system reboot trap message.
	snmp-server trap-source	Specifies the source address for sending a trap message.
	snmp-server trap-format private	Enables a trap message to carry private fields when the message is sent.
Shielding the Agent Function	(Optional) It is used to shield the agent function when the agent service is not required.	
	no snmp-server	Shields the agent function.
Setting SNMP Control Parameters	(Optional) It is used to set or modify SNMP control parameters.	
	snmp-server contact	Sets the device contact mode.
	snmp-server location	Sets the device location.
	snmp-server logging	Sets the logging function.
	snmp-server chassis-id	Sets the serial number of the device.

Configuration	Description and Command	
	snmp-server net-id	Sets NE information about the device.
	snmp-server packetsize	Modifies the maximum packet length.
	snmp-server udp-port	Modifies the UDP port ID of the SNMP service.
	snmp-server queue-length	Modifies the length of a trap message queue.
	snmp-server trap-timeout	Modifies the interval for sending a trap message.

1.4.1 Configuring Basic SNMP Functions

Configuration Effect

Enable users to access the agent through the NMS.

Notes

- By default, no authentication name is set on network devices and SNMPv1 or SNMPv2C cannot be used to access the MIB of network devices. When an authentication name is set, if no access permission is specified, the default access permission is read-only.

Configuration Steps

Configuring an SNMP View

- Optional
- An SNMP view needs to be configured when the View-based Access Control Model (VACM) is used.

Configuring an SNMP User Group

- Optional
- An SNMP user group needs to be configured when the VACM is used.

Configuring an Authentication Name and Access Permission

- Mandatory
- An authentication name must be set on the agent when SNMPv1 and SNMPv2C are used to manage network devices.

Configuring an SNMP User

- Mandatory
- A user must be set when SNMPv3 is used to manage network devices.

Enabling the Agent Function

- Optional
- By default, the agent function is enabled. When the agent function needs to be enabled again after it is disabled, this command must be used.

Enabling the SNMP Attack Protection and Detection Function

- Optional

- By default, the SNMP attack protection and detection function is disabled. When malicious attacks need to be prevented, the configuration item must be used on the agent.

Verification

Run the **show snmp** command to check the SNMP function on devices.

Related Commands

Configuring an SNMP View

Command	**snmp-server view** *view-name oid-tree* { **include** | **exclude** }
Parameter Description	*view-name*: View name *oid-tree*: MIB objects associated with a view, which are displayed as an MIB subtree. **include:** Indicates that the MIB object subtree is included in the view. **exclude:** Indicates that the MIB object subtree is not included in the view.
Command Mode	Global configuration mode
Usage Guide	Specify a view name and use it for view-based management.

Configuring an SNMP User Group

Command	**snmp-server group** *groupname* { **v1** | **v2c** | **v3** { **auth** | **noauth** | **priv** } } [**read** *readview*] [**write** *writeview*] [**access** { **ipv6** *ipv6-aclname* | *aclnum* | *aclname* }]
Parameter Description	**v1** | **v2c** |**v3:** Specifies the SNMP version. **auth:** Messages sent by users in the group need to be verified but data confidentiality is not required. This configuration is valid for SNMPv3 only. **noauth:** Messages sent by users in the group do not need to be verified and data confidentiality is not required. This configuration is valid for SNMPv3 only. **priv:** Messages sent by users in the group need to be verified and confidentiality of transmitted data is required. This configuration is valid for SNMPv3 only. *readview*: Associates one read-only view. *writeview*: Associates one read/write view. *aclnum*: ACL number. The specified ACL is associated and the range of IPv4 NMS addresses from which access to the MIB is allowed is specified. *aclname*: ACL name. The specified ACL is associated and the range of IPv4 NMS addresses from which access to the MIB is allowed is specified. *ipv6-aclname*: IPv6 ACL name. The specified ACL is associated and the range of IPv6 NMS addresses from which access to the MIB is allowed is specified.
Command Mode	Global configuration mode
Usage Guide	Associate certain users with a group and associate the group with a view. Users in a group have the same access permission. In this way, you can determine whether managed objects associated with an operation are in the allowable range of a view. Only managed objects in the range of a view can be accessed.

Configuring an Authentication Name and Access Permission

Command	**snmp-server community** [*0* | *7*] *string* [**view** *view-name*] [[**ro** | **rw**] [**host** *ipaddr*]] [**ipv6** *ipv6-aclname*] [*aclnum* |

	aclname]
Parameter Description	*0*: Indicates that the input community string is a plaintext string. *7*: Indicates that the input community string is a ciphertext string. *string*: Community string, which is equivalent to the communication password between the NMS and the SNMP agent. *view-name*: Specifies a view name for view-based management. **ro:** Indicates that the NMS can only read variables of the MIB. **rw:** The NMS can read and write variables of the MIB. *aclnum*: ACL number. The specified ACL is associated and the range of IPv4 NMS addresses from which access to the MIB is allowed is specified. *aclname*: ACL name. The specified ACL is associated and the range of IPv4 NMS addresses from which access to the MIB is allowed is specified. *ipv6-aclname*: ACL name. The specified ACL is associated and the range of IPv6 NMS addresses from which access to the MIB is allowed is specified. *ipaddr*: Associates NMS addresses and specifies NMS addresses for accessing the MIB.
Command Mode	Global configuration mode
Usage Guide	This command is the first important command for enabling the SNMP agent function. It specifies community attributes and NMS scope where access to the MIB is allowed. To disable the SNMP agent function, run the **no snmp-server** command.

Configuring an SNMP User

Command	**snmp-server user** *username groupname* { **v1** | **v2c** | **v3** [**encrypted**] [**auth** { **md5** | **sha** } *auth-password*] [**priv des56** *priv-password*] } [**access** { **ipv6** *ipv6-aclname* | *aclnum* | *aclname* }]
Parameter Description	*username*: User name. *groupname*: Specifies the group name for a user. **v1** | **v2c** | **v3:** Specifies the SNMP version. Only SNMPv3 supports later security parameters. **encrypted:** The specified password input mode is ciphertext input. Otherwise, plaintext is used for input. If ciphertext input is selected, enter a key consisting of continuous hexadecimal digits. An MD5 protocol authentication key consists of 16 bytes and an SHA authentication protocol key consists of 20 bytes. Two characters stand for one byte. Encrypted keys are valid for this engine only. **auth:** Specifies whether authentication is used. **md5:** Specifies the MD5 authentication protocol. **sha** specifies the SHA authentication protocol. *auth-password*: Configures a password string (not more than 32 characters) used by the authentication protocol. The system converts the passwords into the corresponding authentication keys. **priv:** Specifies whether confidentiality is used. **des56** specifies the use of the 56-bit DES encryption protocol. *priv-password*: Configures a password string (not more than 32 characters) used for encryption. The system converts the password into the corresponding encryption key. *aclnum*: ACL number. The specified ACL is associated and the range of IPv4 NMS addresses from which access to the MIB is allowed is specified. *aclname*: ACL name. The specified ACL is associated and the range of IPv4 NMS addresses from which access to the MIB is allowed is specified. *ipv6-aclname*: IPv6 ACL name. The specified ACL is associated and the range of IPv6 NMS addresses from which access to the MIB is allowed is specified.
Command	Global configuration mode

Mode	
Usage Guide	Configure user information so that the NMS can communicate with the agent by using a valid user. For an SNMPv3 user, you can specify the security level, authentication algorithm (MD5 or SHA), authentication password, encryption algorithm (at present, only DES is available), and encryption password.

Enabling the Agent Function

Command	**enable service snmp-agent**
Parameter Description	N/A
Configuration mode	Privileged mode.
Usage Guide	This command is used to enable the SNMP agent function of a device.

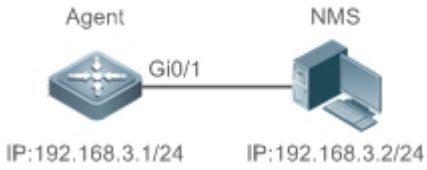
Enabling the SNMP Attack Protection and Detection Function

Command	**snmp-server authentication attempt** *times* **exceed** { **lock** | **lock-time** *minutes* | **unlock** }
Parameter Description	*times*: Number of continuous failed attempts. **lock**: After continuous authentication fails, the source IP address is permanently forbidden to initiate authentication for access. The administrator needs to manually unlock the IP address. **lock-time** *minutes*: After continuous authentication fails, the source IP address is forbidden to initiate authentication for access in a period of time. Beyond the period, the source IP address can be authenticated for access again. **unlock**: After continuous authentication fails, the source IP address is allowed to access the MIB continuously, which is equivalent to the fact that the SNMP attack protection and detection function is not configured.
Command Mode	Global configuration mode
Usage Guide	Configure the SNMP attack protection and detection function so that the corresponding measure can be taken after continuous authentication fails. The permanently forbidden source IP addresses can be authenticated for access again only after the administrator manually unlocks the IP addresses. The source IP address that are forbidden to access the MIB in a period of time can be authenticated for access again after the period expires or after the administrator manually unlocks the IP addresses.

Displaying the SNMP Status Information

Command	**show snmp** [**mib** | **user** | **view** | **group| host | process-mib-time**]
Parameter Description	**mib:** Displays information about the SNMP MIB supported in the system. **user:** Displays information about an SNMP user. **view:** Displays information about an SNMP view. **group:** Displays information about an SNMP user group. **host:** Displays information about user configuration. **process-mib-time:** Displays the MIB node with the longest processing time.
Configuration mode	Privileged mode.
Usage Guide	N/A

Configuring SNMPv3 Configuration

Scenario **Figure 1-5**	 ● The NMS manages network devices (agents) based on the user authentication and encryption mode, for example, the NMS uses user1 as the user name, MD5 as the authentication mode, 123 as the authentication password, DES56 as the encryption algorithm, and 321 as the encryption password. ● Network devices can control the operation permission of users to access MIB objects. For example, the user named user1 can read MIB objects under the system node (1.3.6.1.2.1.1) and can only write MIB objects under the SysContact node (1.3.6.1.2.1.1.4.0). ● Network devices can actively send authentication and encryption messages to the NMS.
Configuration Steps	● Configure a MIB view and a MIB group. Create a MIB view “view1”, which includes the associated MIB object (1.3.6.1.2.1.1); then create a MIB view “view2”, which includes the associated MIB object (1.3.6.1.2.1.1.4.0). Create a group “g1”, select the version “v3”, set the security level to the authentication and encryption mode “priv”, and configure permissions to read the view “view1” and write the view “view2”. ● Configure an SNMP user. Create a user named “user1” under group “g1”, select “v3” as the version, and set the authentication mode to “md5”, authentication password to “123”, encryption mode to “DES56”, and encryption password to “321”. ● Configure the SNMP host address. Set the host address to 192.168.3.2, select “3” as the version, set the security level to the authentication and encryption mode “priv”, and associate the user name “user1”. Enable the agent to actively send a trap message to the NMS. ● Set the IP address of the agent. Set the address of the Gi0/1 interface to 192.168.3.1/24.
Agent	FS(config)#snmp-server view view1 1.3.6.1.2.1.1 include FS(config)#snmp-server view view2 1.3.6.1.2.1.1.4.0 include FS(config)#snmp-server group g1 v3 priv read view1 write view2 FS(config)#snmp-server user user1 g1 v3 auth md5 123 priv des56 321 FS(config)#snmp-server host 192.168.3.2 traps version 3 priv user1 FS(config)#snmp-server enable traps FS(config)#interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)#ip address 192.168.3.1 255.255.255.0 FS(config-if-gigabitEthernet 0/1)#exit
Verification	1. Run the **show running-config** command to display configuration information of the device. 2. Run the **show snmp user** command to display the SNMP user. 3. Run the **show snmp view** command to display the SNMP view. 4. Run the **show snmp group** command to display the SNMP group. 5. Run the **show snmp host** command to display the host information configured by the user. 6. Install MIB-Browser.

Agent	FS# show running-config
	!
	interface gigabitEthernet 0/1
	  no ip proxy-arp
	  ip address 192.168.3.1 255.255.255.0
	!
	snmp-server view view1 1.3.6.1.2.1.1 include
	snmp-server view view2 1.3.6.1.2.1.1.4.0 include
	snmp-server user user1 g1 v3 encrypted auth md5 7EBD6A1287D3548E4E52CF8349CBC93D priv des56 D5CEC4884360373ABBF30AB170E42D03
	snmp-server group g1 v3 priv read view1 write view2
	snmp-server host 192.168.3.2 traps version 3 priv user1
	snmp-server enable traps

	FS# show snmp user User name: user1 Engine ID: 800013110300d0f8221120 storage-type: permanent active Security level: auth priv Auth protocol: MD5 Priv protocol: DES Group-name: g1
	FS#show snmp view view1(include) 1.3.6.1.2.1.1 view2(include) 1.3.6.1.2.1.1.4.0 default(include) 1.3.6.1
	FS# show snmp group groupname: g1 securityModel: v3 securityLevel:authPriv readview: view1 writeview: view2 notifyview:
	FS#show snmp host

Notification host: 192.168.3.2

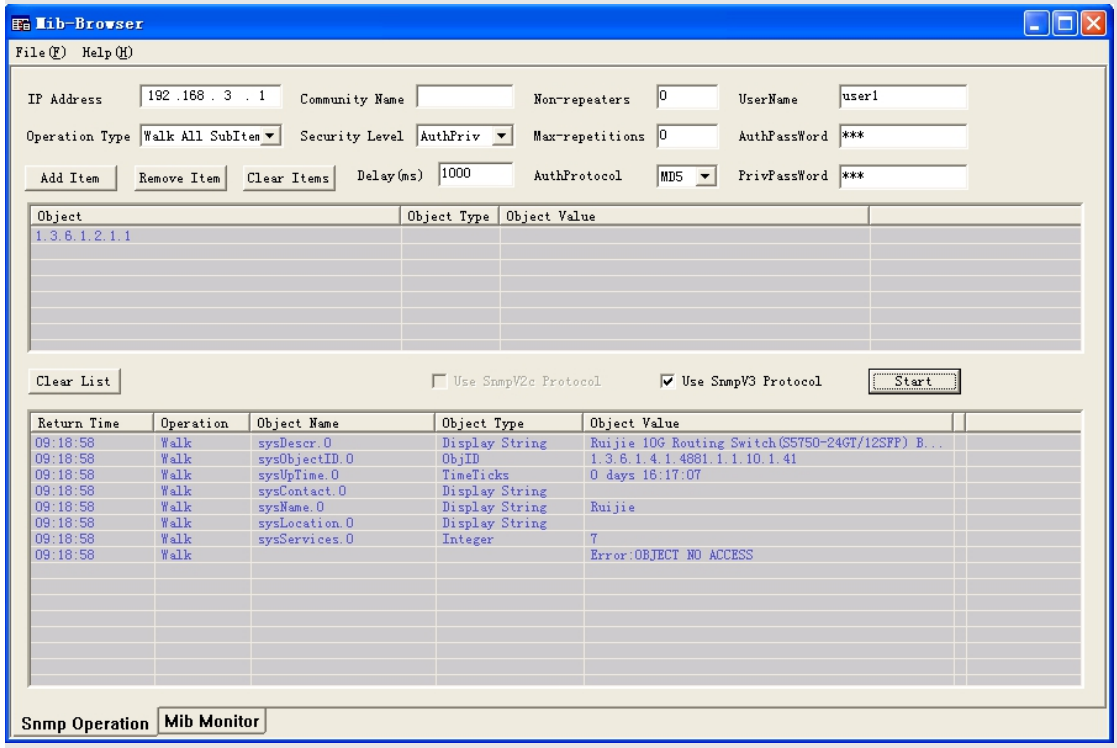
udp-port: 162

type: trap

user: user1

security model: v3 authPriv

Install MIB-Browser, enter IP address 1**92.168.3.1** in **IP Address** and **user1** in **UserName**, select **AuthPriv** for **Security Level**, enter **123** in **AuthPassWord**, select **MD5** for **AuthProtocol**, and enter **321** in **PrivPassWord**. Click **Add Item** and select a management unit for which the MIB needs to be queried, for example, **System** in the following figure. Click **Start**. The MIB is queried for network devices. The lowest pane in the following figure shows query results.



Common Errors

-

1.4.2 Enabling the Trap Function

Configuration Effect

Enable the agent to actively send a trap message to the NMS.

Notes

N/A

Configuration Steps

Configuring the SNMP Host Address

- Optional

- Configure the host address of the NMS when the agent is required to actively send messages.

Enabling the Agent to Actively Send a Trap Message to the NMS

- Optional
- Configure this item on the agent when the agent is required to actively send a trap message to the NMS.

Enabling the Function of Sending a Link Trap Message on an Interface

- Optional
- Configure this item on the agent when a link trap message needs to be sent on an interface.

Enabling the Function of Sending a System Reboot Trap Message

- Optional
- Configure this item on the agent when the FSOS system is required to send a trap message to the NMS to notify system reboot before reloading or reboot of the device.

Specifying the Source Address for Sending a Trap Message

- Optional
- Configure this item on the agent when it is required to permanently use a local IP address as the source SNMP address to facilitate management.

Enabling a Trap Message to Carry Private Fields when the Message Is Sent

- Optional
- Configure this item on the agent when private fields need to be carried in a trap message.

Verification

Run the **show snmp** command to display the SNMP status.

Run the **show running-config** command to display configuration information of the device.

Related Commands

Setting the NMS Host Address

Command	**snmp-server host** { *host-addr* | **ipv6** *ipv6-addr* } [**traps** | **inrorms**] [**version** { **1** | **2c** | **3** { **auth** | **noauth** | **priv** }] *community-string* [**udp-port** *port-num*] [*notification-type*]
Parameter Description	*host-addr*: Address of the SNMP host. *ipv6-addr*: (IPv6) address of the SNMP host. **traps | informs:** Configures the host to send a trap message or an inform message. **version:** SNMP version, which can be set to **V1**, **V2C**, or **V3**. **auth** | **noauth** | **priv:** Sets the security level of V3 users. *community-string*: Community string or user name (V3). *port-num*: Configures the port ID of the SNMP host. *notification-type*: Type of trap messages that are actively sent, for example, snmp. If no trap type is specified, all trap messages are sent.

Command Mode	Global configuration mode
Usage Guide	This command is used with the **snmp-server enable traps** command to actively send trap messages to the NMS.

Enabling the Agent to Actively Send a Trap Message to the NMS

Command	**snmp-server enable traps** [*notification-type*]
Parameter Description	*notification-type*: Enables trap notification for the corresponding events, including the following types: snmp: Enables trap notification for SNMP events. bridge: Enables trap notification for bridge events. mac-notification: Enables trap notification for MAC events. ospf: Enables trap notification for OSPF events. vrrp: Enables trap notification for VRRP events. web-auth: Enables trap notification for Web authentication events.
Command Mode	Global configuration mode
Usage Guide	This command must be used with the **snmp-server host** command to so that trap messages can be actively sent.

Enabling the Function of Sending a Link Trap Message on an Interface

Command	**snmp trap link-status**
Parameter Description	-
Configuration mode	Interface configuration mode
Usage Guide	For interfaces (Ethernet interface, AP interface, and SVI interface), when this function is enabled, the SNMP sends a Link Trap message if the link status on the interfaces changes. Otherwise, the SNMP does not send the message.

Enabling the Function of Sending a System Reboot Trap Message

Command	**snmp-server system-shutdown**
Parameter Description	-
Configuration mode	Global configuration mode
Usage Guide	When the function of notification upon SNMP system reboot is enabled, a trap message is sent to the NMS to notify system reboot before reloading or reboot of the device.

Specifying the Source Address for Sending a Trap Message

Command	**snmp-server trap-source** *interface*
Parameter Description	*interface*: Used as the interface for the SNMP source address.
Configuration mode	Global configuration mode
Usage Guide	By default, the IP address of the interface where SNMP packets are sent is used as the source address. To facilitate management and identification, this command can be run to permanently use one local IP address as the source SNMP

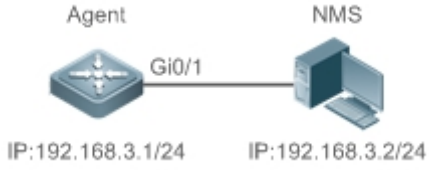
	address.

Enabling a Trap message to Carry Private Fields when the Message Is Sent

Command	**snmp-server trap-format private**
Parameter Description	N/A
Configuration mode	Global configuration mode
Usage Guide	This command can be used to enable a trap message to carry private fields when the message is sent. At present, supported private fields include the alarm generation time. For the specific data types and data ranges of the fields, see FS-TRAP-FORMAT-MIB.mib.

Configuration Example

Enabling the Trap Function

Scenario **Figure 1-6**	 ● The NMS manages network devices (agents) based on the community authentication mode, and network devices can actively send messages to the NMS.
Configuration Steps	1. Perform configuration to enable the agent to actively send messages to the NMS. Set the SNMP host address to 192.168.3.2, the message format to Version2c, and the authentication name to user1. Enable the agent to actively send trap messages. 2. Set the IP address of the agent. Set the address of the Gi0/1 interface to 192.168.3.1/24.
Agent	FS(config)#snmp-server host 192.168.3.2 traps version 2c user1 FS(config)#snmp-server enable traps FS(config)#interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)#ip address 192.168.3.1 255.255.255.0 FS(config-if-gigabitEthernet 0/1)#exit
Verification	● Run the **show running-config** command to display configuration information of the device. ● Run the **show snmp** command to display the SNMP status.
Agent	FS# show running-config ip access-list standard a1 10 permit host 192.168.3.2 interface gigabitEthernet 0/1 no ip proxy-arp

	ip address 192.168.3.1 255.255.255.0 snmp-server view v1 1.3.6.1.2.1.1 include snmp-server location fuzhou snmp-server host 192.168.3.2 traps version 2c user1 snmp-server enable traps snmp-server contact FS.com.cn snmp-server community user1 view v1 rw a1 snmp-server chassis-id 1234567890
	FS#show snmp Chassis: 1234567890 0 SNMP packets input 0 Bad SNMP version errors 0 Unknown community name 0 Illegal operation for community name supplied 0 Encoding errors 0 Number of requested variables 0 Number of altered variables 0 Get-request PDUs 0 Get-next PDUs 0 Set-request PDUs 0 SNMP packets output 0 Too big errors (Maximum packet size 1472) 0 No such name errors 0 Bad values errors 0 General errors 0 Response PDUs 0 Trap PDUs SNMP global trap: enabled SNMP logging: disabled SNMP agent: enabled

Common Errors

N/A

1.4.3 Shielding the Agent Function

Configuration Effect

Shield the agent function when the agent service is not required.

Notes

- Run the **no snmp-server** command to shield the SNMP agent function when the agent service is not required.
- Different from the shielding command, after the **no enable service snmp-agent** command is run, all SNMP services are directly disabled (that is, the SNMP agent function is disabled, no packet is received, and no response packet or trap packet is sent), but configuration information of the agent is not shielded.

Configuration Steps

Shielding the SNMP Agent Function for the Device

- Optional
- To shield the configuration of all SNMP agent services, use this configuration.

Disabling the SNMP Agent Function for the Device

- Optional
- To directly disable all services, use this configuration.

Verification

Run the **show services** command to check whether SNMP services are enabled or disabled.

Run the **show snmp** command to display the SNMP status.

Run the **show running-config** command to display configuration information of the device.

Related Commands

Shielding the SNMP Agent Function for the Device

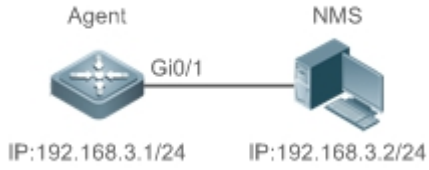
Command	**no snmp-server**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	By default, the SNMP agent function is disabled. When SNMP agent parameters (for example, NMS host address, authentication name, and access permission) are set, the SNMP agent service is automatically enabled. The **enable service snmp-agent** command must also be run at the same time so that the SNMP agent service can take effect. If the SNMP agent service is disabled or the **enable service snmp-agent** command is not run, the SNMP agent service does not take effect. Run the **no snmp-server** command to disable SNMP agent services of all versions supported by the device. After this command is run, all SNMP agent service configurations are shielded (that is, after the **show running-config** command is run, no configuration is displayed. Configurations are restored after the SNMP agent service is enabled again). After the **enable service snmp-agent** command is run, the SNMP agent configurations are not shielded.

Disabling the SNMP Agent Function for the Device

Command	**no enable service snmp-agent**
Parameter Description	N/A
Configuration mode	Global configuration mode
Usage Guide	This command can be used to disable the SNMP service, but it will not shield SNMP agent parameters.

Configuration Example

Enabling the SNMP Service

Scenario Figure 1-7	 After the SNMP service is enabled and the SNMP agent server is set, the NMS can access devices based on SNMP.
Configuration Steps	1. Enable the SNMP service. 2. Set parameters for the SNMP agent server to make the SNMP service take effect.
A gent	FS(config)#enable service snmp-agent
Verification	1. Run the **show services** command to check whether the SNMP service is enabled or disabled.
Agent	FS#show service web-server : disabled web-server(https): disabled snmp-agent : enabled ssh-server : disabled telnet-server : enabled

Common Errors

N/A

1.4.4 Setting SNMP Control Parameters

Configuration Effect

Set basic parameters of the SNMP agent, including the device contact mode, device location, serial number, and parameters for sending a trap message. By accessing the parameters, the NMS can obtain the contact person of the device and physical location of the device.

Notes

N/A

Configuration Steps

Setting the System Contact Mode

- Optional
- When the contact mode of the system needs to be modified, configure this item on the agent.

Setting the System Location

- Optional
- When the system location needs to be modified, configure this item on the agent.

Setting the System Serial Number

- Optional
- When the system serial number needs to be modified, configure this item on the agent.

Setting NE Information about the Device

- Optional
- When the NE code needs to be modified, configure this item on the agent.

Setting the Maximum Packet Length of the SNMP Agent

- Optional
- When the maximum packet length of the SNMP agent needs to be modified, configure this item on the agent.

Setting the UDP Port ID of the SNMP Service

- Optional
- When the UDP port ID of the SNMP service needs to be modified, configure this item on the agent.

Setting the Queue Length of Trap Messages

- Optional
- When the size of the message queue needs to be adjusted to control the message sending speed, configure this item on the agent.

Setting the Interval for Sending a Trap Message

- Optional
- When the interval for sending a trap message needs to be modified, configure this item on the agent.

Configuring SNMP Flow Control

- Optional
- If a large number of SNMP request packets result in high CPU usage for SNMP tasks, configure SNMP flow control to limit the number of request packets processed per second in each SNMP task, so as to control the CPU usage for SNMP tasks.

Verification

Run the **show snmp** command to display the SNMP status.

Run the **show running-config** command to display configuration information of the device.

Related Commands

Setting the System Contact Mode

Command	**snmp-server contact** *text*
Parameter Description	*text*: String that describes the system contact mode.
Command Mode	Global configuration mode
Usage Guide	N/A

Setting the System Location

Command	**snmp-server location** *text*
Parameter Description	*text*: String that describes system information.
Configuration mode	Global configuration mode
Usage Guide	N/A

Setting the System Serial Number

Command	**snmp-server chassis-id** *text*
Parameter Description	*text*: Text of the system serial number, which may be digits or characters.
Configuration mode	Global configuration mode
Usage Guide	In general, the device serial number is used as the SNMP serial number to facilitate identification of the device.

Setting NE Information about the Device

Command	**snmp-server net-id** *text*
Parameter Description	*text*: Text that is used to set the device NE code. The text is a string that consists of 1 to 255 characters that are case-sensitive and may include spaces.
Configuration mode	Global mode.
Usage Guide	Set the NE code of the device.

Setting the Maximum Packet Length of the SNMP Agent

Command	**snmp-server packetsize** *byte-count*
Parameter Description	*byte-count*: Packet size, ranging from 484 bytes to 17,876 bytes.
Configuration mode	Global mode.
Usage Guide	N/A

Setting the UDP Port ID of the SNMP Service

Command	**snmp-server udp-port** *port-num*
Parameter Description	*port-num*: Specifies the UDP port ID of the SNMP service, that is, the ID of the protocol port that receives SNMP packets.
Configuration mode	Global mode.
Usage Guide	Specify the protocol port ID for receiving SNMP packets.

Setting the Length of a Trap Message Queue

Command	**snmp-server queue-length** *length*
Parameter Description	*length*: Queue length, ranging from 1 to 1,000.
Configuration mode	Global configuration mode
Usage Guide	Adjust the size of the message queue to control the message sending speed.

Setting the Interval for Sending a Trap Message

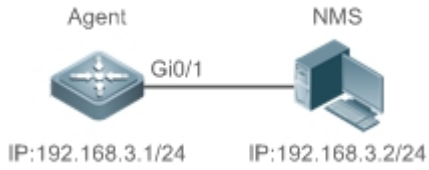
Command	**snmp-server trap-timeout** *seconds*
Parameter Description	*seconds*: Interval (unit: second). The value range is 1 to 1,000.
Configuration mode	Global configuration mode
Usage Guide	Adjust the interval for sending a message to control the message sending speed.

Configuring SNMP Flow Control

Command	**snmp-server flow-control pps** [*count*]
Parameter Description	*count***:** Number of SNMP request packets processed per second. The value range is 50 to 65,535.
Command Mode	Global configuration mode
Usage Guide	If a large number of SNMP request packets result in high CPU usage for SNMP tasks, configure SNMP flow control to limit the number of request packets processed per second in each SNMP task, so as to control the CPU usage for SNMP tasks.

Configuration Example

Setting SNMP Control Parameters

Scenario **Figure 1-8**	 ● The NMS manages network devices (agents) based on the community authentication mode and can obtain basic

	system information about the devices, for example, system contact mode, location, and serial number.
Configuration Steps	1. Set SNMP agent parameters. Set the system location, contact mode, and serial number. 2. Set the IP address of the agent. Set the address of the Gi0/1 interface to 192.168.3.1/24.
Agent	FS(config)#snmp-server location fuzhou FS(config)#snmp-server contact FS.com.cn FS(config)#snmp-server chassis-id 1234567890 FS(config)#interface gigabitEthernet 0/1 FS(config-if-gigabitEthernet 0/1)#ip address 192.168.3.1 255.255.255.0 FS(config-if-gigabitEthernet 0/1)#exit
Verification	1. Check the configuration information of the device. 2. Check the SNMP view and group information.

Agent	FS# show running-config
	ip access-list standard a1
	  10 permit host 192.168.3.2
	interface gigabitEthernet 0/1
	  no ip proxy-arp
	  ip address 192.168.3.1 255.255.255.0
	snmp-server view v1 1.3.6.1.2.1.1 include
	snmp-server location fuzhou
	snmp-server host 192.168.3.2 traps version 2c user1
	snmp-server enable traps
	snmp-server contact FS.com.cn
	snmp-server community user1 view v1 rw a1
	snmp-server chassis-id 1234567890

	FS#show snmp view v1(include) 1.3.6.1.2.1.1 default(include) 1.3.6.1 FS#show snmp group groupname: user1 securityModel: v1 securityLevel:noAuthNoPriv readview: v1

	writeview: v1 notifyview: groupname: user1 securityModel: v2c securityLevel:noAuthNoPriv readview: v1 writeview: v1 notifyview:

Common Errors

N/A

1.5 Monitoring

Clearing

Description	Command
Clears the list of source IP addresses that are locked after continuous authentication fails.	**clear snmp locked-ip** [**ipv4** *ipv4-address* | **ipv6** *ipv6-address*]

Displaying

Description	Command
Displays the SNMP status.	**show snmp** [**mib** | **user** | **view** | **group**| **host**]

2 Configuring RMON

2.1 Overview

The Remote Network Monitoring (RMON) aims at resolving problems of managing local area networks (LANs) and remote sites by using one central point. In RMON, network monitoring data consists of a group of statistics and performance indicators, which can be used for monitoring the network utilization, so as to facilitate network planning, performance optimization, and network error diagnosis.

RMON is mainly used by a managing device to remotely monitor and manage managed devices.

Protocols and Standards

STD 0059 / RFC 2819: Remote Network Monitoring Management Information Base

RFC4502: Remote Network Monitoring Management Information Base Version 2

RFC 3919: Remote Network Monitoring (RMON) Protocol Identifiers for IPv6 and Multi Protocol Label Switching (MPLS)

RFC 3737: IANA Guidelines for the Registry of Remote Monitoring (RMON) MIB Modules

RFC 3434: Remote Monitoring MIB Extensions for High Capacity Alarms

RFC 3395: Remote Network Monitoring MIB Protocol Identifier Reference Extensions

RFC 3287: Remote Monitoring MIB Extensions for Differentiated Services

RFC 3273: Remote Network Monitoring Management Information Base for High Capacity Networks

RFC 2896: Remote Network Monitoring MIB Protocol Identifier Macros

RFC 2895: Remote Network Monitoring MIB Protocol Identifier Reference

2.2 Applications

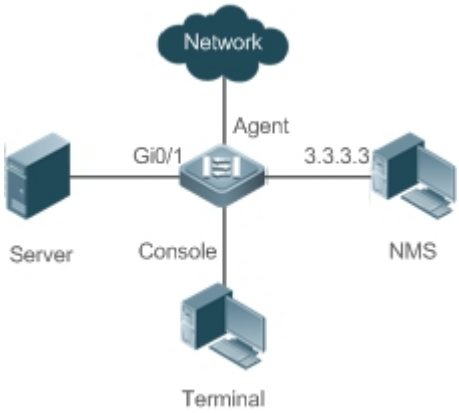
Application	Description
Collecting Statistics on Information of a Monitored Interface	Applies four functions of RMON to an interface to monitor the network communication of the interface.

2.2.1 Collecting Statistics on Information of a Monitored Interface

Scenario

The RMON Ethernet statistics function is used to monitor accumulated information of an interface, the history statistics function is used to monitor the packet count of an interface within each monitoring interval, and the alarm function is used to immediately acquire packet count exceptions of an interface. The following figure shows the networking topology.

Figure 2- 1



Deployment

Interface is monitored to accumulatively collect statistics on the packet count of the interface and collect statistics on the packet count and bandwidth utilization of the interface within the monitoring interval. If a packet count exception occurs on the interface, an alarm is reported to the network management system (NMS). The configuration key points are as follows:

- Configure the RMON Ethernet statistics function on interface.
- Configure the RMON history statistics function on interface.
- Configure the RMON alarm table and define RMON event processing actions in configuration mode. Monitored objects of alarms are the object identifier (OID) values of specific fields in the RMON Ethernet statistical table configured for interface.

2.3 Features

Basic Concepts

RMON defines multiple RMON groups. FS products support the statistics group, history group, alarm group, and event group, which are described as follows:

Statistics Group

The statistics group is used to monitor and collect statistics on Ethernet interface traffic information, which is accumulated from the entry creation time to the current time. The statistical items include discarded data packets, broadcast data packets, cyclic redundancy check (CRC) errors, large and small blocks, and collisions. Statistical results are stored in the Ethernet statistical table.

History Group

The history group is used to periodically collect network traffic information. It records accumulated values of network traffic information and the bandwidth utilization within each interval, and saves them in the history control table. It includes two small groups:

- The HistoryControl group is used to set the sampling interval, sampling data source, and other control information.
- The EthernetHistory group provides administrators with historical data, including statistics on network segment traffic, error packets, broadcast packets, utilization, and number of collisions.

Alarm Group

The alarm group is used to monitor a specified Management Information Base (MIB) object. When the value of a MIB object exceeds the preset upper limit or is lower than the preset lower limit, an alarm is triggered and the alarm is processed as an event.

Event Group

The event group is used to define the event processing mode. When a monitored MIB object meets alarm conditions, an event is triggered. An event can be processed in any of the following modes:

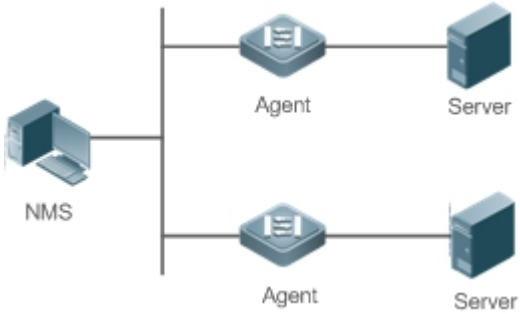
- none: No action is taken.
- log: Event-relevant information is recorded in the log record table so that administrators can view it at any time.
- snmp-trap: A trap message is transmitted to the NMS to notify the NMS of the event occurrence.
- log-and-trap: Event-relevant information is recorded in the log record table and a trap message is transmitted to the NMS.

Working Principle

RMON supports multiple monitors and two data collection methods. Method 1: A dedicated RMON probe is used to collect data and the NMS can directly acquire all information about the RMON MIB from the RMON probe. Method 2: RMON agents are built into network devices (such as switches and routers) so that the devices have the RMON probe function. The NMS uses basic commands of the Simple Network Management Protocol (SNMP) to exchange data with the RMON agents and collect network management information. This method, however, is limited by device resources and information of only four groups rather than all data of the RMON MIB is acquired.

The following figure shows an example of communication between the NMS and RMON agents. The NMS, through the RMON agents running on devices, can acquire information about overall traffic, error statistics, and performance statistics of the network segment where a managed network device interface is, thereby implementing remote management of network devices.

Figure 2- 2



Overview

Feature	Description
RMON Ethernet Statistics	Collects statistics on the packet count, byte count, and other data of a monitored Ethernet interface accumulatively.
RMON History Statistics	Records the counts of packets, bytes, and other data communicated by an Ethernet interface within the configured interval and calculates the bandwidth utilization within the interval.
RMON Alarm	Samples values of monitored variables at intervals. The alarm table is used in combination with the event table. When the upper or lower limit is reached, a relevant event table is triggered to perform event processing or no processing is performed.

2.3.1 RMON Ethernet Statistics

Working Principle

The RMON Ethernet statistics function accumulatively collects statistics on network traffic information of an Ethernet interface from the entry creation time to the current time.

Related Configuration

Configuring RMON Statistical Entries

- The RMON Ethernet statistics function is disabled by default.
- Run the **rmon collection stats** command to create Ethernet statistical entries on a specified Ethernet interface.
- After statistical entries are successfully created on a specified interface, the statistics group collects statistics on the traffic information of the current interface. The statistical items are variables defined in the RMON Ethernet statistical table, and recorded information is the accumulated values of variables from the creation time of the RMON statistical table to the current time.

2.3.2 RMON History Statistics

Working Principle

The RMON history statistics function records accumulated statistics on traffic information of an Ethernet interface within each interval.

Related Configuration

Configuring RMON Historical Control Entries

- The RMON history statistics function is disabled by default.
- Run the **rmon collection history** command to create historical control entries on an Ethernet interface.
- The RMON history group collects statistics on variables defined in the RMON history table and records accumulated values of variables within each interval.

2.3.3 RMON Alarm

Working Principle

The RMON alarm function periodically monitors value changes of alarm variables. If the value of an alarm variable reaches the specified upper threshold or lower threshold, a corresponding event is triggered for processing, for example, a trap message is transmitted or one logTable entry record is generated. If a lower threshold or upper threshold is reached multiple times consecutively, only one corresponding event is triggered and another event is triggered till a reverse threshold is reached.

Related Configuration

Configuring the Event Table

- The RMON event group function is disabled by default.
- Run the **rmon event** command to configure the event table.

Configuring Alarm Entries

- The RMON alarm group function is disabled by default.
- Run the **rmon event** command to configure the event table and run the **rmon alarm** command to configure the RMON alarm table.
- The RMON alarm function is implemented by the alarm table and event table jointly. If a trap message needs to be transmitted to a managing device in the case of an alarm event, the SNMP agent must be correctly configured first. For the configuration of the SNMP agent, see the *Configuring SNMP*.

- If a configured alarm object is a field node in the RMON statistics group or history group, the RMON Ethernet statistics function or RMON history statistics function need to be configured on a monitored Ethernet interface first.

2.4 Configuration

Configuration	Description and Command	
Configuring RMON Ethernet Statistics	(Mandatory) It is used to accumulatively collect statistics on traffic information of an Ethernet interface.	
	rmon collection stats	Configures Ethernet statistical entries.
Configuring RMON History Statistics	(Mandatory) It is used to collect, at intervals, statistics on traffic information of an Ethernet interface and the bandwidth utilization within the interval.	
	rmon collection history	Configures historical control entries.
Configuring RMON Alarm	(Mandatory) It is used to monitor whether data changes of a variable is within the valid range.	
	rmon event	Configures event entries.
	rmon alarm	Configures alarm entries.

2.4.1 Configuring RMON Ethernet Statistics

Configuration Effect

Acquire accumulated statistics on traffic information of a monitored Ethernet interface from the entry creation time to the current time.

Notes

This function cannot be configured in batch interface configuration mode.

Configuration Steps

Configuring RMON Statistical Entries

- Mandatory.
- If statistics and monitoring are required for a specified interface, Ethernet statistical entries must be configured on this interface.

Verification

Run the **show rmon stats** command to display Ethernet statistics.

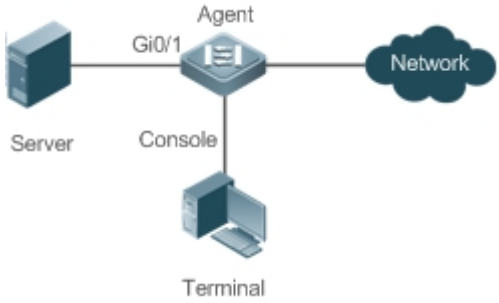
Related Commands

Configuring RMON Statistical Entries

Command	**rmon collection stats** *index* [**owner** *ownername*]
Parameter Description	*index*: Indicates the index number of a statistical entry, with the value ranging from 1 to 65,535. **owner** *ownername*: Indicates the entry creator, that is, *ownername*, which is a case-sensitive string of 1-63 characters.
Command Mode	Interface configuration mode
Usage Guide	The values of statistical entry parameters cannot be changed.

Configuration Example

Configuring RMON Ethernet Statistics

Scenario **Figure 2- 3**	
	As shown in the preceding figure, the RMON agent is connected to the server, and the NMS requires the RMON statistics group to conduct performance statistics on received packets of interface Gi0/1. Administrators can view the statistics at any time to understand data about received packets of an interface and take measures in a timely manner to handle network exceptions.
Configuration Steps	● Configure a statistical table instance on interface GigabitEthernet 0/1 to collect statistics on the traffic of this interface.
Agent	FS# configure terminal FS (config)# interface gigabitEthernet 0/1 FS (config-if-GigabitEthernet 0/1)# rmon collection stats 1 owner admin
Verification	Run the **show rmon stats** command to display Ethernet statistics.
Agent	FS# show rmon stats ether statistic table: index = 1 interface = GigabitEthernet 0/1 owner = admin status = 1 dropEvents = 0 octets = 25696 pkts = 293 broadcastPkts = 3 multiPkts = 0 crcAlignErrors = 0 underSizePkts = 0 overSizePkts = 0 fragments = 0

	jabbers = 0 collisions = 0 packets64Octets = 3815 packets65To127Octets = 1695 packets128To255Octets = 365 packets256To511Octets = 2542 packets512To1023Octets = 152 packets1024To1518Octets = 685

Common Errors

Statistical table entries are re-configured or configured statistical table entries are modified.

2.4.2 Configuring RMON History Statistics

Configuration Effect

Acquire accumulated statistics on the traffic of a monitored Ethernet interface and the bandwidth utilization within each interval.

Notes

This function cannot be configured in batch interface configuration mode.

Configuration Steps

- Mandatory.
- If network statistics on a specified interface need to be collected, RMON historical control entries must be configured on the interface.

Verification

Run the **show rmon history** command to display history group statistics.

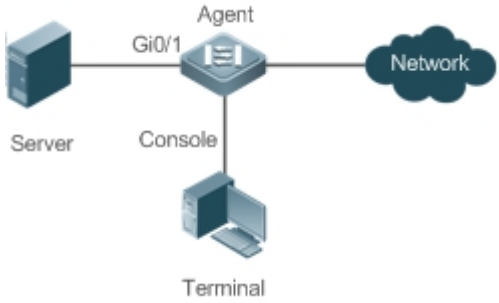
Related Commands

Configuring RMON Historical Control Entries

Command	**rmon collection history** *index* [**owner** *ownername*] [**buckets** *bucket-number*] [**interval** *seconds*]
Parameter Description	*index*: Indicates the index number of a history statistical entry, with the value ranging from 1 to 65,535. **owner** *ownername*: Indicates the entry creator, that is, *ownername*, which is a case-sensitive string of 1-63 characters. **buckets** *bucket-number*: Sets the capacity of the history table in which a history statistical entry exists, that is, sets the maximum number of records (*bucket-number*) that can be accommodated in the history table. The value of *bucket-number* ranges from 1 to 65,535 and the default value is **10**. **interval** *seconds*: Sets the statistical interval, with the unit of seconds. The value ranges from 1 second to 3,600 seconds and the default value is **1,800** seconds.
Command Mode	Interface configuration mode
Usage Guide	The values of history statistical entry parameters cannot be changed.

Configuration Example

Configuring RMON History Statistics

Scenario **Figure 2- 4**	
	As shown in the preceding figure, the RMON agent is connected to the server, and the NMS needs to collect statistics on received packets of interface Gi0/1 through the RMON history group at an interval of 60 seconds, in an effort to monitor the network and understand emergency data.
Configuration Steps	● Configure the history control table on interface GigabitEthernet 0/1 to periodically collect statistics on the traffic of this interface.
Agent	FS# configure terminal FS(config)# interface gigabitEthernet 0/1 FS(config-if-GigabitEthernet 0/1)# rmon collection history 1 buckets 5 interval 300 owner admin
Verification	Run the **show rmon history** command to display history group statistics.
Agent	FS# show rmon history rmon history control table: index = 1 interface = GigabitEthernet 0/1 bucketsRequested = 5 bucketsGranted = 5 interval = 60 owner = admin stats = 1 rmon history table: index = 1 sampleIndex = 786 intervalStart = 6d:18h:37m:38s dropEvents = 0 octets = 2040

```
pkts = 13
broadcastPkts = 0
multiPkts = 0
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 787
intervalStart = 6d:18h:38m:38s
dropEvents = 0
octets = 1791
pkts = 16
broadcastPkts = 1
multiPkts = 0
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 788
intervalStart = 6d:18h:39m:38s
dropEvents = 0
octets = 432
pkts = 6
broadcastPkts = 0
```

```
multiPkts = 0
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 789
intervalStart = 6d:18h:40m:38s
dropEvents = 0
octets = 432
pkts = 6
broadcastPkts = 0
multiPkts = 0
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 790
intervalStart = 6d:18h:41m:38s
dropEvents = 0
octets = 86734
pkts = 934
broadcastPkts = 32
multiPkts = 23
crcAlignErrors = 0
```

	underSizePkts = 0 overSizePkts = 0 fragments = 0 jabbers = 0 collisions = 0 utilization = 0

Common Errors

History control table entries are re-configured or configured history control table entries are modified.

2.4.3 Configuring RMON Alarm

Configuration Effect

Periodically monitor whether value changes of alarm variables are within the specified valid range.

Notes

If a trap message needs to be transmitted to a managing device when an alarm event is triggered, the SNMP agent must be correctly configured. For the configuration of the SNMP agent, see the *Configuring SNMP*.

If an alarm variable is a MIB variable defined in the RMON statistics group or history group, the RMON Ethernet statistics function or RMON history statistics function must be configured on the monitored Ethernet interface. Otherwise, an alarm table fails to be created.

Configuration Steps

Configuring Event Entries

- Mandatory.
- Complete the configuration in global configuration mode.

Configuring Alarm Entries

- Mandatory.
- Complete the configuration in global configuration mode.

Verification

- Run the **show rmon event** command to display the event table.
- Run the **show rmon alarm** command to display the alarm table.

Related Commands

Configuring the Event Table

Command	**rmon event** *number* [**log**] [**trap** *community*] [**description** *description-string*] [**owner** *ownername*]
Parameter Description	*number*: Indicates the index number of an event table, with the value ranging from 1 to 65,535. **log**: Indicates a log event. The system logs a triggered event. **trap** *community*: Indicates a trap event. When an event is triggered, the system transmits a trap message with the

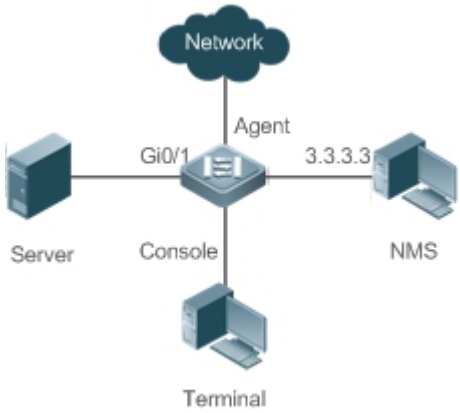
	community name of *community*. **description** *description-string*: Sets the description information about an event, that is, *description-string*. The value is a string of 1-127 characters. **owner** *ownername*: Indicates the entry creator, that is, *ownername*, which is a case-sensitive string of 1-63 characters.
Command Mode	Global configuration mode
Usage Guide	The values of configured event entry parameters can be changed, including the event type, trap community name, event description, and event creator.

Configuring the RMON Alarm Group

Command	**rmon alarm** *number variable interval* {**absolute | delta**} **rising-threshold** *value* [*event-number*] **falling-threshold** *value* [*event-number*] [**owner** *ownername*]
Parameter Description	*number*: Indicates the index number of an alarm entry, with the value ranging from 1 to 65,535. *variable*: Indicates an alarm variable, which is a string of 1-255 characters and is represented in dotted format using the node OID (format: entry.integer.instance; example: 1.3.6.1.2.1.2.1.10.1). *Interval*: Indicates the sampling interval, with the unit of seconds and the value ranging from 1 to 2,147,483,647. **absolute**: Indicates that the sampling type is absolute value sampling, that is, variable values are directly extracted when the sampling time is up. **delta**: Indicates that the sampling type is changing value sampling, that is, changes in the variable values within the sampling interval are extracted when the sampling time is up. **rising-threshold** *value*: Sets the upper limit of the sampling quantity (*value*), with the value ranging from -2,147,483,648 to +2,147,483,647. *event-number:* Indicates that an event with the event number of *event-number* is triggered when the upper limit or lower limit is reached. **falling-threshold** *value*: Sets the lower limit of the sampling quantity (*value*), with the value ranging from -2,147,483,648 to +2,147,483,647. **owner** *ownername*: Indicates the entry creator, that is, *ownername*, which is a case-sensitive string of 1-63 characters.
Command Mode	Global configuration mode
Usage Guide	Values of configured alarm entry parameters can be changed, including alarm variables, sampling type, entry creator, sampling interval, upper/lower limit of the sampling quantity, and relevant trigger events.

Configuration Example

Configuring RMON Alarm

Scenario **Figure 2- 5**	
	Assume that SNMPv1 runs on the NMS, the community name used for accessing the settings is public, with the attribute of read-write, and the IP address used by the NMS to receive trap messages is 3.3.3.3. Assume that the OID value of unknown protocol packets received by monitored interface GigabitEthernet0/3 is 1.3.6.1.2.1.2.2.1.15.3, the sampling mode is relative sampling, and the sampling interval is 60 seconds. When the relative sampling value is larger than 100 or lower than 10, event 1 and event 2 are triggered respectively. In event 1, a trap message is transmitted and the event is logged. In event 2, the event is only logged. The configuration of the RMON agent is completed on the terminal. The RMON agent is connected to the NMS and is connected to the server through interface GI0/1. The RMON agent needs to monitor the count of unknown protocol packets received by interface GI0/1. The sampling interval is 60 seconds. When the absolute sampling value is smaller than 10, the event is only logged. When the absolute sampling value is larger than 100, the event is logged and a trap message is transmitted to the NMS.
Configuration Steps	● Configure the host address for receiving trap messages. ● Configure an event group to process alarm trigger. ● Configure the alarm function.
Agent	FS# configure terminal Enter configuration commands, one per line. End with CNTL/Z. FS(config)# snmp-server community public rw FS(config)# snmp-server host 3.3.3.3 trap public FS(config)# rmon event 1 description rising-threshold-event log trap public owner admin FS(config)# rmon event 2 description falling-threshold-event log owner admin FS(config)# rmon alarm 1 1.3.6.1.2.1.2.2.1.15.3 60 delta rising-threshold 100 1 falling-threshold 10 2 owner admin
Verification	● Run the **show rmon event** command to display the event table. ● Run the **show rmon alarm** command to display the alarm table.
Agent	FS# show rmon event rmon event table: index = 1 description = rising-threshold-event type = 4

```
                    community = public
                    lastTimeSent = 0d:0h:0m:0s
                    owner = admin
                    status = 1

                    index = 2
                    description = falling-threshold-event
                    type = 2
                    community =
                    lastTimeSent = 6d:19h:21m:48s
                    owner = admin
                    status = 1

rmon log table:
                    eventIndex = 2
                    index = 1
                    logTime = 6d:19h:21m:48s
                    logDescription = falling-threshold-event
FS# show rmon alarm
rmon alarm table:
          index: 1,
          interval: 60,
          oid = 1.3.6.1.2.1.2.2.1.15.3
          sampleType: 2,
          alarmValue: 0,
          startupAlarm: 3,
          risingThreshold: 100,
          fallingThreshold: 10,
          risingEventIndex: 1,
          fallingEventIndex: 2,
          owner: admin,
          stauts: 1
```

Common Errors

- The entered OID of a monitored object is incorrect, the variable corresponding to the OID does not exist, or the type is not an integer or unsigned integer.
- The upper threshold is smaller than or equal to the lower threshold.

2.5 Monitoring

Displaying

Description	Command
Displays all RMON configuration information.	**show rmon**
Displays the Ethernet statistical table.	**show rmon stats**
Displays the history control table.	**show rmon history**
Displays the alarm table.	**show rmon alarm**
Displays the event table.	**show rmon event**

3 Configuring NTP

3.1 Overview

The Network Time Protocol (NTP) is an application-layer protocol that enables network devices to synchronize time. NTP enables network devices to synchronize time with their servers or clock sources and provides high-precision time correction (the difference from the standard time is smaller than one millisecond in a LAN and smaller than decades of milliseconds in a WAN). In addition, NTP can prevent attacks by using encrypted acknowledgment.

Currently, FS devices can be used both as NTP clients and NTP servers. In other words, a FS device can synchronize time with a time server, and be used as a time server to provide time synchronization for other devices. When a FS device is used as a server, it supports only the unicast server mode.

Protocols and Standards

- RFC 1305 : Network Time Protocol (Version 3)

3.2 Applications

Application	Description
Synchronizing Time Based on an External Reference Clock Source	A device is used as a client that synchronizes time with an external clock source. After successful synchronization, it is used as a server to provide time synchronization for other devices.
Synchronizing Time Based on a Local Reference Clock Source	A device uses a local clock as a reliable NTP reference clock source and is also used as a server to provide time synchronization for other devices.

3.2.1 Synchronizing Time Based on an External Reference Clock Source

Scenario

As shown in Figure 3- 1:

- DEVICE-A is used as a reliable reference clock source to provide time synchronization for external devices.
- DEVICE-B specifies DEVICE-A as the NTP server and synchronizes time with DEVICE-A.
- After successful synchronization, DEVICE-B provides time synchronization for DEVICE-C.

Figure 3- 1

Deployment

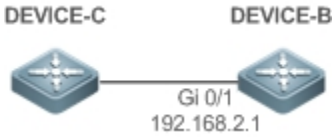
Configure DEVICE-B to the NTP external reference clock mode.

3.2.2 Synchronizing Time Based on a Local Reference Clock Source

Scenario

As shown in Figure 3- 2, DEVICE-B uses a local clock as the NTP reference clock source and provides time synchronization for DEVICE-C.

Figure 3- 2



Deployment

Configure DEVICE-B to the NTP local reference clock mode.

3.3 Features

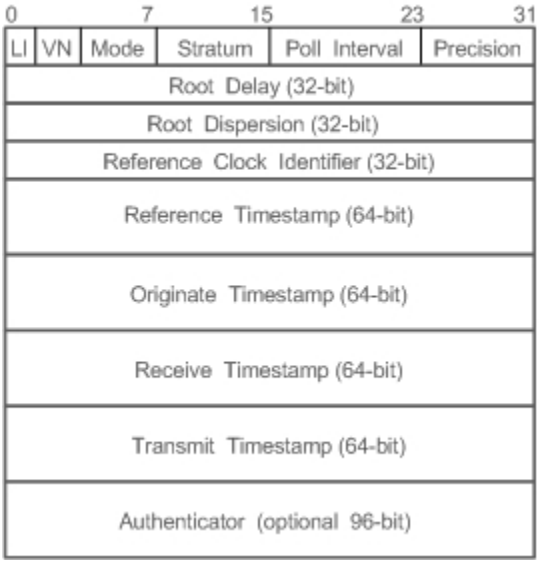
Basic Concepts

NTP Packet

As defined in RFC1305, NTP uses User Datagram Protocol (UDP) packets for transmission and the used UDP port ID is 123.

Figure 3- 3 shows the format of an NTP time synchronization packet.

Figure 3- 3 Format of an NTP Time Synchronization Packet



- Leap Indicator(LI): indicates a 2-bit leap second indicator.

00: indicates no warning information; 01: indicates that there are 61 seconds in the previous minute; 10: indicates that there are 59 seconds in the previous minute; 11: indicates that the clock is not synchronized.

- Version Number(VN): indicates a 3-bit NTP version number. The current version number is 3.
- Mode: indicates a 3-bit NTP working mode.

0: indicates no definition; 1: indicates symmetric active; 2: indicates symmetric passive; 3: indicates a client; 4: indicates a server; 5: indicates broadcasting; 6: indicates control information; 7: reserved.

- Stratum: indicates the 8-bit stratum of a local clock. 0: indicates no definition; 1: indicates the master reference clock source; other values: indicate slave reference clock sources.
- Poll Interval: indicates the poll interval (seconds), which is a 8-bit integer.

- Precision: indicates the time precision (seconds) of a local clock, which is a 8-bit integer.
- Root Delay: indicates the round-trip time to the master reference clock source, which is a 32-bit integer.
- Root Dispersion: indicates the largest difference from the master reference clock source, which is a 32-bit integer.
- Reference Clock Identifier: indicates the 32-bit identifier of a reference clock source.
- Reference Timestamp: indicates a 64-bit timestamp, namely, the time that is set or corrected at the last time.
- Originate Timestamp: indicates a 64-bit timestamp, namely, the local time when a time synchronization request leaves from a client.
- Receive Timestamp: indicates a 64-bit timestamp, namely, the local time when a time synchronization request packet arrives at a server.
- Transmit Timestamp: indicates a 64-bit timestamp, namely, the local time when a time synchronization response packet leaves from a server.
- Authenticator (optional): indicates authentication information.

NTP Server

A device uses a local clock as the reference clock source to provide time synchronization for other devices in the network.

NTP Client

A device is used as an NTP client that synchronizes time with an NTP server in the network.

Stratum

In NTP, "stratum" is used to describe the hops from a device to an authority clock source. An NTP server whose stratum is 1 has a directly connected atomic clock or radio controlled clock; an NTP server whose stratum is 2 obtains time from the server whose stratum is 1; an NTP server whose stratum is 3 obtains time from the server whose stratum is 2; and so on. Therefore, clock sources with lower stratums have higher clock precisions.

Hardware Clock

A hardware clock operates based on the frequency of the quartz crystal resonator on a device and is powered by the device battery. After the device is shut down, the hardware clock continues running. After the device is started, the device obtains time information from the hardware clock as the software time of the device.

Overview

Feature	Description
NTP Time Synchronization	Network devices synchronize time with their servers or reliable clock sources to implement high-precision time correction.
NTP Security Authentication	The NTP packet encryption authentication is used to prevent unreliable clock sources from time synchronization interference on a device.
NTP Access Control	An Access Control List (ACL) is used to filter sources of received NTP packets.

3.3.1 NTP Time Synchronization

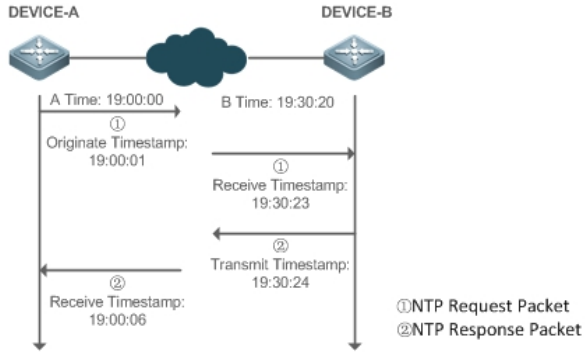
Working Principle

NTP time synchronization is implemented by interaction of NTP packets between a client and a server:

- The client sends a time synchronization packet to all servers every 64 seconds. After receiving response packets from the servers, the client filters and selects the response packets from all servers, and synchronizes time with an optimum server.
- After receiving the time synchronization request packet, a server uses the local clock as the reference source, and fills the local time information into the response packet to be sent to the client based on the protocol requirement.

Figure 3- 4 shows the format of an NTP time synchronization packet.

Figure 3- 4 Working Principle of NTP



DEVICE-B (B for short) is used as an NTP reference clock source, DEVICE-A (A for short) is used as an NTP client that synchronizes time with DEVICE-B. At a time point, the local clock of A is 19:00:00 and the local clock of B is 19:30:20.

10. A sends an NTP request packet. The local time (T0) when the packet leaves from A is 19:00:00 and is filled in Originate Timestamp.
11. After a 2-second network delay, the local time (T1) when B receives the request packet is 19:30:23 and is filled in Receive Timestamp.
12. B processes the NTP request and sends an NTP response packet one second later. The local time (T2) when the response packet leaves from B is 19:30:24 and is filled in Transmit Timestamp.
13. After a 2-second network delay, A receives the response packet. The local time (T3) when the response packet arrives at A is 19:00:06.

The specific calculations for time synchronization are as follows:

- A obtains the time difference of 30 minutes and 20 seconds between B and A by using the formula ((T1-T0)+(T2-T3))/2.
- A obtains the packet round-trip delay of four seconds between A and B by using the formula (T3-T0)-(T2-T1).

NTP Working Mode

- External clock reference mode

In this mode, a device is used as both a server and a client. If receiving time synchronization requests from other clients, the device must synchronize time with the specified server first and provide time synchronization for the clients after successful synchronization.

- Local clock reference mode

In this mode, a device uses the default local clock as the reliable clock source and provides time synchronization directly for other clients.

Related Configuration

Configuring an NTP Server

- The NTP function is disabled by default.
- Run the **ntp server** command to specify an NTP server (external clock reference source), which can enable NTP.

- After the configuration, the device works in the external clock reference mode.

Real-time Synchronization

- A device performs time synchronization every 64 seconds by default.

Updating a Hardware Clock

- By default, a device does not update synchronized time to the hardware clock.
- Run the **ntp update-calendar** command to enable a device to automatically update the hardware clock after successfully synchronizing time each time.

Configuring the NTP Master Clock

- By default, a device works in the external clock reference mode.
- Run the **ntp master** command to configure a device to the local clock reference mode.

3.3.2 NTP Security Authentication

To prevent malicious damage on an NTP server, NTP uses the authentication mechanism to check whether the time synchronization information is really from the announced server and check the information return path to provide an anti-interference protection mechanism.

Working Principle

An NTP client and an NTP server are configured with the same key. When sending request and response packets, a device calculates the hash values of the packets by using the MD5 algorithm based on the specified key and NTP packet content, and fills the hash values into the packet authentication information. The receiving device checks whether the packets are sent by a trusted device or modified based on the authentication information.

Related Configuration

Configuring a Global Security Authentication Mechanism for NTP

- By default, no NTP security authentication mechanism is enabled.
- Run the **ntp authenticate** command to enable the NTP security authentication mechanism.

Configuring a Global Authentication Key for NTP

- By default, no global authentication key is configured.
- Run the **ntp authentication-key** command to enable an NTP global authentication key.

Configuring a Globally Trusted Key ID for NTP

- By default, no globally trusted key is configured.
- Run the **ntp trusted-key** command to configure a device as the reference clock source to provide a trusted key for time synchronization externally.

Configuring a Trusted Key ID for an External Reference Clock Source

- Run the **ntp server** command to specify an external reference source and the trusted key of this clock source as well.

3.3.3 NTP Access Control

Working Principle

Provide a minimum security measure by using an ACL.

Related Configuration

Configuring the Access Control Rights for NTP Services

- By default, there is no access control right for NTP.
- Run the **ntp access-group** command to configure the access control rights for NTP.

3.4 Configuration

Configuration	Description and Command	
Configuring Basic Functions of NTP	(Mandatory) It is used to enable NTP. After NTP is enabled, a device works in the external clock reference mode.	
	ntp server	Configures an NTP server.
	ntp update-calendar	Automatically updates a hardware clock.
	(Optional) It is used to configure a device to the local clock reference mode.	
	ntp master	Configures the NTP master clock.
	(Optional) It is used to disable NTP.	
	no ntp	Disables all functions of NTP and clears all NTP configurations.
	ntp disable	Disables receiving of NTP packets from a specified interface.
Configuring NTP Security Authentication	(Optional) It is used to prevent unreliable clock sources from performing time synchronization interference on a device.	
	ntp authenticate	Enables a security authentication mechanism.
	ntp authentication-key	Configures a global authentication key.
	ntp trusted-key	Configures a trusted key for time synchronization.
	ntp server	Configures a trusted key for an external reference clock source.
Configuring NTP Access Control	(Optional) It is used to filter the sources of received NTP packets.	
	ntp access-group	Configures the access control rights for NTP.

3.4.1 Configuring Basic Functions of NTP

Configuration Effect

External Clock Reference Mode

- Use a device as a client to synchronize time from an external reference clock source to the local clock.
- After the time synchronization is successful, use the device as a time synchronization server to provide time synchronization.

Local Clock Reference Mode

- Use the local clock of a device as the NTP reference clock source to provide time synchronization.

Notes

- In the client/server mode, a device can be used as a time synchronization server to provide time synchronization only after successfully synchronizing time with a reliable external clock source.
- Once the local clock reference mode is configured, the system will not synchronize time with a clock source with a higher stratum.
- Configuring a local clock as the master clock (especially when specifying a lower stratum) may overwrite an effective clock source. If this command is used for multiple devices in a network, the clock difference between the devices may cause unstable time synchronization of the network.
- Before a local clock is configured as the master clock, if the system never synchronizes time with an external clock source, you may need to manually calibrate the system clock to ensure that there is no excessive difference. For details about how to manually calibrate the system clock, refer to the system time configuration section in the configuration guide.

Configuration Steps

Configuring an NTP Server

- (Mandatory) At least one external reference clock source must be specified (A maximum of 20 different external reference clock sources can be configured).
- If it is necessary to configure an NTP key, you must configure NTP security authentication before configuring the NTP server.

Automatically Updating a Hardware Clock

- Optional.
- By default, the system updates only the system clock, but not the hardware clock after successful time synchronization.
- After this command is configured, the system automatically updates the hardware clock after successful time synchronization.

Configuring the NTP Master Clock

- To switch a device to the local clock reference mode, run this command.

Disabling NTP

- To disable NTP and clear NTP configurations, run the **no ntp** command.
- By default, all interfaces can receive NTP packets after NTP is enabled. To disable NTP for a specified interface, run the **ntp disable** command.

Verification

- Run the **show ntp status** command to display the NTP configuration.
- Run the **show clock** command to check whether time synchronization is completed.

Related Commands

Configuring an NTP Server

Command	**ntp server** { *ip-addr* | *domain* | **ip** *domain* | **ipv6** *domain*}[**version** *version*][**source** *if-name*][**key** *keyid*][**prefer**]

Parameter Description	*ip-addr*: Indicates the IPv4/IPv6 address of the reference clock source. *domain*: Indicates the IPv4/IPv6 domain name of the reference clock source. *version*: Indicates the NTP version number, ranging from 1 to 3. *if-name*: Indicates the interface type, including AggregatePort, Dialer GigabitEthernet, Loopback, Multilink, Null, Virtual-ppp, Virtual-template and Vlan. *keyid*: Indicates the key used for communicating with the reference clock source, ranging from 1 to 4294967295. **prefer**: Indicates whether the reference clock source has a high priority.
Command Mode	Global configuration mode
Usage Guide	By default, no NTP server is configured. FS client system supports interaction with up to 20 NTP servers. You can configure an authentication key for each server (after configuring global authentication and the related key) to initiate encrypted communication with the servers. If it is necessary to configure an authentication key, you must configure NTP security authentication before configuring an NTP server. The default version of NTP for communicating with a server is NTP version 3. In addition, you can configure the source interface for transmitting NTP packets and specify that the NTP packets from a corresponding server can be received only on the transmitting interface.

Updating a Hardware Clock

Command	**ntp update-calendar**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Local Reference Clock Source

Command	**ntp master**[*stratum*]
Parameter Description	*stratum*: specifies the stratum of a local clock, ranging from 1 to 15. The default value is 8.
Command Mode	Global configuration mode
Usage Guide	N/A

Disabling NTP

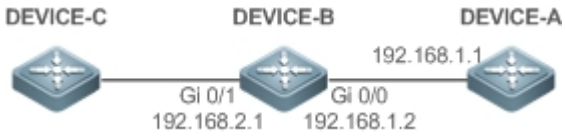
Command	**no ntp**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	This command can be used to fast disable all functions of NTP and clear all NTP configurations.

Disabling Receiving of NTP Packets on an Interface

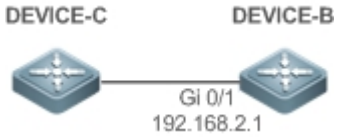
Command	**ntp disable**
Parameter Description	N/A
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuration Example

External Clock Reference Mode of NTP

Scenario **Figure 3- 5**	
	● DEVICE-B is configured to the NTP external clock reference mode. ● DEVICE-A is used as the reference clock source of DEVICE-B. ● DEVICE-C synchronizes time with DEVICE-B.
Configuration Steps	● DEVICE-A configures the local clock as the NTP reference clock source. ● DEVICE-B configures DEVICE-A as the reference clock source. ● DEVICE-C configures DEVICE-B as the reference clock source.
DEVICE-A	A#configure terminal A(config)# ntp master A(config)#exit
DEVICE-B	B#configure terminal B(config)# ntp server 192.168.1.1 B(config)# exit
DEVICE-C	C#configure terminal C(config)# ntp server 192.168.2.1 C(config)# exit
Verification	● Run the **show ntp status** command on DEVICE-B to display the NTP configuration. ● DEVICE-B sends a time synchronization packet to 192.168.1.1 in order to synchronize time with DEVICE-A. ● After successfully synchronizing time with DEVICE-A, DEVICE-B can respond to the time synchronization request from DEVICE-C. ● Run the **show clock** command on DEVICE-B and DEVICE-C to check whether the time synchronization is successful.

Local Clock Reference Mode of NTP

Scenario **Figure 3- 6**	
	● DEVICE-B configures the local clock as the NTP reference clock source. ● DEVICE-C synchronizes time with DEVICE-B.
Configuration Steps	● DEVICE-B configures the local clock as the NTP reference clock source. ● DEVICE-C configures DEVICE-B as the reference clock source.
DEVICE-B	B#configure terminal B(config)# ntp master B(config)# exit
DEVICE-C	C#configure terminal C(config)# ntp server 192.168.2.1 C(config)# exit
Verification	● Run the **show clock** command on DEVICE-C to check whether the time synchronization is successful.

3.4.2 Configuring NTP Security Authentication

Configuration Effect

Synchronizing Time from a Trusted Reference Clock Source

Use a device as a client to synchronize time only from a trusted external reference clock source to the local clock.

Providing Time Synchronization for a Trusted Device

Use the local clock of a device as the NTP reference clock source to provide time synchronization for only a trusted device.

Notes

The authentication keys of the client and server must be the same.

Configuration Steps

Configuring a Global Security Authentication Mechanism for NTP

- Mandatory.
- By default, a device disables the security authentication mechanism.

Configuring a Global Authentication Key for NTP

- Mandatory.
- By default, a device is not configured with an authentication key.

Configuring a Globally Trusted Key ID for NTP

- Optional.

- To provide time synchronization for a trusted device, you must specify a trusted authentication key by using the key ID.
- Only one trusted key can be configured. The specified authentication key must be consistent with that of the trusted device.

Configuring an Authentication Key ID for an External Reference Clock Source

- Optional.
- To synchronize time with a trusted reference clock source, you must specify a trusted authentication key by using the key ID.
- Each trusted reference clock source is mapped to an authentication key. The authentication keys must be consistent with the keys of trusted reference clock sources.

Verification

- Run the **show run** command to verify the NTP configuration.
- Run the **show clock** command to check whether time is synchronized only with a trusted device.

Related Commands

Enabling a Security Authentication Mechanism

Command	**ntp authenticate**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	By default, a client does not use a global security authentication mechanism. If no security authentication mechanism is used, communication will not be encrypted. A global security indicator is not enough to imply that the communication between the client and server is implemented in an encrypted manner. Other global keys and an encryption key for the server must also be configured for initiating encrypted communication between the client and server.

Configuring a Global Authentication Key

Command	**ntp authentication-key** *key-id* **md5** *key-string* [*enc-type*]
Parameter Description	*key-id*: indicates the ID of a global authentication key, ranging from 1 to 4294967295. *key-string*: indicates a key string. *enc-type*: (optional) indicates whether an entered key is encrypted. 0 indicates no encryption, and 7 indicates simple encryption. The default setting is no encryption.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Trusted Key for NTP

Command	**ntp trusted-key** *key-id*
Parameter Description	*key-id*: Indicates the ID of a trusted key, ranging from 1 to 4294967295.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Trusted Key for an External Reference Clock Source

Refer to the section “Related Commands”.

Configuration Example

Security Authentication

Scenario **Figure 3-7**	
	● DEVICE-B is configured to the NTP client/server mode and provides NTP services requiring security authentication for DEVICE-C. The authentication key is "abcd". ● DEVICE-A is used as the reference clock source of DEVICE-B. ● DEVICE-C synchronizes time with DEVICE-B.
Configuration Steps	● DEVICE-B configures DEVICE-A as the reference clock source. ● DEVICE-C configures DEVICE-B as the reference clock source.
DEVICE-B	B#configure terminal B(config)# ntp authentication-key 1 md5 abcd B(config)# ntp trusted-key 1 B(config)# ntp server 192.168.1.1 B(config)# exit
DEVICE-C	C#configure terminal C(config)# ntp authentication-key 1 md5 abcd C(config)# ntp server 192.168.2.1 key 1 C(config)# exit
Verification	● DEVICE-B sends a time synchronization packet that carries authentication information to 192.168.1.1 in order to synchronize time with DEVICE-A. ● Run the **show clock** command on DEVICE-B to check whether the time synchronization is successful.

3.4.3 Configuring NTP Access Control

Configuration Effect

Access control for NTP services provides a minimum security measure. A more secure method is to use an NTP authentication mechanism.

Notes

- Currently, the system does not support control query (used to control NTP servers by using network management devices, such as setting the leap second indicator or monitoring its working status). Though rule matching is implemented in the preceding sequence, no request related to control query is supported.

- If no access control rule is configured, all accesses are allowed. If any access control rule is configured, only accesses allowed by the rule can be implemented.

Related Configuration

Configuring the Access Control Rights for NTP

- Optional.
- Run the **ntp access-group** command to configure the access control rights and a corresponding ACL for NTP.

Verification

Run the **show run** command to verify the NTP configuration.

Related Commands

Configuring the Access Control Rights for NTP Services

Command	**ntp access-group** { **peer** | **serve** |**serve-only** | **query-only** }*access-list-number* | *access-list-name*
Parameter Description	**peer**: allows time request and control query for local NTP services, and allows a local device to synchronize time with a remote system (full access rights). **serve**: allows time request and control query for local NTP services, but does not allow a local device to synchronize time with a remote system. **serve-only**: allows only time request for local NTP services. **query-only**: allows only control query for local NTP services. *access-list-number*: indicates the number of an IP ACL, ranging from 1 to 99 and from 1300 to 1999. For details about how to create an IP ACL, refer to the *Configuring ACL*. *access-list-name*: indicates the name of an IP ACL. For details about how to create an IP ACL, refer to the *Configuring ACL*.
Command Mode	Global configuration mode
Usage Guide	Configure NTP access control rights. When an access request arrives, the NTP service matches rules in the sequence from the minimum access restriction to the maximum access restriction and uses the first matched rule. The matching sequence is peer, serve, serve-only, and query-only.

Configuration Example

Configuring NTP Access Control Rights

Configuration Steps	Allow only the device with the IP address of 192.168.1.1 to send a time synchronization request to a local device.
	FS(config)# access-list 1 permit 192.168.1.1 FS(config)# ntp access-group serve-only 1

3.5 Monitoring

Displaying

Description	Command
show ntp status	Displays the current NTP information.

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
debug ntp	Enables debugging.
no debug ntp	Disables debugging.

4 Configuring SNTP

4.1 Overview

The Simple Network Time Protocol (SNTP) is a simplified version of Network Time Protocol (NTP), which is used to synchronize the clocks of computers on the Internet. SNTP is applied in scenarios where it is unnecessary to use all NTP functions.

NTP uses a complex algorithm and has higher requirements for the system whereas SNTP uses a simpler algorithm and provides higher performance. Generally, SNTP precision can reach about 1s, which meets the basic requirements of most scenarios. Since SNTP packets are the same as NTP packets, the SNTP client implemented on a device is fully compatible with an NTP server.

Protocols and Standards

- RFC 2030: Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI

4.2 Applications

Application	Description
Synchronizing Time with an NTP Server	A device is used as a client to synchronize time with an NTP server.

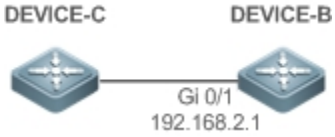
4.2.1 Synchronizing Time with an NTP Server

Scenario

As shown in Figure 4- 1, DEVICE-B uses a local clock as the NTP clock reference source and provides time synchronization for DEVICE-C.

DEVICE-C is used as an SNTP client to synchronize time with DEVICE-B.

Figure 4- 1



Deployment

- Specify DEVICE-B as the SNTP server of DEVICE-C.
- Enable SNTP for DEVICE-C.

4.3 Features

Basic Concepts

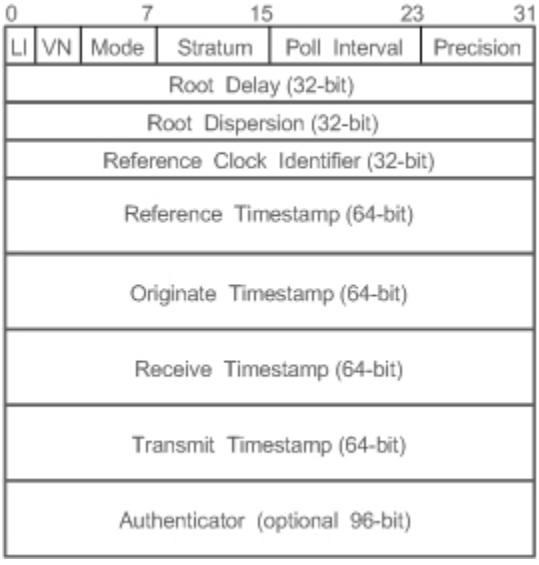
SNTP Packet

SNTPV4 is developed from NTP, which is intended to simplify the functions of NTP. It does not change the NTP specifications and the original implementation of NTP. The message format of SNTPV4 is the same as that of NTP defined in RFC1305, with only some data fields initialized into preset values.

As defined in RFC1305, SNTP uses User Datagram Protocol (UDP) packets for transmission and the used UDP port ID is 123.

Figure 4- 2 shows the format of an SNTP time synchronization packet.

Figure 4- 2 Format of an SNTP Time Synchronization Packet



- Leap Indicator(LI): indicates a 2-bit leap second indicator.

00: indicates no warning information; 01: indicates that there are 61 seconds in the previous minute; 10: indicates that there are 59 seconds in the previous minute; 11: indicates that the clock is not synchronized.

- Version Number(VN): indicates a 3-bit NTP/SNTP version number. The current version number is 3.
- Mode: indicates a 3-bit SNTP/NTP working mode.

0: indicates no definition; 1: indicates symmetric active; 2: indicates symmetric passive; 3: indicates a client; 4: indicates a server; 5: indicates broadcasting; 6: indicates control information; 7: reserved.

- Stratum: indicates the 8-bit stratum of a local clock. 0: indicates no definition; 1: indicates the master clock reference source; other values: indicate slave clock reference sources.
- Poll Interval: indicates the poll interval (seconds), which is a 8-bit integer.
- Precision: indicates the time precision (seconds) of a local clock, which is a 8-bit integer.
- Root Delay: indicates the round-trip time to the master clock reference source, which is a 32-bit integer.
- Root Dispersion: indicates the largest difference from the master reference clock source, which is a 32-bit integer.
- Reference Clock Identifier: indicates the 32-bit identifier of a reference clock source.
- Reference Timestamp: indicates a 64-bit timestamp, namely, the time that is set or corrected at the last time.
- Originate Timestamp: indicates a 64-bit timestamp, namely, the local time when a time synchronization request leaves from a client.
- Receive Timestamp: indicates a 64-bit timestamp, namely, the local time when a time synchronization request packet arrives at a server.
- Transmit Timestamp: indicates a 64-bit timestamp, namely, the local time when a time synchronization response packet leaves from a server.
- Authenticator (optional): indicates authentication information.

Overview

Feature	Description
SNTP Time Synchronization	Synchronizes time from an SNTP/NTP server to a local device.

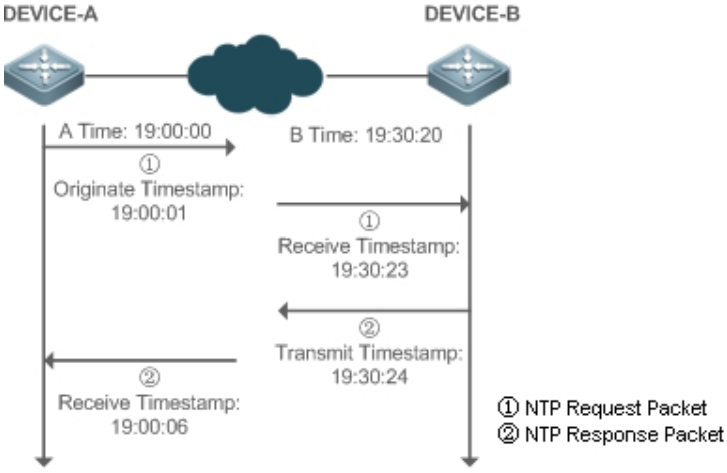
4.3.1 SNTP Time Synchronization

Working Principle

SNTP time synchronization is implemented by interaction of SNTP/NTP packets between a client and a server. The client sends a time synchronization packet to the server at intervals (half an hour by default). After receiving a response packet from the server, the client synchronizes time.

Figure 4- 3 shows the format of an SNTP time synchronization packet.

Figure 4- 3 Working Principle of SNTP



DEVICE-B (B for short) is used as an NTP reference clock source, DEVICE-A (A for short) is used as an SNTP client that synchronizes time with DEVICE-B. At a time point, the local clock of A is 19:00:00 and the local clock of B is 19:30:20.

14. A sends an SNTP/NTP request packet. The local time (T0) when the packet leaves from A is 19:00:00 and is filled in Originate Timestamp.
15. After a 2-second network delay, the local time (T1) when B receives the request packet is 19:30:23 and is filled in Receive Timestamp.
16. B processes the NTP request and sends an NTP response packet one second later. The local time (T2) when the response packet leaves from B is 19:30:24 and is filled in Transmit Timestamp.
17. After a 2-second network delay, A receives the response packet. The local time (T3) when the response packet arrives at A is 19:00:06.

The specific calculations for time synchronization are as follows:

- A obtains the time difference of 30 minutes and 20 seconds between B and A by using the formula ((T1-T0)+(T2-T3))/2.
- A obtains the packet round-trip delay of four seconds between A and B by using the formula (T3-T0)-(T2-T1).

Related Configuration

Enabling SNTP

- SNTP is disabled by default.
- Run the **sntp enable** command to enable SNTP.

Configuring an SNTP Server

- By default, no SNTP server is configured.
- Run the s**ntp server** command to specify an SNTP server.

Configuring the SNTP Time Synchronization Interval

- By default, the SNTP time synchronization interval is 1,800s.
- Run the **sntp interval** command to specify the time synchronization interval.

4.4 Configuration

Configuration	Description and Command	
Configuring SNTP	(Mandatory) It is used to enable SNTP.	
	sntp enable	Enables SNTP.
	sntp server	Configures the IP address of an SNTP server.
	(Optional) It is used to configure the SNTP time synchronization interval.	
	sntp interval	Configures the SNTP time synchronization interval.

4.4.1 Configuring SNTP

Configuration Effect

An SNTP client accesses an NTP server at fixed intervals to correct the clock regularly.

Notes

All time obtained through SNTP communication is Greenwich Mean Time (GMT). To obtain precise local time, you need to set the local time zone for alignment with GMT.

Configuration Steps

Enabling SNTP

- (Mandatory) SNTP is disabled by default.

Configuring the IP address of an SNTP Server

- (Mandatory) No SNTP/NTP server is configured by default.

Configuring the SNTP Time Synchronization Interval

- Optional.

- By default, a device synchronizes time every half an hour.

Verification

Run the **show sntp** command to display SNTP-related parameters.

Related Commands

Enabling SNTP

Command	**sntp enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	SNTP is disabled by default. Run the **no sntp enable** global configuration command to disable SNTP.

Configuring the IP address of an SNTP Server

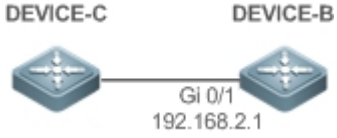
Command	**sntp server** { *ip- address* | *domain* }
Parameter Description	*ip-address*: indicates the IP address of an SNTP server. No SNTP server is configured by default. *domain:* domain name of the SNTP server. No SNTP server is configured by default.
Command Mode	Global configuration mode
Usage Guide	Since SNTP is fully compatible with NTP, the server can be configured as a public NTP server on the Internet. Since SNTP packets are the same as NTP packets, the SNTP client is fully compatible with the NTP server. There are many NTP servers on the Internet. You can select an NTP server with a shorter delay as the SNTP server on your device.

Configuring the SNTP Time Synchronization Interval

Command	**sntp interval** *seconds*
Parameter Description	*seconds*: Indicates the time synchronization interval, ranging from 60s to 65,535s. The default value is 1,800s.
Command Mode	Global configuration mode
Usage Guide	Run this command to set the interval for an SNTP client to synchronize time with an NTP/SNTP server. The interval configured here does not take effect immediately. To make it take effect immediately, run the **sntp enable** command.

Configuration Example

SNTP Time Synchronization

Scenario **Figure 4- 4**	
	- DEVICE-B indicates an NTP server on the Internet.

	● DEVICE-C synchronizes time with DEVICE-B.
Configuration Steps	Enable SNTP for DEVICE-C and configure DEVICE-B as an NTP server.
DEVICE-C	C#configure terminal C(config)# sntp server 192.168.2.1 C(config)# sntp enable C(config)# exit
Verification	● Run the **show clock** command on DEVICE-C to check whether the time synchronization is successful. ● Run the **show sntp** command on DEVICE-C to display the SNTP status and check whether the server is successfully configured.

4.5 Monitoring

Displaying

Description	Command
show sntp	Displays SNTP-related parameters.

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
debug sntp	Enables debugging.

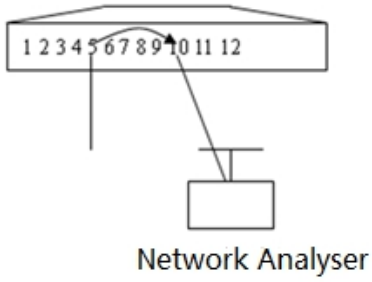
5 Configuring SPAN-RSPAN

5.1 Overview

The Switched Port Analyzer (SPAN) is to copy packets of a specified port to another switch port that is connected to a network monitoring device, so as to achieve network monitoring and troubleshooting.

All input and output packets of a source port can be monitored through SPAN. For example, as shown in the following figure, all packets on Port 5 are mapped to Port 10, and the network analyzer connected to Port 10 receives all packets that pass through Port 5.

Figure 5- 1 SPAN Configuration Instance



The SPAN function is mainly applied in network monitoring and troubleshooting scenarios, to monitor network information and rectify network faults.

The Remote SPAN (RSPAN), an extension to SPAN, is capable of remotely monitoring multiple devices. Each RSPAN session is established in a specified remote VLAN. RSPAN breaks through the limitation that a mirrored port and a mirroring port must reside on the same device, and allows a mirrored port to be several network devices away from a mirroring port. Users can observe data packets of the remote mirrored port by using an analyzer in the central equipment room.

The application scenarios of RSPAN are similar to those of SPAN. RSPAN allows users to conduct real-time data monitoring without staying in the equipment room, providing great convenience for users.

VLAN SPAN (VSPAN) considers data streams of some VLANs as data sources and mirrors them to a destination port. The configuration is similar to that of the port-based SPAN. VSPAN has the following features:

- A VLAN that is not a remote VLAN can be specified as the data source of VSPAN.
- Some VLANs that are not remote VLANs can be specified as the data sources of VSPAN.
- When a VLAN is configured as a data source, packets only in the Rx direction can be mirrored.

5.2 Applications

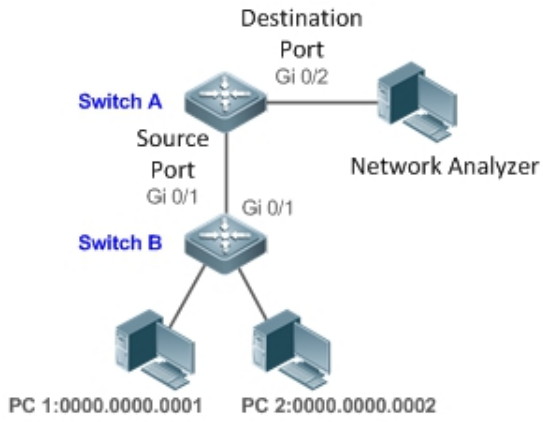
Application	Description
Stream-based SPAN	Data streams with certain characteristics need to be monitored, for example, data streams using a specified access control list (ACL) policy need to be monitored.
One-to-Many RSPAN	Multiple users need to monitor data of the same port.
RSPAN Basic Applications	Packets on the mirroring source device need to be mirrored to the destination device for monitoring.

5.2.1 Stream-based SPAN

Scenario

As shown in the following figure, the network analyzer can be configured to can monitor all data streams forwarded by Switch A to Switch B and specific data streams of Switch B (for example, data streams from PC1 and PC2).

Figure 5- 2 SPAN Simple Application Topology



Remarks	0000.0000.0001 is the MAC address of PC1. 0000.0000.0002 is the MAC address of PC2.

Deployment

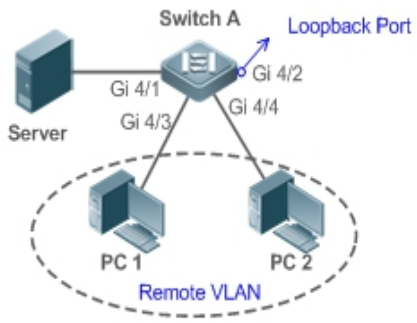
- In the preceding figure, configure the SPAN function on Switch A connected to the network analyzer, set port Gi 0/1 connected to Switch B as the SPAN source port, and set port Gi 0/2 that is directly connected to the network analyzer as the SPAN destination port.
- Configure stream-based SPAN (only data streams of PC1 and PC2 are allowed) for the source port Gi 0/1 of SPAN.

5.2.2 One-to-Many RSPAN

Scenario

As shown in the following figure, one-to-many RSPAN can be implemented on a single device, that is, both PC 1 and PC 2 can be configured to monitor the transmitted and received traffic of the port connected to the server. Users can make proper configuration (for example, remote VLAN and port MAC loopback) to monitor data streams that pass through port Gi 4/1 on PC 1 and PC 2, thereby monitoring data streams of the server.

Figure 5- 3 Application Topology of One-to-Many RSPAN



Deployment

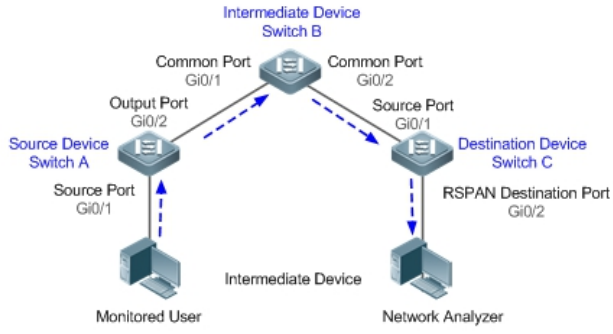
- Create a remote VLAN on Switch A.
- Configure Switch A as the source device of RSPAN and configure the port Gi 4/1 that is directly connected to the server as the RSPAN source port. Select a port that is in the Down state, Gi 4/2 in this example, as the RSPAN output port, add this port to the remote VLAN, and configure MAC loopback (run the **mac-loopback** command in interface configuration mode).
- Add ports that are directly connected to PC 1 and PC 2 to the remote VLAN.

5.2.3 RSPAN Basic Applications

Scenario

As shown in the following figure, the RSPAN function enables the network analyzer to monitor the STA connected to the source device Switch A from the destination device Switch C through the intermediate device Switch B. The devices can normally exchange data with each other.

Figure 5- 4 Basic Application Topology of RSPAN



Deployment

- Configure a remote VLAN on Switch A, Switch B, and Switch C.
- On Switch A, configure port Gi 0/1 directly connected to the STA as the source port, configure port Gi 0/2 connected to Switch B as the output port, and configure the switching function for the output port.
- On Switch B, configure port Gi 0/1 connected to Switch A and port Gi 0/2 connected to Switch C as common ports.

- On Switch C, configure port Gi0/1 connected to Switch B as a common source port, configure port Gi 0/2 connected to the network analyzer as the RSPAN destination port, and configure the switching function for the RSPAN destination port.

5.3 Features

Basic Concepts

SPAN Session

A SPAN session is data streams between the SPAN source port and the destination port, which can be used to monitor the packets of one or more ports in the input, output, or both directions. Switched ports, routed ports, and aggregate ports (APs) can be configured as source ports or destination ports of SPAN sessions. Normal operations on a switch are not affected after ports of the switch are added to a SPAN session.

Users can configure a SPAN session on a disabled port but the SPAN session is inactive. A SPAN session is in the active state only after the port on which the SPAN session is configured is enabled. In addition, a SPAN session does not take effect after a switch is powered on. It is active only after the destination port is in the operational state. Users can run the **show monitor** [**session** *session-num*] command to display the operation status of a SPAN session.

SPAN Data Streams

A SPAN session covers data streams in three directions:

- Input data streams: All packets received by a source port are copied to the destination port. Users can monitor input packets of one or more source ports in a SPAN session. Some input packets of a source port may be discarded for some reasons (for example, for the sake of port security). It does not affect the SPAN function and such packets are still mirrored to the destination port.
- Output data streams: All packets transmitted by a source port are copied to the destination port. Users can monitor output packets of one or more source ports in a SPAN session. Packets transmitted from other ports to a source port may be discarded for some reasons and such packets will not be transmitted to the destination port. The format of output packets of a source port may be changed for some reasons. For example, after routing, packets transmitted from the source port are changed in source MAC addresses, destination MAC addresses, VLAN IDs, and TTLs, and their formats are also changed after copied to the destination port.
- Bidirectional data streams: Bidirectional data streams include input data streams and output data streams. In a SPAN session, users can monitor data streams of one or more source ports in the input and output directions.

Source Port

A source port is called a monitored port. In a SPAN session, data streams of the source port are monitored for network analysis and troubleshooting. In a single SPAN session, users can monitor the input, output, and bidirectional data streams, and the number of source ports is not restricted.

A source port has the following features:

- A source port can be a switched port, routed port, or AP.
- A source port cannot be used as a destination port simultaneously.
- A source port and a destination port can belong to the same VLAN or different VLANs.

Destination Port

A SPAN session has one destination port (called a monitoring port) for receiving packets copied from a source port.

A destination port has the following features:

- A destination port can be a switched port, routed port, or AP.
- A destination port cannot be used as a source port simultaneously.

Overview

Feature	Description
SPAN	Configures mirroring of ports on the same device.
RSPAN	Configures mirroring of ports on different devices.

5.3.1 SPAN

SPAN is used to monitor data streams on switches. It copies frames on one port to another switch port that is connected to a network analyzer or RMON analyzer so as to analyze the communication of the port.

Working Principle

When a port transmits or receive packets, SPAN, after checking that the port is configured as a SPAN source port, copies the packets transmitted and received by the port to the destination port.

Configuring a SPAN Source Port

Users need to specify a SPAN session ID and source port ID to configure a SPAN source port, and set the optional SPAN direction item to determine the direction of SPAN data streams or specify an ACL policy to mirror specific data streams.

Configuring a SPAN Destination Port

Users need to specify a SPAN session ID and destination port ID to configure a SPAN destination port, and set the optional switching function item to determine whether to enable the switching function and tag removal function on the SPAN destination port.

Related Configuration

The SPAN function is disabled by default. It is enabled only after a session is created, and the SPAN source and destination ports are configured. A SPAN session can be created when a SPAN source port or destination port is configured.

Configuring a SPAN Source Port

A SPAN session does not have a SPAN source port by default. Users can run the following command to configure a SPAN source port:

monitor session *session-num* **source interface** *interface-id* [**both** | **rx** | **tx**] [**acl** *name*]

In the preceding command:

session-num: Indicates the SPAN session ID. The number of supported SPAN sessions varies with products.

interface-id: Indicates the SPAN source port to be configured.

rx: Indicates that only packets received by the source port are monitored after **rx** is configured.

tx: Indicates that only packets transmitted by the source port are monitored after **tx** is configured.

both: Indicates that packets transmitted and received by the source port are copied to the destination port for monitoring after **both** is configured, that is, **both** includes **rx** and **tx**. If none of **rx**, **tx**, and **both** is selected, **both** is enabled by default.

acl: Specifies an ACL policy. After this option is configured, packets allowed by the ACL policy on the source port are monitored. This function is disabled by default.

Configuring a SPAN Destination Port

A SPAN session does not have a SPAN destination port by default. Users can run the following command to configure a SPAN destination port:

monitor session *session-num* **destination interface** *interface-id* [**switch**]

In the preceding command:

switch: Indicates that the SPAN destination port only receives packets mirrored from the SPAN source port and discards other packets if this option is disabled, and receives both packets mirrored from the SPAN source port and packets from non-source ports if this option is enabled, that is, the communication between this destination port and other devices is not affected.

When the SPAN destination port is configured, the relevant function is disabled by default if **switch** is not configured.

Configuring Stream-based SPAN

This function is disabled by default. Users can run the **monitor session** *session-num* **source interface** *interface-id* [**rx**] **acl** *acl-name* command to configure stream-based SPAN.

Pay attention to the following points when using SPAN:

- The SPAN destination port is used for the Spanning Tree Protocol (STP) calculation.
- SPAN is unavailable if a source port or destination port is disabled.
- If a source port or destination port is added to an AP, the source port or destination port exits from a SPAN session.
- If a VLAN (or VLAN list) is used as a SPAN source, ensure that the destination port has sufficient bandwidth for receiving mirrored data of the VLAN (or VLAN list).
- Not all products support all options of the preceding commands because of product differences.

5.3.2 RSPAN

RSPAN is capable of monitoring multiple devices. Each RSPAN session is established in a specified remote VLAN. RSPAN breaks through the limitation that a mirrored port and a mirroring port must reside on the same device, and allows a mirrored port to be several network devices away from a mirroring port.

Working Principle

A remote VLAN is created for the source device, intermediate device, and destination device, all ports involved in an RSPAN session need to be added to the remote VLAN. Mirrored packets are broadcasted in the remote VLAN so that they are transmitted from the source port of the source switch to the destination port of the destination switch.

Configuring a Remote VLAN

Packets from an RSPAN source port are broadcasted in a remote VLAN so as to be copied from the local switch to the remote switch. The RSPAN source port, output port, reflection port, transparent transmission ports of the intermediate device (packet input port and output port of the intermediate device), destination port and input port of the destination port must be added to the remote VLAN. The RSPAN function requires configuring a VLAN as a remote VLAN in VLAN mode.

Configuring an RSPAN Session

The configuration of the RSPAN source port and destination port are similar to that of the SPAN source port and destination port, but the mirroring session ID specified during configuration must be the ID of an RSPAN session.

Configuring an RSPAN Source Port

The configuration of an RSPAN source port is the same as that of a SPAN source port, but the specified mirroring session ID must be the ID of an RSPAN session.

Configuring an RSPAN Output Port

The output port is located on the source device and must be added to a remote VLAN. Mirrored packets of a source port are broadcasted in this remote VLAN. The source device transmits packets to the intermediate switch or destination switch through the output port.

Configuring an RSPAN Destination Port

When an RSPAN destination port is configured, an RSPAN session ID, remote VLAN, and port name must be specified so that packets from the source port are copied to the destination port through the remote VLAN.

Configuring Stream-based RSPAN

RSPAN is an extension to SPAN and also supports stream-based mirroring. The configuration is the same as that of stream-based SPAN. Stream-based RSPAN does not affect normal communication.

Users can configure an ACL in the input direction of a source port on an RSPAN source device. Standard ACLs, extended ACLs, MAC ACLs, and user-defined ACLs are supported.

Users can configure a port ACL in the input direction of a source port on an RSPAN source device, and configure a port ACL in the output direction of the destination port on the RSPAN destination device. Users can also configure an ACL in the output direction of a remote VLAN on an RSPAN source switch and configure an ACL in the input direction of the remote VLAN on the RSPAN destination switch.

Configuring One-to-Many RSPAN

If data streams of one source port need to be mirrored to multiple destination ports, users can configure an RSPAN session, configure the source port of the RSPAN session as a one-to-many mirroring source port and select another Ethernet port as the forwarding port (output port on the source device).In addition, the MAC loopback function needs to be configured on the RSPAN forwarding port in interface configuration mode, the expected RSPAN output port and RSPAN forwarding port need to be added to the remote VLAN. Then, mirrored packets are looped back on the RSPAN forwarding port and then broadcasted in the remote VLAN, thereby implementing one-to-many RSPAN.

Related Configuration

The RSPAN function is disabled by default. It is enabled only after an RSPAN session is created, and a remote VLAN, RSPAN source port, and RSPAN destination port are configured.

Configuring a Remote VLAN

No remote VLAN is specified for RSPAN by default. Users can run the **remote-span** command in VLAN mode to configure a VLAN as a remote VLAN. One remote VLAN corresponds to one RSPAN session.

Configuring an RSPAN Source Device

This function is disabled by default. Users can run the **monitor session** *session-num* **remote-source** command in global configuration mode to configure a device as the remote source device of a specified RSPAN session.

Configuring an RSPAN Destination Device

This function is disabled by default. Users can run the **monitor session** *session-num* **remote-destination** command in global configuration mode to configure a device as the remote destination device of a specified RSPAN session.

Configuring an RSPAN Source Port

A source port of an RSPAN session is configured on the source device. The configuration is the same as that of a SPAN source port but an RSPAN session ID needs to be specified. This function is disabled by default.

Configuring an Output Port on the RSPAN Source Device

This function is disabled by default. Users can run the **monitor session** *session-num* **destination remote vlan** *remote-vlan* **interface** *interface-name* [**switch**] command in global configuration mode to configure an output port on the RSPAN source device. If the option **switch** is configured, the output port can participate in normal data packet switching. It is not configured by default. The output port must be added to a remote VLAN.

Configuring a Destination Port on the RSPAN Destination Device

This function is disabled by default. Users can run the **monitor session** *session-num* **destination remote vlan** *remote-vlan* **interface** *interface-name* [**switch**] command in global configuration mode to configure a destination port on the RSPAN destination device. If the option **switch** is configured, the destination port can participate in normal data packet switching. It is not configured by default. The destination port must be added to a remote VLAN.

Pay attention to the following points when using RSPAN:

- A remote VLAN must be configured on each device, their VLAN IDs must be consistent, and all ports that participate in a session must be added to the VLAN.
- It is not recommended that common ports be added to a remote VLAN.
- Do not configure a port that is connected to an intermediate switch or destination switch as an RSPAN source port. Otherwise, traffic on the network may be in chaos.

5.4 Configuration

Configuration	Description and Command	
Configuring SPAN Basic Functions	(Mandatory) It is used to create SPAN.	
	monitor session *session-num* **source interface** *interface-id* [**both | rx | tx**]	Configures a SPAN source port.
	monitor session *session-num* **destination interface** *interface-id*[**switch**]	Configures a SPAN destination port.
	monitor session *session-num* **source interface** *interface-id* **rx acl** *acl-name*	Configures stream-based SPAN.
	monitor session *session-num* **source vlan** *vlan-id* [**rx**]	Specifies a VLAN as the data source of SPAN.
	monitor session *session-num* **source filter vlan** *vlan-id-list*	Specifies some VLANs as the data sources of SPAN.
Configuring	(Mandatory) It is used to create RSPAN.	

RSPAN Basic Functions	**monitor session** *session-num* **remote-source**	Configures an RSPAN session ID and specifies a source device.
	monitor session *session-num* **remote-destination**	Configures an RSPAN session ID and specifies a destination device.
	remote-span	Configures a remote VLAN.
	monitor session *session-num* **source interface** *interface-id* [**both | rx | tx**]	Configures an RSPAN source port.
	monitor session *session-num* **destination remote vlan** *remote-vlan-id* **interface** *interface-id* [**switch**]	Configures an output port on the RSPAN source device or a destination port on the RSPAN destination device.

5.4.1 Configuring SPAN Basic Functions

Configuration Effect

- Configure a source and destination ports for a SPAN session.
- Configure a destination port to monitor any packets transmitted and received by a source port.

Notes

- If a source port or destination port is added to an AP, the source port or destination port exits from a SPAN session.
- If the switch function is disabled on a SPAN destination port, the destination port receives only mirrored packets and discards other packets that pass through the port. After the switch function is enabled, the destination port can receive non-mirrored packets.

Configuration Steps

Configuring a SPAN Session

- Global configuration mode. Mandatory.
- You can configure a SPAN session when configuring a SPAN source port or destination port, or when configuring a specified VLAN or some VLANs as a data source or data sources of SPAN.

Configuring a SPAN Source Port

- Global configuration mode. Mandatory.
- You can select the SPAN direction when configuring a SPAN source port. The **both** direction is configured by default, that is, both transmitted and received packets are monitored.

Configuring a SPAN Destination Port

Global configuration mode. Mandatory.

A SPAN session is active only when a SPAN source port is configured (or a VLAN is specified as the data source of SPAN) and a SPAN destination port is configured.

Verification

- Run the **show monitor** command or the **show running** command to verify the SPAN configuration. Alternatively, conduct packet capture analysis on the SPAN destination port and check whether the SPAN function takes effect according to the captured packets.

Related Commands

Configuring a SPAN Source Port

Command	**monitor session** *session-num* **source interface** *interface-id* [**both** | **rx** |**tx**]
Parameter Description	*session-num:* Indicates the ID of a SPAN session. *interface-id:* Indicates the interface ID. **both:** Indicates that packets in the input and output directions are monitored. It is the default value. **rx:** Indicates that packets in the input direction are monitored. **tx:** Indicates that packets in the output direction are monitored.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a SPAN Destination Port

Command	**monitor session** *session-num* **destination interface** *interface-id*[**switch**]
Parameter Description	*session-num:* Indicates the ID of a SPAN session. *interface-id*: Indicates the interface ID. **switch:** Indicates that the switching function is enabled on the SPAN destination port. It is disabled by default.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring Stream-based SPAN

Command	**monitor session** *session-num* **source interface** *interface-id* **rx acl** *acl-name*
Parameter Description	*session-num:* Indicates the ID of a SPAN session. *interface-id:* Indicates the interface ID. *acl-name:* Indicates an ACL name.
Command Mode	Global configuration mode
Usage Guide	N/A

Specifying a VLAN as the Data Source of SPAN

Command	**monitor session** *session-num* **source vlan** *vlan-id* [**rx**]
Parameter Description	*session-num:* Indicates the ID of a SPAN session. *vlan-id:* Indicates a specified VLAN ID. **rx:** Indicates that packets in the input direction are monitored.
Command Mode	Global configuration mode

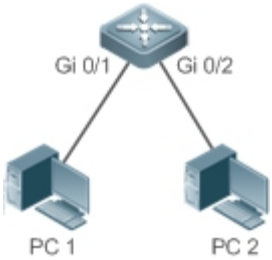
Usage Guide	N/A

Specifying Some VLANs as the Data Sources of SPAN

Command	**monitor session** *session-num* **source filter vlan** *vlan-id-list*
Parameter Description	*session-num:* Indicates the ID of a SPAN session. *vlan-id-list*: Indicates some specified VLAN IDs.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

The following uses SPAN as an example.

Scenario Figure 5- 3	
Configuration Steps	● As shown in Figure 1-5, add ports Gi 0/1 and Gi 0/2 of Device A to VLAN 1. ● Create SVI 1 and set the address of SVI 1 to 10.10.10.10/24. ● Set IP addresses of PC 1 and PC 2 to 10.10.10.1/24 and 10.10.10.2/24 respectively. ● Configure SPAN for Device A and configure ports Gi 0/1 and Gi 0/2 as the source port and destination port of SPAN respectively.
A	FS# configure FS(config)# vlan 1 FS(config-vlan)# exit FS(config)# interface vlan 1 FS(config-if-VLAN 1)# ip address 10.10.10.10 255.255.255.0 FS(config-if-VLAN 1)# exit FS(config)# monitor session 1 source interface gigabitEthernet 0/1 FS(config)# monitor session 1 destination interface gigabitEthernet 0/2
Verification	Run the **show monitor** command to check whether SPAN is configured correctly. After successful configuration, PC 1 sends ping packets to SVI 1 and PC 2 conducts monitoring by using the packet capture tool.
A	FS# show monitor sess-num: 1 span-type: LOCAL_SPAN

	src-intf: GigabitEthernet 0/1 frame-type Both dest-intf: GigabitEthernet 0/2

Common Errors

- The session ID specified during configuration of the SPAN source port is inconsistent with that specified during configuration of the SPAN destination port.
- Packet loss may occur if packets of a port with large bandwidth are mirrored to a port with small bandwidth.

5.4.2 Configuring RSPAN Basic Functions

Configuration Effect

- Configure a source port and destination port on the source device of an RSPAN session and configure the destination port on the destination device.
- Configure the destination port on the RSPAN destination device to monitor any packets that are transmitted or received by the source port.

Notes

- If a source port or destination port is added to an AP, the source port or destination port exits from a SPAN session.
- If the switch function is disabled on an RSPAN destination port, the destination port receives only mirrored packets and discards other packets that pass through the port. After the switch function is enabled, the destination port can receive non-mirrored packets.
- All ports involved in RSPAN must be added to a remote VLAN.
- A remote VLAN must be created on an intermediate device and transparent transmission ports must be added to the remote VLAN.

Configuration Steps

Configuring an RSPAN Session

- Global configuration mode. Mandatory.
- The same session ID needs to be configured on the RSPAN source device and RSPAN destination device.

Configuring an RSPAN Source Device

- Global configuration mode. Mandatory.
- It is used to specify a device to be monitored by RSPAN.

Configuring an RSPAN Destination Device

- Global configuration mode. Mandatory.
- It is used to specify the destination device for outputting RSPAN packets.

Configuring an RSPAN Source Port

- Global configuration mode. Mandatory.
- Complete the configuration on an RSPAN source device. After configuration, RSPAN monitoring can be conducted on packets of the RSPAN source port. You can specify RSPAN to monitor remote VLAN packets in the input direction, output direction, or both directions of the RSPAN source port.

Configuring an RSPAN Output Port

- Global configuration mode. Mandatory.
- Complete the configuration on an RSPAN source device. After configuration, mirrored packets received by the ports added to the remote VLAN can be transmitted to the RSPAN destination device through the output port.

Configuring an RSPAN Destination Port

- Global configuration mode. Mandatory.
- Complete the configuration on the RSPAN destination device. After configuration, the RSPAN destination device forwards mirrored packets received by the ports added to the remote VLAN to the monitoring device through the destination port.

Verification

- Run the **show monitor** command or the **show running** command to check whether RSPAN is successfully configured on each device, or conduct packet capture on the destination mirroring port on the RSPAN destination device to check whether packets mirrored from the source port of the RSPAN source device are captured.

Related Commands

Configuring an RSPAN Source Device

Command	**monitor session** *session-num* **remote-source**
Parameter Description	*session-num:* Indicates the ID of an RSPAN session.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring an RSPAN Destination Device

Command	**monitor session** *session-num* **remote-destination**
Parameter Description	*session-num:* Indicates the ID of an RSPAN session.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Remote VLAN

Command	**remote-span**
Parameter Description	N/A
Command	VLAN mode

Mode	
Usage Guide	N/A

Configuring an RSPAN Source Port

Command	**monitor session** *session-num* **source interface** *interface-id* [**both | rx | tx**][**acl** *acl-name*]
Parameter Description	*session-num:* Indicates the ID of an RSPAN session. *interface-id:* Indicates the interface ID. **both**: Indicates that packets in the input and output directions are monitored. It is the default value. **rx**: Indicates that packets in the input direction are monitored. **tx**: Indicates that packets in the output direction are monitored. *acl-name*: Indicates an ACL name.
Command Mode	Global configuration mode
Usage Guide	The configuration is the same as that of a SPAN source port but an RSPAN session ID needs to be specified.

Configuring an Output Port on the RSPAN Source Device

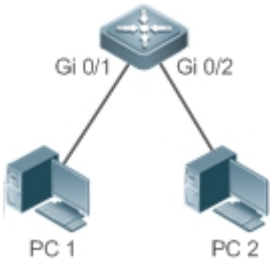
Command	**monitor session** *session-num* **destination remote vlan** *remote-vlan* **interface** *interface-id* [**switch**]
Parameter Description	*session-num:* Indicates the ID of an RSPAN session. *remote-vlan*: Indicates a remote VLAN. *interface-id:* Indicates the interface ID. **switch**: Indicates whether the port participates in packet switching.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Destination Port on the RSPAN Destination Device

Command	**monitor session** *session-num* **destination remote vlan** *remote-vlan* **interface** *interface-id* [**switch**]
Parameter Description	*session-num:* Indicates the ID of an RSPAN session. *remote-vlan*: Indicates a remote VLAN. *interface-id:* Indicates the interface ID. **switch**: Indicates whether the port participates in packet switching.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Configuring One-to-Many RSPAN

Scenario **Figure 5- 4**	
Configuration Steps	● As shown in the preceding figure, configure a remote VLAN on Switch A, Switch B, and Switch C. ● Configure the source port, output port, and MAC loopback port on Switch A. ● Configure the destination port on Switch B and Switch C.
A	FS# configure FS(config)# vlan 7 FS(config-vlan)# remote-span FS(config-vlan)# exit FS(config)# monitor session 1 remote-source FS(config)# monitor session 1 source interface fa 0/1 both FS(config)# monitor session 1 destination remote vlan 7 interface fa 0/2 switch FS(config)# interface fa0/2 FS(config-if-FastEthernet 0/2)# mac-loopback FS(config-if)# switchport access vlan 7 FS(config-if)# exit FS(config)# interface range fa0/3-4 FS(config-if-range)# switchport mode trunk
B, C	FS(config)# vlan 7 FS(config-vlan)# remote-span FS(config-vlan)# exit FS(config)# monitor session 1 remote-destination FS(config)# monitor session 1 destination remote vlan 7 interface fa 0/2 FS(config)# interface fa0/1 FS(config-if)#switchport mode trunk
Verification	Run the **show monitor** command or the **show running** command on Switch A, Switch B, and Switch C to check whether RSPAN is configured successfully.
A	FS# show monitor sess-num: 1

	span-type: SOURCE_SPAN src-intf: FastEthernet 0/1 frame-type Both dest-intf: FastEthernet 0/2 Remote vlan 7 mtp_switch on
B	FS# show monitor sess-num: 1 span-type: DEST_SPAN dest-intf: FastEthernet 0/2 Remote vlan 7 mtp_switch on
C	FS# show monitor sess-num: 1 span-type: DEST_SPAN dest-intf: FastEthernet 0/2 Remote vlan 7 mtp_switch on

Common Errors

- A remote VLAN must be configured on the source device, intermediate device, and destination device, and their VLAN IDs must be consistent.
- Packet loss may occur if packets of a port with large bandwidth are mirrored to a port with small bandwidth.
- One MAC loopback port and multiple output ports need to be configured to implement one-to-many RSPAN.

5.5 Monitoring

Displaying

Description	Command
Displays all mirroring sessions existing in the system.	**show monitor**
Displays a specified mirroring session.	**show monitor session** *session-id*

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs SPAN.	**debug span**

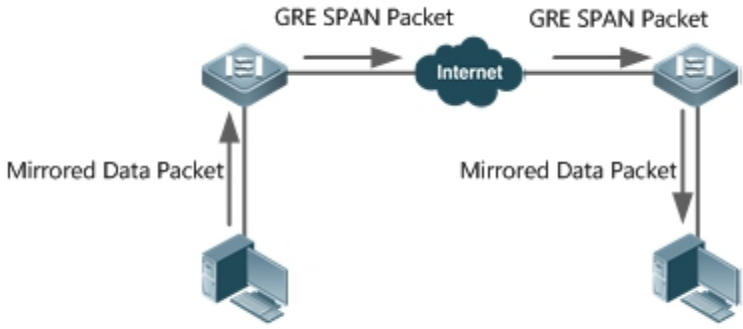
6 Configuring ERSPAN

6.1 Overview

Encapsulated Remote Switched Port Analyzer (ERSPAN) is an extension to Remote Switched Port Analyzer (RSPAN). SPAN data packets of common RSPANs can be transmitted only within Layer 2 and cannot pass through routing networks. However, an ERSPAN can transmit SPAN packets between routing networks.

An ERSPAN encapsulates all SPAN packets into IP packets through a generic routing encapsulation (GRE) tunnel, and routes them to the destination port of an RSPAN device. The following figure shows the topology of a typical application:

Figure 6- 1 Topology of a Typical ERSPAN Application



There are two kinds of roles played by the devices in the figure:

- Source switch: A source switch refers to the switch where the ERSPAN source port resides. It copies the packets on the source port, outputs the copies from the output port, encapsulates them into IP packets, and forwards the IP packets to the destination switch.
- Destination switch: A destination switch refers to the switch where the ERSPAN destination port resides. It puts the received SPAN packets through the SPAN destination port, decapsulates them into GRE packets, and then forwards the GRE packets to the monitoring device.

To implement ERSPAN, the GRE-encapsulate IP packets must be able to be normally routed to the destination SPAN device.

6.2 Applications

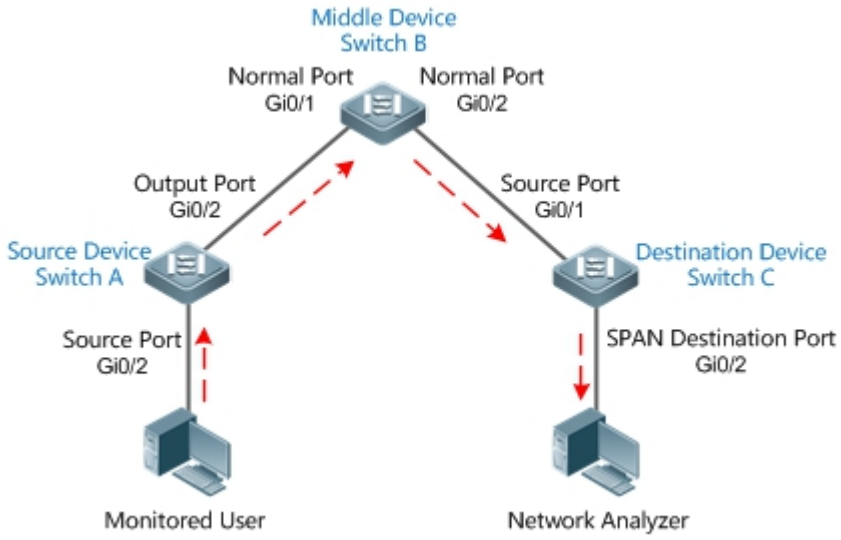
Application	Description
Basic ERSPAN Applications	Packets on the SPAN source device need to be mirrored to the destination device for monitoring.

6.2.1 Basic ERSPAN Applications

Scenario

As shown in the following figure, ERSPAN enables the network analyzer to monitor the users connected to the source device Switch A. The devices can normally exchange data with each other.

Figure 6- 2 Topology of Basic ERSPAN Applications



Deployment

- On Switch A, configure the port directly connected to users (Gi 0/1) as a source port, and configure the port connected to Switch B (Gi 0/2) as an output port.
- On Switch B, the ports connected to Switch A and Switch C (Gi 0/1 and Gi 0/2) are respectively member interfaces of switch virtual interface (SVI) interfaces of two network segments, ensuring interworking between the two IP network segments.

6.3 Features

Basic Concepts

ERSPAN Session

SPAN data packets of common RSPANs can be transmitted only within Layer 2 and cannot pass through routing networks. However, ERSPAN mirroring allows SPAN packets to be transmitted between routing networks. An ERSPAN encapsulates all SPAN packets into IP packets through a GRE tunnel, and routes them to the destination port of an RSPAN device. An ERSPAN can monitor input, output, and bidirectional packets of one or more ports. Ports such as a switched port, routed port and aggregate port (AP) can be configured as a source port for an ERSPAN session. The switch is not affected after the port is added to an ERSPAN session.

Source Port

A source port is also called a monitored port. In an ERSPAN session, data streams of the source port are monitored for network analysis and troubleshooting. In a single ERSPAN session, users can monitor the input, output, and bidirectional data streams, and the number of source ports is not limited. A source port has the following features:

- A source port can be a switched port, routed port, or an AP.
- It supports mirroring of multiple source ports on the source device to the designated output ports.
- The source port and output port cannot be on the same port; when the SPAN source port is a Layer-3 interface, both Layer-2 and Layer-3 packets are monitored.
- When multiple ports are bidirectionally monitored, a packet is input from a port and output from the other. Such monitoring is considered correct if only one packet is monitored.

- When the status of enabled Spanning Tree Protocol (STP) port is in block state, the input and output packets on the port can be monitored;
- Source port and destination port can belong to the same VLAN or different VLANs.

Overview

Feature	Description
ERSPAN	Configures SPAN on different Internet ports.

6.3.1 ERSPAN

Encapsulated ERSPAN is an extension of RSPAN. SPAN data packets of common RSPANs can be transmitted only within Layer 2 and cannot pass through routing networks. However, an ERSPAN can transmit SPAN packets between routing networks.

Working Principle

All the mirrored packets are encapsulated into IP packets through a GRE tunnel, and routed to the destination port of an RSPAN device.

Configuring an ERSPAN Session

Configure ERSPAN of the switch, and distinguish between attributes of ERSPAN switch of the device. You need to designate an ERSPAN session ID, and enter the ERSPAN configuration mode after configuration succeeds.

Configuring a Source Port

After entering the ERSPAN configuration mode, you need to name the source port to configure the SPAN source port, and determine the direction of SPAN data streams according to optional configurations of SPAN direction.

Enabling an ERSAN Session

By default, enabling an ERSPAN session is to enable ERSPAN mirroring. Only enabled ERSPAN sessions take effect.

Encapsulating the Origin IP Address

Encapsulating an origin IP address aims to configure the origin IP address of an encapsulated GRE packet.

Encapsulating the Destination IP Address

Encapsulating a destination IP address aims to configure the destination IP address of an encapsulated GRE packet and ensure normal routing of SPAN packets on the network.

Encapsulating IP TTL/DSCP

Encapsulate Time to Live (TTL) and Differentiated Services Code Point (DSCP) values of IP packets.

Related Configuration

By default, an SPAN is disabled. It is enabled only after a session is created, and source SPAN port, origin IP and destination IP addresses are configured.

Configuring an ERSPAN Session

FS(config)# **monitor session** *session_num* **erspan-source**

Wherein,

session-num: Indicates that the number of SPAN sessions supported by SPAN session IDs varies with products.

Configuring a Source Port

FS(config-mon-erspan-src)# **source interface** *single_interface* {[**rx** | **tx** | **both**]}

Wherein,

single_interface: Indicates the SPAN source port to be configured.

rx: Indicates that only the packets received by the source port are monitored after **rx** is configured.

tx: Indicates that only the packets sent from the source port are monitored after **tx** is configured.

both: Indicates that after **both** is configured, the packets sent and received by the source port are transmitted to the destination port to be monitored; that is to say, **both** includes **rx** and **tx**. If none of **rx**, **tx**, or **both** is configured, **both** is enabled by default.

Configuring Stream-based SPAN

The function is disabled by default.Run the FS(config-mon-erspan-src)# **source interface** *interface-id* **rx acl** *acl-name*

command to configure stream-based SPANs.

Enabling an ERSAN Session

FS (config-mon-erspan-src)# **shutdown**

This command is used to disable ERSPAN mirroring.(By default) Run the **no shutdown** command to enable ERSPAN mirroring.

Encapsulating the Destination IP Address

FS(config-mon-erspan-src)# **destination ip address** *ip-address*

Wherein,

ip-address: Encapsulates the destination IP address.

Encapsulating the Origin IP Address

FS(config-mon-erspan-src)# **origin ip address** *ip-address*

Wherein,

ip-address: Encapsulates the origin IP address.

Encapsulating IP TTL

FS(config-mon-erspan-src)# **ip ttl** *ttl_value*

Wherein,

ttl_value: Configures the TTL value of an encapsulated IP address. The TTL value ranges from 0 to 255, and the default value is 64.

Encapsulating IP DSCP

FS(config-mon-erspan-src)# **ip dscp** *dscp_value*

Wherein,

dscp_value: Configures the DSCP value of an encapsulated IP address. The DSCP value ranges from 0 to 63, and the default value is 0. The function takes effect only after trusting DSCP is configured on the SPAN source port.

Pay attention to the following issues during use:

- Confirm the Layer-3 routing connectivity from source switch to destination switch.

- ERSPAN is unavailable if a source port is disabled.
- If a source port or destination port is added to an AP, the source port or destination port egresses an ERSPAN session.
- As a result of product differences, not all products support all options of the above-mentioned commands.

6.4 Configuration

Configuration	Description and Command	
Configuring Basic ERSPAN Functions	(Mandatory) It is used to create ERSPAN mirroring.	
	monitor session *erspan_source_session_number* **erspan-source**	Configures an ERSPAN session ID, and enters the configuration mode of the source ERSPAN device.
	source interface *single_interface* {[**rx** | **tx** | **both**]}	Associates the source ERSPAN port, and selects an SPAN direction.
	source interface *single_interface* **rx acl** *acl-name*	Configures the stream-based SPAN source for ERSPAN.
	shutdown	Disables ERSPAN mirroring.
	destination ip address *ip_address*	Configures the destination IP address for an ERSPAN stream. The address must be the interface address of the destination device.
	original ip address *ip_address*	Configures the encapsulated origin IP address for ERSPAN.
	ip ttl *ttl_value*	(Optional) Configures the TTL value of an encapsulated IP address for ERSPAN.
	ip dscp *dscp_value*	(Optional) Configures the DSCP field value of an encapsulated IP address for ERSPAN.

6.4.1 Configuring Basic ERSPAN Functions

Configuration Effect

- RSPAN enables a network analyzer to monitor users.
- Devices can normally exchange data with each other.

Notes

- If a source port is added to an AP, the source port egresses an ERSPAN session.
- The Layer-3 routing connectivity from source switch to destination switch must be ensured.

Configuration Steps

ERSPAN Session

- Global configuration mode. Mandatory.
- The session ID configured with local SPAN or RSPAN cannot be used for an ERSPAN session. Enter the ERSPAN mode after configuration.

Source Port

- Global configuration mode. Mandatory.

- An SPAN direction can be selected during configuration of the SPAN source port. The direction is **both** by default; that is, both reception and transmission of packets are monitored.

Enabling an ERSPAN Session

- Global configuration mode. Mandatory.
- By default, enabling an ERSPAN session is to enable ERSPAN mirroring. Only enabled ERSPAN sessions take effect.

Encapsulating the Origin IP Address

- Global configuration mode. Mandatory.
- It is used to encapsulate origin IP addresses of SPAN packets.

Encapsulating the Destination IP Address

- Global configuration mode. Mandatory.
- It is used to encapsulate destination IP addresses of SPAN packets.

Encapsulating IP TTL/DSCP

- Global configuration mode. Optional.
- It is used to encapsulate DSCP values of SPAN IP packets.

Verification

- Run the **show monitor** command or the **show running** command to verify the SPAN configuration. You can also conduct packet capture analysis on the SPAN destination port and check whether SPAN takes effect according to the captured packets.

Related Commands

Configuring an ERSPAN Session

Command	**monitor session** *session_number* **erspan-source**
Parameter Description	*session-num:* Indicates the SPAN session ID.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuring a Source Port

Command	**source interface** *single_interface* {[**rx** | **tx** | **both**]}
Parameter Description	*single_interface*: Indicates the SPAN session ID. **both**: Monitors both input and output packets by default. **rx**: Monitors only input packets. **tx**: Monitors only output packets.
Command Mode	ERSPAN session mode
Usage Guide	N/A

Configuring Stream-based SPAN

Command	**source interface** *interface-id* ***rx*** **acl** *acl-name*
Parameter Description	*interface-id*: Indicates the interface name. *acl-name*: Indicates the ACL name.
Command Mode	ERSPAN session mode
Usage Guide	N/A

Enabling an ERSAN Session

Command	**shutdown**
Parameter Description	
Command Mode	ERSPAN session mode
Usage Guide	N/A

Encapsulating the Origin IP Address

Command	**original ip address** *ip_address*
Parameter Description	*ip_address:* Indicates the origin IP address to be encapsulated.
Command Mode	ERSPAN session mode
Usage Guide	

Encapsulates the Destination IP Address

Command	**destination ip address** *ip_address*
Parameter Description	*ip_address:* Indicates the destination IP address to be encapsulated.
Command Mode	ERSPAN session mode

Encapsulating IP TTL

Command	**ip ttl** *ttl_value*
Parameter Description	*ttl_value:* Configures the TTL value of an encapsulated IP address for ERSPAN.
Command Mode	ERSPAN session mode
Usage Guide	-

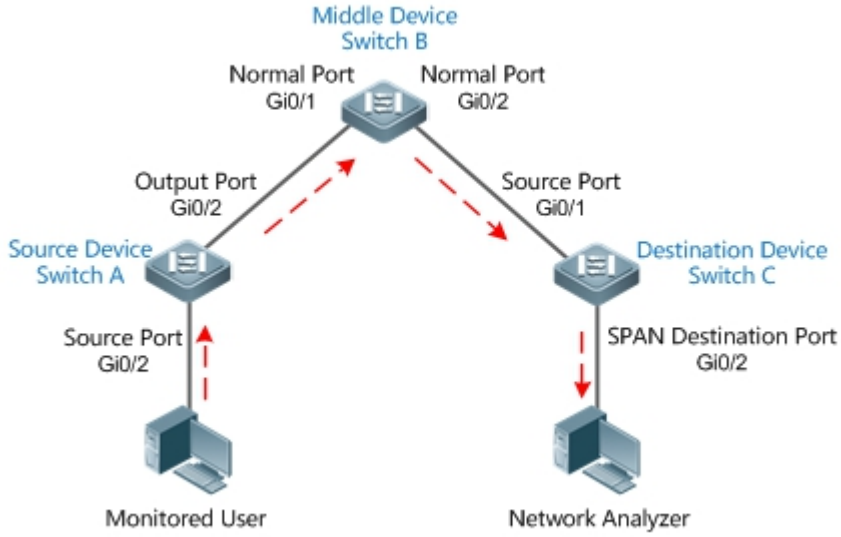
Encapsulating DSCP

Command	**ip dscp** *dscp_value*
Parameter Description	*dscp_value:* Configures the DSCP field value of an encapsulated IP address for ERSPAN.

Command Mode	ERSPAN session mode
Usage Guide	-

Configuration Example

The following uses a SPAN as an example.

Scenario Figure 6- 3	
Configuration Steps	● As shown in Figure 1-2, on Switch A, create ERSPAN Session 1 and configure it as the source device, and configure Gi 0/1 as the source port.
	SwitchA(config)#monitor session 1 erspan-source SwtichA(config-mon-erspan-src)#source interface gigabitEthernet 0/1 both SwtichA(config-mon-erspan-src)#origin ip address 10.1.1.2 SwtichA(config-mon-erspan-src)#destination ip address 12.1.1.2
Verification	Step 1: Check the configuration of the device.
	SwitchA#show running-config ! monitor session 1 erspan-src source interface GigabitEthernet 0/1 both origin ip address 10.1.1.2 destination ip address 12.1.1.2

Step 2: Check the ERSPAN information of the device.

```
SwitchA#show monitor
sess-num: 1                                          //ERSPAN Session
span-type: ERSPAN_SOURCE                             //ERSPAN source device
src-intf:                                            //ERSPAN source port information
GigabitEthernet 0/1        frame-type Both     TX status: Inactive     RX status: Inactive
dest-intf:                                           //ERPSAN output port information
GigabitEthernet 0/2
orgin ip address 10.1.1.2
destination ip address 12.1.1.2
ip ttl 64
ip dscp 0
```

Common Errors

- The session ID used to configure ERSPAN mirroring is configured with RSPAN or LOCAL SPAN.
- Layer-3 routing interworking between source switch and destination switch fails.

6.5 Monitoring

Displaying

Description	Command
Displays all SPAN sessions in the system.	**show monitor**
Displays specific SPAN sessions.	**show monitor session** *session-id*

Debugging

System resources are occupied when debugging information is output. Therefore, disable debugging immediately after use.

Description	Command
Debugs SPAN.	**debug span**

7 Configuring sFlow

7.1 Overview

sFlow is a network monitoring technology jointly developed by InMon, HP, and FoundryNetworks in 2001. This technology has been standardized. It can provide complete traffic flows of Layer 2 to Layer 4, and it is applicable to traffic analysis in the extra-large network. This technology helps users analyze the performance, trend, and existence of network traffic flows in a detailed manner in real time.

sFlow has the following advantages:

- Accurate: sFlow supports accurate monitoring of traffic on a Gigabit network or a network with higher bandwidth.
- Scalable: One sFlow Collector can monitor thousands of sFlow Agents, and it has high scalability.
- Low cost: sFlow Agent is embedded in a network device, and its cost is low.

Protocol Specification

- sFlow Version 5
- RFC 1014

7.2 Applications

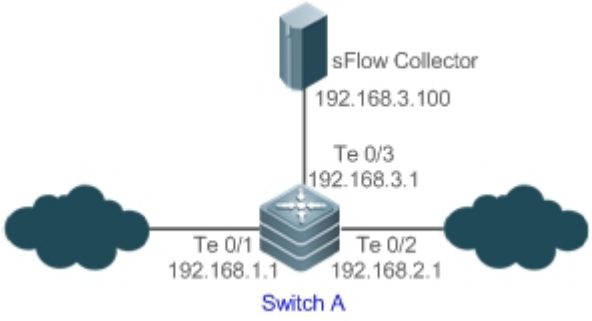
Typical Application	Scenario
Monitoring the LAN Traffic	Regard the device as an sFlow Agent, perform sampling of interface traffic in the LAN, and send the sFlow datagrams to an sFlow Collector for traffic analysis, thereby achieving the purpose of network monitoring.

7.2.1 Monitoring the LAN Traffic

Application Scenario

As shown in Figure 7- 1, start switch A that serves as an sFlow Agent, enable flow sampling and counter sampling on port Te 0/1, monitor the traffic in the 192.168.1.0 network segment, encapsulate the sampling data into sFlow datagrams at regular intervals or when the buffer is full, and sent the sFlow data to the sFlow Collector for traffic analysis.

Figure 7- 1



Function Deployment

- Configure the addresses of sFlow Agent and sFlow Collector on switch A.
- Enable flow sampling and counter sampling on port Te 0/1 of switch A.

Lots of server software supports sFlow. You can obtain software supporting sFlow at http://www.sflow.org/products/collectors.php. The software sflowtrend is free of charge.

7.3 Features

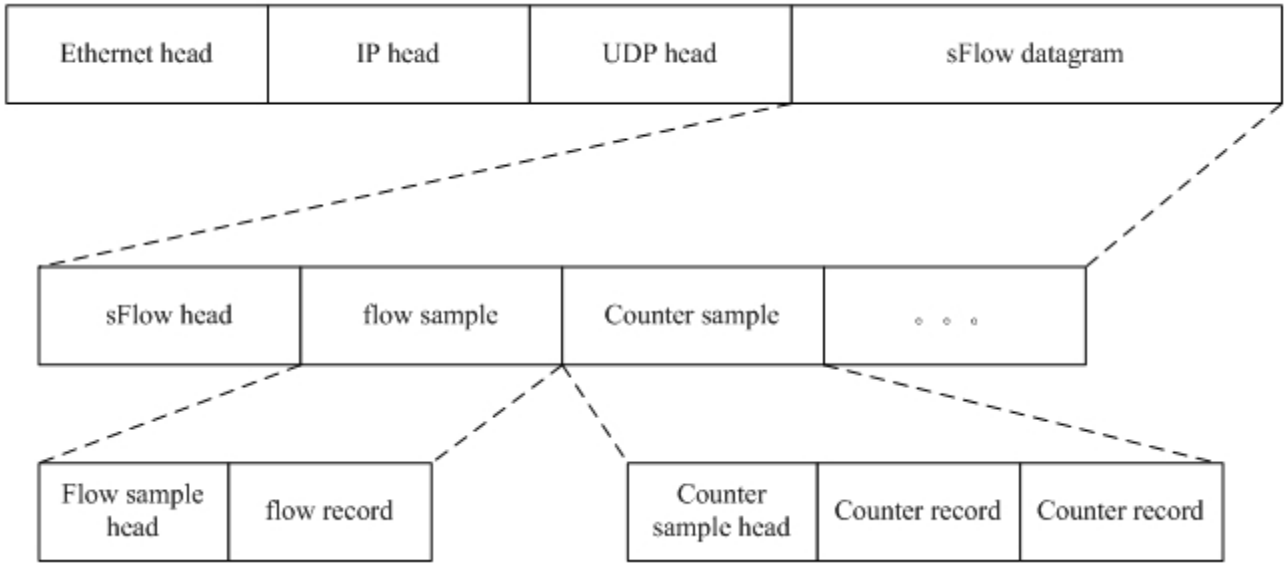
Basic Concepts

sFlow Agent

sFlow Agent is embedded in a network device. Generally, one network device can serves as an sFlow Agent. sFlow Agent can perform flow sampling and counter sampling, encapsulate sampled data into sFlow datagrams, and send the sFlow datagrams to the sFlow Collector.

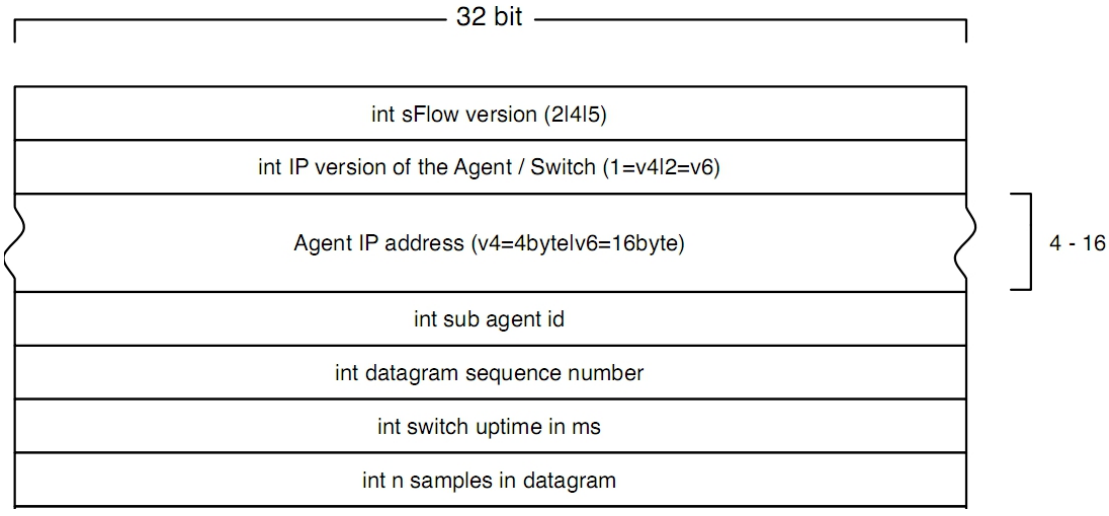
sFlow datagrams are encapsulated in UDP. Figure 7- 2 shows the sFlow datagram format.

Figure 7- 2 sFlow Datagram Format



One sFlow datagram may contain one or multiple flow samples and counter samples.

Figure 7- 3 sFlow Header



sFlow Geader Description:

Field	Description
sFlow version	sFlow version. V2, V4, and V5 are available. Currently, FS supports V5 only.
IP version of the agent/switch	IP address version of the sFlow Agent
Agent IP address	IP address of the sFlow Agent
Sub agent id	Sub-agent ID
Datagram sequence number	Serial number of the sFlow datagram
Switch uptime	Duration from the startup time of the switch to the current time
n samples in datagram	The number of samples in the an sFlow datagram. One sFlow datagram may contain one or multiple flow samples and counter samples.

sFlow Collector

sFlow Collector receives and analyzes the sFlow datagram sent from the sFlow Agent. sFlow Collector may be a PC or server. A PC or server installed with the application software for sFlow datagram analysis can be regarded as an sFlow Collector.

Flow Sampling

Based on the specified sampling rate, the sFlow Agent device performs flow sampling on the traffic flowing through an interface, including copying the header of the packet, extracting the Ethernet header and IP header of the packet, and obtaining the route information of the packet.

Figure 7- 4 Flow Sample Header

1 8	32
int data format sample data (20 bit enterprise & 12 bit format) (standard enterprise 0, formats 1)	
int sample length byte	
int sample sequence number	
int source id type (0=ifIndex| 1=smonVlanDataSource| 2=entPhysicalEntry)	int source id index value
int sampling rate	
int sample pool (total number of packets that could have been sampled)	
int drops (packets dropped due to a lack of resources)	
int input (SNMP ifIndex of input interface, 0 if not known)	
int output (SNMP ifIndex of output interface, 0 if not known) broadcast or multicast are handled as follows: the first bit indicates multiple destinations, the lower order bits number of interfaces	
int n * flow records	

Counter Sampling

In counter sampling, an sFlow Agent periodically obtains the statistics and CPU usage on a specified interface. The statistics on the interface include the number of packets input through the interface and the number of packets output through the interface.

Figure 7- 5 Counter Sample Header

1 8 32	
int data format sample data (20 bit enterprise & 12 bit format) (standard enterprise 0, formats 2)	
int sample length byte	
int sample sequence number	
int source id type (0=ifIndex| 1=smonVlanDataSource| 2=entPhysicalEntry)	int source id index value
int n * counter records	

Functions and Features

Feature	Description
Flow Sampling	Sample the traffic flowing through the interface, and send the encapsulated sFlow datagram to the sFlow Collector for analysis.
Counter Sampling	Periodically send the statistics on the interface to the sFlow Collector for analysis.

7.3.1 Flow Sampling

Sample the traffic flowing through the interface, and send the encapsulated sFlow datagram to the sFlow Collector for analysis.

Working Principle

Based on the specified sampling rate, the sFlow Agent device performs flow sampling on the traffic flowing through an interface, including copying the header of the packet, extracting the Ethernet header and IP header of the packet, and obtaining the route information of the packet. Then, the sFlow Agent encapsulates the flow sampling data into an sFlow datagram and sends the datagram to the sFlow Collector for analysis.

7.3.2 Counter Sampling

Periodically send the statistics on the interface to the sFlow Collector for analysis.

Working Principle

The sFlow Agent performs interface polling on a regular basis. For an interface whose counter sampling interval expires, the sFlow Agent obtains the statistics on this interface, encapsulates the statistics into an sFlow datagram, and sends the datagram to the sFlow Collector for analysis.

7.4 Configuration

Configuration Item	Suggestion & Related Command	
Configuring Basic Functions of sFlow	Mandatory configuration. Establish communication connections between sFlow Agent and sFlow Collector.	
	sflow agent {address | interface}	Configures the sFlow Agent address.
	sflow collector *collector-id* **destination**	Configures the sFlow Collector address.
	Mandatory configuration. Enable flow sampling and counter sampling.	
	sflow counter collector	Enables the sFlow Agent to send counter samples to the sFlow Collector.
	sflow flow collector	Enables the sFlow Agent to send flow samples to the sFlow Collector .
	sflow enable	Enables sFlow sampling for the configuration interface, that is, enables counter sampling and flow sampling.
Configuring Optional Parameters of sFlow	Optional configuration. Sets the optional parameter attributes of sFlow.	
	sflow collector *collector-id* **max-datagram-size**	Configures the maximum length of the sFlow datagram.
	sflow counter interval	Configures the counter sampling interval.
	sflow flow max-header	Configures the maximum length of the packet header copied during flow sampling.
	sflow sampling-rate	Configures the sampling rate of flow sampling.

7.4.1 Configuring Basic Functions of sFlow

Configuration Effect

- sFlow Agent and sFlow Collector can communicate with each other.
- Traffic flowing through the interface are sampled based on the default sampling rate and sent to the sFlow Collector for analysis.
- Statistics of the interface are periodically sent to the sFlow Collector based on the default sampling interval for analysis.

Notes

- Flow sampling can be configured on only physical interfaces.
- To enable the sFlow Collector to analyze the flow sampling results, the IP address of the sFlow Collector on the sFlow Agent device is required.

Configuration Steps

Configuring sFlow Agent Address

- Mandatory configuration.
- Use the **sflow agent address** command to configure the address of the sFlow Agent.
- The sFlow Agent address must be a valid address. That is, the sFlow Agent address must not be a multicast or broadcast address. It is recommended that the IP address of the sFlow Agent device be used.

Command	**sflow agent** { **address** {*ip-address* | **ipv6** *ipv6-address* }} | { **interface** { *interface-name* | **ipv6** *interface-name* }}
Parameter Description	**address**: Configures the IP address of the sFlow agent. *ip-address*: sFlow Agent IPv4 address **ipv6** *ipv6-address*: sFlow Agent IPv6 address **interface**: Configures the interface of the sFlow agent. *interface-name*: Interface of IPv4 address. **ipv6** *interface-name*: Interface of IPv6 address.
Defaults	No sFlow Agent address is configured by default
Command Mode	Global configuration mode
Configuration Usage	This command is used to configure the **Agent IP address** field in the output sFlow datagram. The datagram not configured with this filed cannot be output. The sFlow Agent address shall be a host address. When a non-host address (for example, a multicast or broadcast address) is configured as the sFlow Agent address, a message indicating configuration failure is displayed. It is recommended that the IP address of the sFlow Agent device be configured as the sFlow Agent address.

Configuring sFlow Collector Address

- Mandatory configuration.
- Use the **sflow collector** command to configure the address of the sFlow Collector.
- The sFlow Collector address must be a valid address. That is, the sFlow Collector address must not be a multicast or broadcast address. sFlow Collector must exist, and the route to it must be reachable.

Command	**sflow collector** *collector-id* **destination** { *ip-address* | **ipv6** *ipv6_address* } *udp-port*
Parameter Description	*collector-id*: sFlow Collector ID. The range is from 1 to 2. *ip-address*: sFlow Agent IPv4 address. It is not configured by default **ipv6** *ipv6-address*: sFlow Agent IPv6 address. It is not configured by default *udp-port*: sFlow Collector listening port number
Command Mode	Global configuration mode
Configuration Usage	This command is used to configure the sFlow Collector address. The sFlow Collector address shall be a host address. When a non-host address (for example, a multicast or broadcast address) is configured as the sFlow Collector address, a message indicating configuration failure is displayed. The sFlow Collector monitors the sFlow datagram on the specified port.

Enabling sFlow Samples Output to the sFlow Collector

- Mandatory configuration.
- You can use the **sflow flow collector** command to enable the sFlow Agent to send flow samples to the sFlow Collector.
- This function must be enabled on the interface to send flow samples to the sFlow Collector. In addition, sFlow Collector must exist, the route to it must be reachable, and the IP address of the corresponding sFlow Collector has been configured on the sFlow Agent device.

Command	**sflow flow collector** *collector-id*
Parameter Description	*collector-id*: sFlow Collector ID. The range is from 1 to 2.
Defaults	Sending the flow samples to the sFlow Collector is disabled by default.
Command Mode	Interface configuration mode
Configuration Usage	This command can be used for physical ports, SVI ports and sub routed ports and aggregate ports. sFlow datagrams can be output only when an IP address is configured for the corresponding sFlow Collector.

Enabling Counter Samples Output to the sFlow Collector

- Mandatory configuration.
- You can use the **sflow counter collector** command to enable the sFlow Agent to send counter samples to the sFlow Collector.
- This must be enabled on the interface to send counter samples to the sFlow Collector. In addition, sFlow Collector must exist, the route to it must be reachable, and the IP address of the corresponding sFlow Collector has been configured on the sFlow Agent device.

Command	**sflow counter collector** *collector-id*
Parameter Description	*collector-id*: sFlow Collector ID. The range is from 1 to 2.
Defaults	Sending counter samples to the sFlow Collector is disabled by default.
Command Mode	Interface configuration mode
Configuration Usage	This command can be used for physical ports, SVI ports and sub routed ports and aggregate ports. sFlow datagrams can be output only when an IP address is configured for the corresponding sFlow Collector.

Enabling Counter Sampling and Flow Sampling

- Mandatory configuration.
- You can use the **sflow enable** command to enable the flow sampling and counter sampling on an interface.
- The forwarding performance of an interface may be affected after flow sampling is enabled.

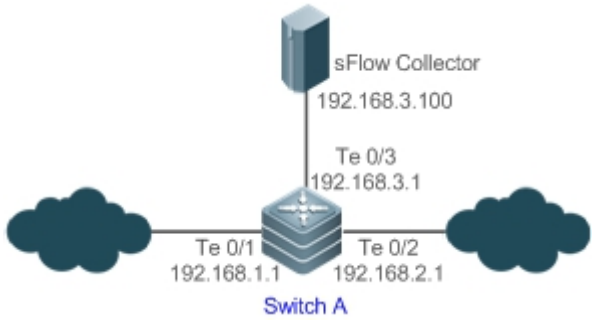
Command	**sflow enable**
Parameter Description	
Defaults	The sFlow sampling function on an interface is disabled by default.
Command Mode	Interface configuration mode
Configuration Usage	This command can be used to enable counter sampling and flow sampling for physical ports and aggregate ports.

Verification

- Use the **show sflow** command to display the sFlow configuration, and check whether the displayed information is consistent with the configuration.

Configuration Examples

Configuring Flow Sampling and Counter Sampling for sFlow Agent

Scenario Figure 7-6	
	As shown in Figure 1-6, start switch A that serves as the sFlow Agent, enable flow sampling and counter sampling on port Te 0/1, monitor the traffic in the 192.168.1.0 network segment, encapsulate the sampling traffic into sFlow datagrams at regular intervals or when the buffer is full, and send the sFlow datagrams to the sFlow Collector for traffic analysis.
Configuration Steps	• Configure 192.168.1.1 as the sFlow Agent address. • Configure 192.168.3.100 as the address of sFlow Collector 1, and 6343 as the port number. • Configure interface TenGigabitEthernet 0/1 to output flow samples and counter samples to sFlow Collector 1, and enable the sFlow sampling function on this interface.
Switch A	FS# configure terminal FS(config)# sflow agent address 192.168.1.1 FS(config)# sflow collector 1 destination 192.168.3.100 6343 FS(config)# interface TenGigabitEthernet 0/1 FS(config-if-TenGigabitEthernet 0/1)# sflow flow collector 1 FS(config-if-TenGigabitEthernet 0/1)# sflow counter collector 1 FS(config-if-TenGigabitEthernet 0/1)# sflow enable FS(config-if-TenGigabitEthernet 0/1)# end
Verification	Use the **show sflow** command to check whether the command output is consistent with the configuration.
	FS# show sflow sFlow datagram version 5 Global information: Agent IP: 192.168.1.1 sflow counter interval:30 sflow flow max-header:64 sflow sampling-rate:8192

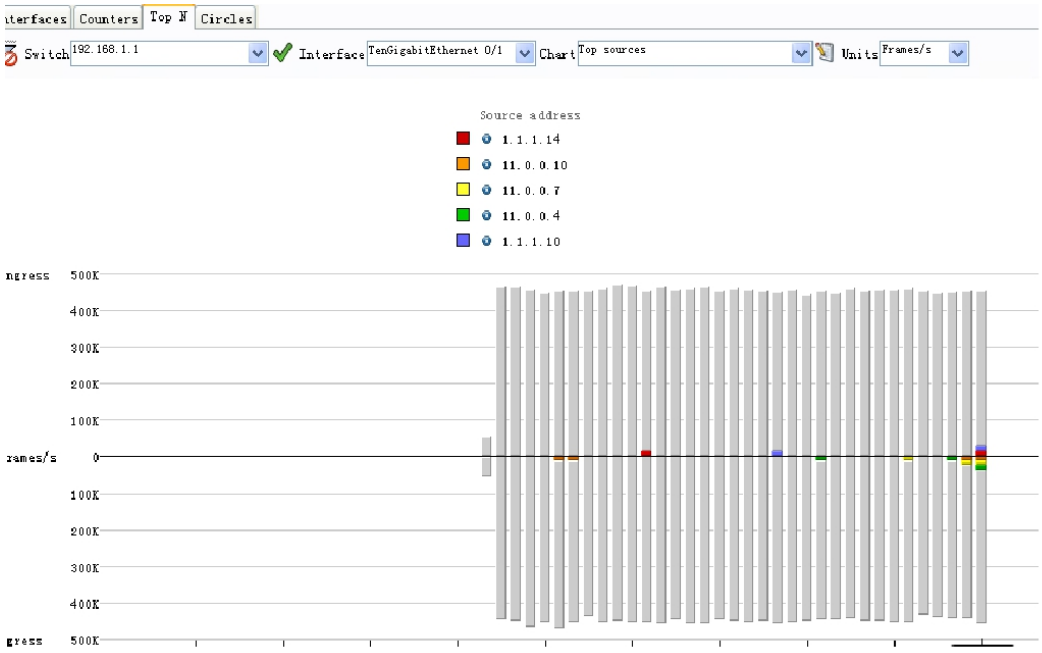
Collector information:

ID	IP	Port	Size	VPN
1	192.168.3.100	6343	1400	
2	NULL	0	1400	

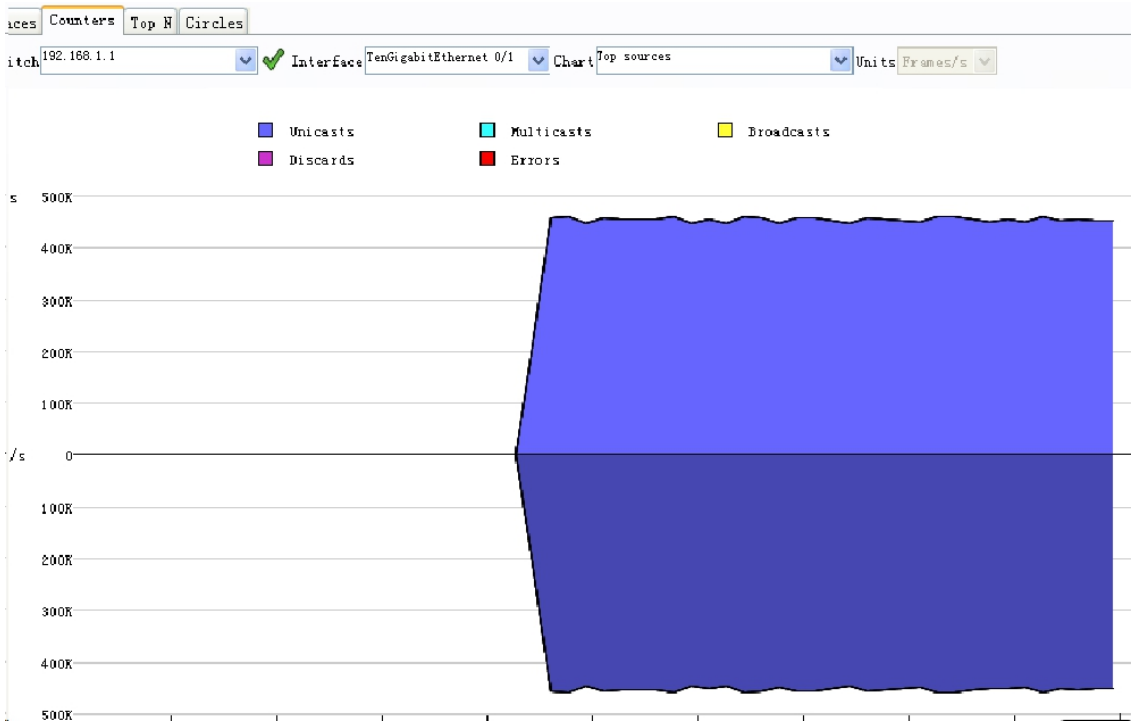
Port information

Interface	CID	FID	Enable
TenGigabitEthernet 0/1	1	1	Y

Information displayed on the sFlowTrend software:



The preceding figure shows the Top N page of the sFlowTrend software. This page displays the flow sampling results and displays the top 5 source IP addresses that involve the largest traffic. The total incoming traffic is about 450 Kpps and the total outgoing traffic is 450 Kpps, which are consistent with the actual traffic.



The preceding figure shows the counters page of the sFlowTrend software. This page displays the counter sampling results. The incoming traffic is 450 Kpps and the outgoing traffic is also 450 Kpps. In addition, all packets are unicast packets.

7.4.2 Configuring Optional Parameters of sFlow

Configuration Effect

You can adjust the data sampling accuracy by modifying relevant parameter attributes of sFlow.

Notes

- The forwarding performance may be affected when the sampling rate is too low.

Configuration Steps

Configuring the Maximum Length of the Output sFlow Datagram

- Optional configuration.
- You can use the **sflow collector** command to configure the length of the sFlow datagram, excluding the Ethernet header, IP header, and UDP header. An sFlow datagram may contain one or multiple flow samples and counter samples. Configuration of the output sFlow datagram's maximum length may lead to the result that the number of sFlow datagrams output during processing of a certain number of flow samples differs from the number of sFlow datagrams output during processing of the same number of counter packets. If the maximum length is greater than MTU, the output sFlow datagrams will be segmented.

Command	**sflow collector** *collector-id* **max-datagram-size** *datagram-size*
Parameter Description	*collector-id*: sFlow Collector ID. The range is from 1 to 2 **max-datagram-size** *datagram-size*: maximum length of the output sFlow datagram. The range is from 200 to 9,000.
Defaults	The default value is 1,400.
Command	Global configuration mode

Mode	
Configuration Usage	-

Configuring the Flow Sampling Rate

- Optional configuration.
- You can use the **sflow sampling-rate** command to configure the global flow sampling rate.
- Configuration of flow sampling rate my affect the sFlow sampling accuracy. A lower sampling rate means a higher accuracy and larger CPU consumption. Therefore, the forwarding performance of the interface may be affected when the sampling rate is low.

Command	**sflow sampling-rate** *rate*
Parameter Description	*rate*: Sampling rate of sFlow sampling. One packet is sampled from every *n* packets (*n* equals the value of **rate**). The range is from 4,096 to16,777,215.
Defaults	The default global flow sampling rate is 8,192.
Command Mode	Global configuration mode
Configuration Usage	This command is used to configure the global sampling rate of sFlow flow sampling, and sFlow flow sampling of all interfaces uses this sampling rate.

Configuring the Maximum Length of the Packet Header Copied During Flow Sampling

- Optional configuration.
- You can use the **sflow flow max-header** command to configure the length of the packet header copied during flow sampling globally.
- Users can use this command to modify the datagram information to be sent to the sFlow Collector. For example, if a user concerns about the IP header, this user can configure the length to 56 bytes. During encapsulation of flow samples, the first 56 bytes of the sample packet are copied to the sFlow datagram.

Command	**sflow flow max-header l***ength*
Parameter Description	*length*: maximum length of the packet header to be copied. The range is from 18 to 256.
Defaults	The default length of the packet header to be copied during global flow sampling is 64 bytes.
Command Mode	Global configuration mode
Configuration Usage	Configure the maximum number of bytes of the packet content copied from the header of the original packet. The copied content is recorded in the generated sample.

Configuring the Sampling Interval

- Optional configuration.
- You can use the **sflow counter interval** command to configure the global counter sampling interval.
- Enable the counter sampling interface to send the statistics on it to the sFlow Collector at the sampling interval.

Command	**sflow counter interval** *seconds*

Parameter Description	*seconds*: time interval. The range is form 3 to 2,147,483,647. The unit is second.
Defaults	The default global counter sampling interval is 30 seconds.
Command Mode	Global configuration mode
Configuration Usage	This command is used to configure the global sFlow counter sampling interval, and sFlow Counter sampling of all interfaces uses this sampling interval.

Verification

- Check whether an sFlow datagram with the flow samples is received on the sFlow Collector.
- Use the **show sflow** command to display the sFlow configuration, and check whether the displayed information is consistent with the configuration.

Configuration Examples

Configuring Optional Parameters of sFlow

Scenario	See Figure 7-6.
	- Set the flow sampling rate to 4,096 in global configuration mode. - Configure the length of the packet header copied during flow sampling to 128 bytes in global configuration mode. - Set the sampling interval to 10 in global configuration mode.
Configuration Steps	FS# configure terminal FS(config)# sflow sampling-rate 4096 FS(config)# sflow flow max-header 128 FS(config)# sflow counter interval 10
	Make traffic pass through interface TenGigabitEthernet 0/1. - Check whether there is traffic on interface TenGigabitEthernet 0/1 on sFlow Collector 1. - Use the **show sflow** command to check whether the command output is consistent with the configuration.

Verification	FS# show sflow sFlow datagram version 5 Global information: Agent IP: 10.10.10.10 sflow counter interval:10 sflow flow max-header:128 sflow sampling-rate:4096 Collector information: ID IP Port Size VPN 1 192.168.2.100 6343 1400 2 NULL 0 1400 Port information Interface CID FID Enable TenGigabitEthernet 0/1 0 1 Y

7.5 Monitoring

Displaying

Function	Command
Displays the sFlow configuration.	**show sflow**

https://www.fs.com

The information in this document is subject to change without notice. FS has made all efforts to ensure the accuracy of the information, but all information in this document does not constitute any kind of warranty.

Copyright © 2009-2022 FS.COM All Rights Reserved.