S5900 SERIES SWITCH SOFTWARE CONFIGURATION GUIDE

CONTENTS

1. Login Switch

1.1 Overview for Login Switch

System supports multiple ways to login switch: serial port, Telnet, SSH, Web browser, netmanager software.

1.2 Login Switch

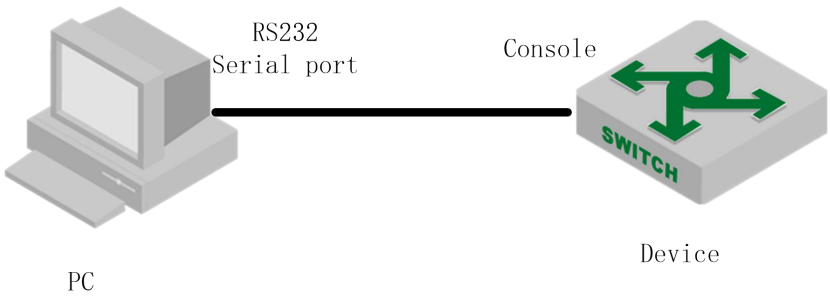
1.2.1 Login Switch via Serial Port

Login via the console port is the most basic way to login to the device.

By default, the user can login to the device directly via the serial port. The baud rate of the switch is 9600bit/s.

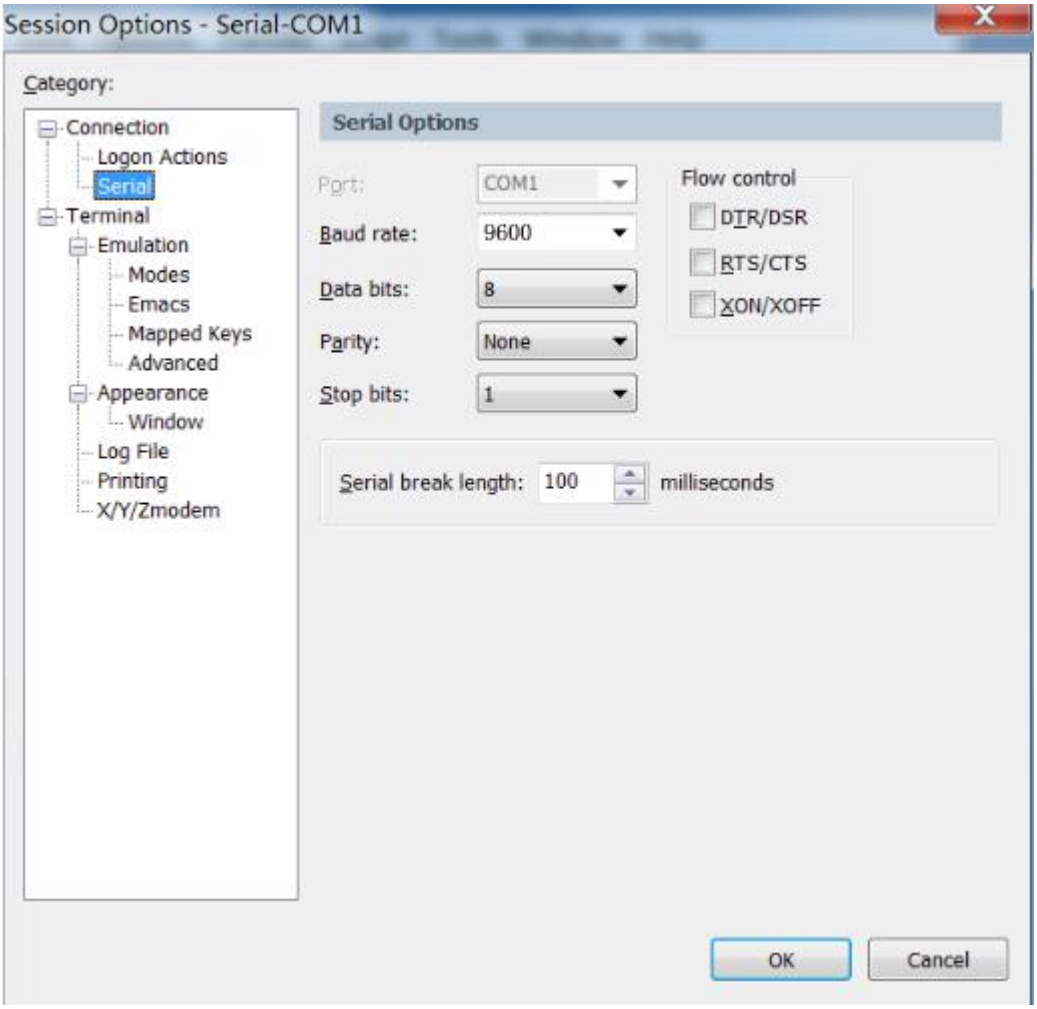
Refer to the following for specific:

(1) As shown below, use a dedicated serial cable (usually the product comes with a serial cable), insert the DB-9 connector of the serial cable into the 9-pin serial port of the PC, and then insert the RJ-45 connector into the console port of the device.



Connect the PC with the DUT via the serial cable

(2) Run the terminal software which supports serial transmission, such as HyperTerminal of SecureCRT/Windows XP. Parameter requirements: baud rate is "9600", the data bit is "8", parity is "no", stop bit is "1", the data flow control is "no", terminal emulation is "automatic detection", as shown in the figure below.



Serial connection parameters

(3) Follow the prompts to key in the user name and password and enter the switch. The default user name is admin, and the default password is 123456. It is recommended that you modify the initial password after you login to the device and remember the modified password (refer to User Management for how to modify the password).

1.2.2 Login Switch via Telnet

(1) Configure equipment to be Telnet server

Login to the switch through telnet, and the device acts as telnet-server. By default, the Telnet-server function is enabled. However, the device does not have a default IP address. That is, the client cannot log in to the device by default. So before use telnet login switch, you need to configure the switch IP via the serial port to ensure the communication between pc and DUT is normal.

Telnet-Server Configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable Telnet-server	telnet enable	optional

Disable Telnet-server	telnet disable	optional
Limit value of login users	telnet limit ***value***	optional
Display the limit value of login users	show telnet limit	optional
Display login client	show telnet client	optional
Enter execution mode	--	
Force users to go offline	stop telnet client [all | ***user-id***]	optional
Client timeout exit function	[no]timeout	optional
Client timeout configuration	timeout ***value***	optional

Note:

As telnet server, DUT will disconnect automatically if the client who login doesn't have any operations for a long time. That is, timeout out. The function is enabled by default, and the timeout value is 20m.

(2) Configure DUT acts as Telnet-client to login other device

The user has successfully logged in to the device and wants to login another device from telnet. Acting as Telnet-client, DUT should ensure the communication between telnet-client and telnet-server is normal when telnet to server.

Telnet-Client Configuration

Operation	Command	Remarks
Enter privilege configuration mode	-	-
telnet login server	telnet[6] ***server-ip*** [***port-number***] [**/localecho**]	required
Enter global configuration mode	-	
Timeout exit function	[no] telnetclient timeout	optional
Timeout configuration	telnetclient timeout ***value***	optional

Note:

As telnet server, DUT will disconnect automatically if the client who login doesn't have any operations for a long time. That is, timeout out. The function is enabled by default, and the timeout value is 20m.

1.2.3 Login Switch via SSH

The SWITCH can act as an SSH server but not an SSH client.

By default, the SSH server function is disabled. Therefore, before you login to the device via SSH, you need to login to the device via the console port, and then enable the SSH server and other attributes so as to ensure normal login to the device through SSH.

SSH Configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable/disable ssh	[no] ssh	required
Display ssh configuration	show ssh	optional
Limit the users number	[no] ssh limit ***value***	optional
Display the number of the users	show ssh limit	optional
Enter privilege configuration mode	-	
Force users to go offline	stop vty [all | ***user-id***]	optional
Configure the default key	crypto key geneate rsa	required
Remove the keyfile	crypto key zeroize rsa	optional
Activate the key	crypto key refresh	optional
Download the key from the external key server to this machine	load keyfile { public | private } tftp inet[6] ***server-ip filename***	optional
	load keyfile { public | private } ftp inet[6] ***server-ip filename username passwd***	optional
Upload the local key to the key server	upload keyfile { public | private } tftp inet[6] ***server-ip filename***	optional
	upload keyfile { public | private } ftp inet[6] ***server-ip filename username passwd***	optional
Display keyfile	show keyfile { public | private }	optional

📖 Note:

1. If you need to use ssh to login DUT, the simplest operation will be a. Open ssh; 2. Configure the default key; 3. Activate key;

2. The key file and configuration are saved in the flash and do not go into decompilation;

1.2.4 Login Switch via Web

You can log on the switch via the web. However, the web function is not perfect, so most of the functions cannot be configured on the web. So it is not recommended to manage the switch in this way.

By default, the function that the switch acts as the http server is disabled. Before login web, you need to enable the http-server function, and then configure the appropriate IP, to ensure the communication between the client and http server is normal.

WEB login configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	required
Enable http-server	**http enable [port** *tcp-port*]	optional
Disable http-server	**http disable**	optional

1.2.5 Manage Switch via Netmanager Software

The SWITCH supports login management via the NMS software. By default, the snmp-server function is enabled and the default community name can be used. Please refer to *snmp user manual* for more detailed configurations.

Snmp configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	required
Enable/disable snmp-server	**snmp-server [enable | disable]**	required

---

Note:

1. Some devices do not have the command to configure snmp-server function. Moreover, the system will be automatically enabled when the system is powered on and it cannot be disabled.

2. By default, the iso view contains two communities which are allowed to access: a. Private community with rw authority; b) public community with ro privilege;

1.3 Command Line Interface

1.3.1 Overview for Command Line Interface

System provides a series of command configurations and command line interfaces, and users can configure switch or manage switch through command line.

Here are the features of command line interface:

- Local configuration through Console port.
- Local/remote configuration through TelNet port.
- Configure protection command to ban unauthorized users from entering system.

- Users can type “?”to get help information at any time
- Provide network test command, like ping, to diagnose if the network is normal.
- Support FTP, TFTP, Xmodem, making it convenience for client to upload file or download file.

There is no need for users to type entire key words because command line interpreter adopts incomplete matching search method which can handle imprecise search requirement. For example, user can get the interface command by typing interf.

1.3.2 Command Line Mode

Protection command was adopted by command line to ban unauthorized users from entering system. Different command mode corresponds to different configuration. For example, after login, users of any level can enter normal user mode with the permission to check information of system operation; however, administrator can type ***enable*** to enter the privilege mode. Under privilege mode, he can enter global mode by typing ***configure terminal***. In the global mode, typing different configuration command corresponds to different command mode. For instance, if you type **vlan *vlan-list*,** you will enter in VLAN configuration mode.

Command line provides the following command modes:

- **User Mode**
- **Privileged Mode**
- **Global Configuration Mode**
- **Ethernet Interface Configuration Mode**
- **VLAN Configuration Mode**
- **AAA Configuration Mode**
- **RADIUS Configuration Mode**
- **Domain Configuration Mode**
- **VLAN Interface Configuration Mode**
- **SuperVLAN Interface Configuration Mode**

The features of each command mode and their entrance details are as follow:

Command Line Mode

command line mode	function	prompt	enter command line	exit command line
User Mode	show switch running status	Switch>	connect with switch, and then input username and password	Key in **exit** to disconnect with SWITCH

Privileged Mode	show switch running status, and then running system management	Switch#	Under User Mode, key in **enable**	Key in **exit** to go back to User Mode; Key in **quit** to disconnect with SWITCH
Global Configuration Mode	configure global parameters	Switch(config)#	Under Privileged Mode, key in **configure terminal**	Key in **exit**、**end** to go back to Privileged Mode; Key in **quit** to disconnect with SWITCH
Ethernet Interface Configuration Mode	configure Ethernet port parameters	Switch(config-if-ethernet-0/0/1)#	Under Global Configuration Mode, key in **interface Ethernet 0/0/1**	Key in **end** to go back to Privileged Mode; Key in **exit** to go back to Global Configuration Mode; Key in **quit** to disconnect with SWITCH
VLAN Configuration Mode	configure VLAN parameters	Switch(config-if-vlan)#	Under Global Configuration Mode, key in **vlan 2**	
AAA Configuration Mode	set up the domain	Switch(config-aaa)#	Under Global Configuration Mode, key in **aaa**	
RADIUS Configuration Mode	configure RADIUS parameters	Switch(config-radius-default)#	Under AAA configuration mode, key in **radius host default**	Key in **end** to go back to Privileged Mode Key in **exit** means go back to AAA configuration mode Key in **quit** to disconnect with SWITCH
Domain Configuration Mode	configure domain parameters	Switch(config-aaa-test.com)#	Under AAA configuration mode, key in **domain test.com**	
VLAN Interface Configuration Mode	configure VLAN L3 interface	Switch(config-if-vlanInterface-22)#	Under Global Configuration Mode, key in **interface vlan-interface 22**	Key in **end** to go back to Privileged Mode; Key in **exit** to go back to Global Configuration Mode; Key in **quit** to disconnect with SWITCH

SuperVLAN Interface Configuration Mode	configure SuperVLAN L3 interface	Switch(config-if-superVLANInterface-1)#	Under Global Configuration Mode, key in **interface supervlan-interface 1**	Key in **end** to go back to Privileged Mode; Key in **exit** to go back to Global Configuration Mode; Key in **quit** to disconnect with SWITCH

1.3.3 Comprehension of Command Syntax

This chapter mainly describes the configuration steps when enter the command line. Please read this chapter and the following chapters for detail information on how to use the command line interface.

The login authentication of switch system console is mainly for operating user identity verification via the matching recognition of the user's username and password, to allow or deny a user login.

The first step: when the following login prompt shows in the command line interface,

Username(1-32 chars):

Please key in the login user name and press the enter button, and then prompt information will be showed:

Password (1-16 chars):

Key in login password, if the password is correct, you can enter the normal user mode, the prompt will be:

Switch>

There are two different permissions in the switch system. One is administrator privilege, and the other is common user permissions. Common users generally can only see the switch configuration information, without the right to modify. However, the administrator can use the specific commands to manage switch configuration.

If login as system administrator, then enter privileged user mode from common user mode:

Switch>enable

Switch#

The second step: key in command name

If the command you key in do not require user to input parameter, you can skip to the third step. If the command you key in requires user to input parameter, please continue the following

steps:

If the command requires a parameter value, please key in the parameter value. You might have to input keywords when you key in parameter value.

Generally，command-line parameter value specifies the parameter what you need to input. It is a numerical value within a certain scope, or a string, or an IP address. If you have any questions, you can enter "?", then input the correct value according to the prompt. Keywords refer to the operand in command.

If the command requires multiple parameter values, please input keywords and each parameter value according to the command prompt until the prompt information "< enter >" appears. At that time, what you should do is just to press the “<enter>” button to end this command.

The third step: key in the complete command, and then press the enter button.

For example:

！ The user does not need to enter parameters

Switch#quit

“quit” is a command without parameter. “quit” is the command name, pressing the enter button is to execute the command after typing this command.

！Users should key in parameters

Switch(config)#vlan 3

“vlan 3”is a command that with parameter and key word. “vlan” is the Key word, and “3” is the parameter value.

1.3.4 Syntax Help

There is a build-in syntax help in the command line interface. If you are not sure of the command syntax, you can key in "?" in any command mode or get all commands and their brief description of this command mode via "help" command; key in the command string you want, with "?" following. Key in "?" after the space character, then the command line will list all the keywords begin with the string you typed. If the Location of "?" is a keyword, the command line will list all the key words and its brief description; if the Location of "?" is a parameter, the command line will list the description of relevant parameter. You can key in commands according to the reminder until the prompt command shows with "< enter >", at that time, what you should do is just to input carriage return for the command execution.

For example:

1. key in "？"under the privilege mode

Switch#?

System mode commands:

cls clear screen

help description of the interactive help

ping ping command

quit disconnect from switch and quit

……

2. key in"？"closely after keywords

Switch(config)#interf?

interface

3. input "？" after command string and the space character

Switch(config)#spanning-tree ?

bpdu-filter	bpdu-filter
bpdu-guard	bpdu-guard
forward-time	config switch delaytime
hello-time	config switch hellotime
max-age	config switch max agingtime
mode	Set state machine mode parameter
mst	Multiple spanning tree configuration
pathcost-standard	Set pathcost-standard
priority	config switch priority
remote-loop-detect	spanning-tree remote loop detect options
root-guard	STP root guard
<enter>	The command end.

4. Parameters range/ parameters format

Switch(config)#spanning-tree forward-time ?

INTEGER<4-30> switch delaytime: <4-30>(second)

5. Ending prompt command line

Switch(config)#spanning-tree ?

<enter> The command end.

1.3.5 History Command

The commands entered by users can be automatically saved by the command line interface and you can invoke or re-execute them at any time later. History command buffer is defaulted as 100. That is, the command line interface can store 100 history commands at the most for each user. You can access the last command by input "Ctrl+P"; you can access the next command by

input “Ctrl+N”

1.3.6 Types of Command Parameter

There are 5 types of parameters:

- **Integer**

The two numbers in the angle brackets (<>), connecting by hyphen (-) mean this parameter is the integer between these two numbers.

For example: INTEGER<1-10> means user can key in any integer which can be more than or equal to 1 and less than or equal to 10, such as 8.

- **IP Address**

A.B.C.D means an IP address.

For example: 192.168.0.100 is a valid IP address.

- **MAC Address**

H:H:H:H:H:H means a MAC address. If a multicast MAC address is needed, there would be corresponded prompt.

For example: 01:02:03:04:05:06 is a valid MAC address.

- **Interface List**

Interface list is prompt as STRING<3-4>. Port parameter interface-num consists of port type and port number. Port type is ethernet and port number is *device/slot-num/port-num*. *device* means stack device, and the value is 0. *slot-num* means slot number, data range is 0-1; *port-num* is the port number in the slot, data range is 1 -12. Port parameter interface-list means multiple ports. Serial ports of the same type can be connected by "to", but the port number behind the “to” must be larger than the one in the front, and this argument only can be repeated up to 3 times. The special declaration of interface parameter interface list will be displayed in the command.

For example: ***show spanning-tree interface ethernet 0/0/1 ethernet 0/0/3 to ethernet 0/0/5*** means showing the spanning-tree information about interface ethernet 0/0/1, ethernet 0/0/3, ethernet 0/0/4 and ethernet 0/5.

- **String**

The prompt STRING<1-19> means a character string which is in the length of1 to 19. Enter “?” to check the parameter description of this command.

1.3.7 Error Message

prompt	Explanation	Causes	Solution
Incomplete command	The command input is incomplete and the system cannot recognize it	There may be more than one command in the system, the system cannot recognize the abbreviation; command is not finished, need to add other parameters	Enter "?" to view the available commands or enter the complete command
Invalid parameter	Invalid parameter	Parameter out of range	Check the range of parameters and re-enter
Unrecognized command	You enter the wrong command, and the system cannot recognize	Misspelled or entered a command that does not exist	Enter "?" to view the available commands

2. Equipment Management

2.1 Overview for User Management Functions

Login the switch to perform management configuration, and it must be authentication & authorization to prevent illegal user accessing and unauthorized accessing. User management is the management on user authentication and authorization. If the authentication and authorization of the user is performed by the switch system itself, it is called local authentication. If the authentication and authorization of a user is performed by a system other than the switch (usually called an authentication server such as radius-server), it is called remote authentication. The user management content in this chapter refers to local authentication.
The switch has three types of permissions:
a. ordinary user

Ordinary users have the lowest privilege level. They can only enter the execution configuration mode, view system configuration information. However, they cannot make other configurations, and cannot modify their own passwords.

b. administrator user

Administrators not only have the rights of ordinary users, but also can configure the switch, modify their own passwords. However, they cannot create users and modify the password of other users.

c. super user

The super user is the default user of the system: admin. The system has only one super user and cannot be deleted. Super user has all permissions: it can do any switch configurations, create users, modify its own password and other users' passwords, delete the users and so forth. The default password for the admin user is 123456. Unless otherwise specified, all the configurations in this user manual are logged in as *admin*.

2.1.1 User Management Configuration

The system can create up to 15 users. After the super administrator account successfully login to the device, you can add new users, modify user passwords, modify user rights, delete accounts, limit the login methods and so forth. Ordinary users cannot modify their own passwords. Administrators can modify their own passwords, but they cannot modify other users' passwords. Super administrators can modify any user's password. Moreover, you can view the user's configurations under all modes.

Permissions: 0-1 for ordinary user, 2-15 for administrator user. Super user (admin), no configuration is required. If you do not enter a permission value when you create a user, the system will automatically assign it with normal permissions.

By default, the user can login via serial port, ssh, telnet, web terminal;

Up to five users are allowed to be online at the same time.

User Management Configuration

Operation	Command	Remarks
Enter global	configure terminal	required

configuration mode		
Create users	username ***username*** privilege ***pri-value*** password { 0|7 } ***password***	optional
The user password is saved in cipher text	service password-encryption	
Modify the user password	username change-password	optional
Modify the user permissions	username ***username*** privilege *new-pri* password **{ 0|7 }** ***password***	optional
Delete user	no username ***username***	optional
Configure the login mode	**username** *username* **terminal { all | console | ssh | telnet | web | none }**	optional
Configure the maximum number of online users	**username online-max *username value***	optional
Display the users information	**show username [***username***]**	optional
Display the online users	**show users**	optional
Enter privilege configuration mode	--	--
Force users to go offline	**stop** {*username* | **vty [all | *user-id*]** }	optional
Configure the timeout value	**[no]timeout *value***	optional

Note:

When you create a user, the password type is divided into 0 and 7, 0 means that the password is in plain text, 7 means that the password is cipher text. Therefore, when you create a user, the password type must be 0. When you configure service password-encryption, the password configured in plain text becomes decrypted in decompilation, and the decrypted password type will change into 7.

For example:
Create user *test*, and password is *123*, and save the password in plain text
switch(config)#username test privilege 0 password 0 123
Add user successfully.
Switch(config)#show running-config oam
![OAM]
username test privilege 0 password 0 123
ipaddress 192.168.1.1 255.255.255.0 0.0.0.0

Save the user password in cipher text
Switch(config)#service password-encryption
Switch(config)#sh running-config oam
![OAM]
service password-encryption
username test privilege 0 password 7 884863d2
ipaddress 192.168.1.1 255.255.255.0 0.0.0.0

By saving the user password in cipher text, you can reduce the risk of password leakage.

2.1.2 Silence Mechanism

System silence mechanism: If the times of consecutive login failures exceed the allowable value, the user is not allowed to try to log in for a certain period of time. The function is disabled by default, and the configuration for times of failed login attempts is enabled.

Silence Mechanism

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	required
Configure the times of consecutive login failures	**[no] username failmax** { ***fail-value*** **|** ***username fail-value*** }	required
Configure the silent time	**username silent-time** ***value***	optional
Display the silent configuration	**show username silent**	optional

2.2 Second-tier Password Authentication

2.2.1 Overview for Second-tier Password Authentication

Normally, the normal users can only enter the execution mode and they cannot enter other configuration modes. Moreover, normal users can only view the configurations information, and they cannot modify the configurations. Second-tier password authentication provides the mechanism to enhance the authority of normal users, if the normal users pass second-tier password authentication, it indicates that normal users have administrator privileges, that is, the normal users can have the authority to carry out other operations.

Second-tier password authentication includes local authentication and remote authentication. If user management uses local authentication, the second-tier password authentication also uses local authentication. Similarly, if user management uses remote authentication, the second-tier password authentication also uses remote authentication. User management and the password authentication use the same authentication server.

2.2.2 Configure the Second-tier Password Authentication

The second-tier password authentication function is disabled by default. If the local user (privilege level 0-1) logs in to the switch and tries to enter the privileged mode, the system prompts for the password. Enter the password of the secondary password, and then the authentication succeeds. If you are using remote authentication, when an ordinary user (privilege level 0-1) logs in to the switch and tries to enter privileged mode, the switch system automatically uses the configured username and password of second-tier password authentication to perform the authentication. If the authentication passes, the second-tier password authentication is considered successful.

When using the remote authentication, please refer to the *User Manual on Remote Authentication* for the configuration of the authentication server.

Configure the Second-tier Password Authentication

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enable/disable the function	[no] username privilege-auth	required
Configure the password of second-tier password authentication	username change-privilege-pwd { 0 | 7 } ***password***	required
Configure the username of second-tier password authentication	[no] username privilege-auth-remote-user ***username***	required
Display the configuration of second-tier password authentication	show username privilege-auth	optional

📖 Note:

If the password is selected as 0, it indicates that the password is in plain text. If you select 7, the password is cipher text. You must use the corresponding plain text for authentication.

2.2.3 Configuration Example for Second-tier Password Authentication

1.Network requirements
Normal users login to the switch via the serial port terminal and have the administrator privileges after passed the second-tier password authentication.

2.Configuration steps
#Use *admin* to login, and then create normal user: test/test
Switch(config)#username test privilege 0 password 0 test

If second-tier password authentication is not configured, log in as a normal user
Switch(config)#quit

Username:test
Password:**** (The password that was set when the user was created: test)

After successful login, tried to enter privileged mode, but failed;
Switch>en
Switch>en

using the **admin** to login so as to configure the related parameters of second-tier password authentication
enable second-tier password authentication
Switch(config)#username privilege-auth

Configure the username of second-tier password authentication（it defaults to local authentication, and the authentication is optional）
Switch(config)#username privilege-auth-remote-user test

Configure the password of second-tier password authentication (When a user enters privileged mode, the password is required)
Switch(config)#username change-privilege-pwd 0 123456
Please input your login password : **** (Check if you have configuration privileges)

Change password successfully.

Exit, and then use test/test(normal user) to login again
Switch(config)#quit
Username:test
Password:**** (The password that was set when the user was created: test)

login successfully，enter privilege configuration mode，and then you are prompted to enter a password
Switch>
enter the wrong second-tier password: test，the system prompts the password to be wrong
Switch>enable
Please input password : **** (enter the wrong second-tier password: test)
Password is error.
Switch>

Enter the correct second-tier password: 123456, successfully authenticated; and then enter privilege configuration mode and global configuration mode, you can configure other parameters.
Switch>enable
Please input password : ****** (enter the password of second-tier password authentication: 123456)
Switch#configure terminal
Switch(config)#

2.3 Remote Authentication

2.3.1 Overview for Remote Authentication

User information can be saved in the database of the switch system or saved to an external server, known as the authentication server, such as radius-server, tacacs +. The user information is stored on the local switch. When the user logs in, the switch can complete the authentication, which is called local authentication. If the user logs in, the switch cannot complete the authentication, and must authenticate via the server. This is called remote authentication. When using remote authentication, ensure that the communication between the switch and the authentication server is normal and that the user who logs in is existent on the authentication server.

2.3.2 Configure the Authentication Mode

The local authentication and remote authentication are used for the switch system authentication, and the local authentication is used by default. The remote authentication supports Radius authentication and Tacacs + authentication. Remote authentication and local authentication can be used in combination, and the remote authentication will take precedence at that time. Moreover, local authentication is attempted only when remote authentication fails.

Configure the Authentication Mode

Operation	Command	Remarks

Enter global configuration mode	configure terminal	required
Configure to use local authentication	muser local	Optional; default configurations
Configure to use RADIUS remote authentication	muser radius ***radius-name*** { pap | chap } [[account] local]	optional
Configure to use tacasc+ remote authentication	muser tacacs+ [[author] [account] [command-account] [local]]	optional
Display authentication configurations	show muser	optional

2.3.3 Configure Radius Remote Authentication

Configuration of Radius Remote Authentication

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure to use RADIUS remote authentication	muser radius ***radius-name*** { pap | chap } [[account] local]	required
Display authentication configurations	show muser	optional

Radius Server Configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enter AAA configuration mode	aaa	required
Configure the radius server name	radius host ***radius-name***	
Configure the RADIUS authentication server	{ primary-auth-ip | second-auth-ip } ***ip-address auth-port***	required
Configure the radius authentication key	auth-secret-key ***key-value***	required
Configure the RADIUS accounting server	{ primary-acct-ip | second-acct-ip } ***ip-address acct-port***	required
Configure the RADIUS accounting key	auth-secret-key ***key-value***	required
Configure the primary server to switch to the primary server after recovery	preemption-time ***value***	Optional,0 by default，means no switching
Display the configurations	show radius host [***radius-name***]	optional

Domain Configuration

Operation	Command	Remarks
Enter AAA configuration mode	aaa	required
Configure the radius domain name	domain ***domain-name***	required
Bind the domain to the radius server	radius host binding ***radius-name***	required
Activate domain	state active	required
Deactivates the domain	state block	optional default configuration

Configure the default domain	default domain-name enable ***domain-name***	optional aaa mode configuration
Delete the default domain	default domain-name disable	optional aaa mode configuration
Display the domain configuration	show domain [***domain-name***]	optional

2.3.4 Configure Tacacs+ Remote Authentication

Tacacs+ Remote Authentication Configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	required
Configure to use tacacs + remote authentication	**muser tacacs+[author][account][command-account] [local]**	required
Password encryption display function	**[no] tacacs+ encrypt-key**	optional，default display clear text
Configure the authentication type	**tacacs+ authentication-type** {**ascii | chap | pap**}	Optional,ascii by default.
Configure the tacacs + server	**tacacs+ { primary | secondary } server** *ip-address* [**encrypt-key** *value*] [**key** *value*] [**port** *port-num*] [**timeout** *value*]	required
Configure the primary server to switch to the primary server after recovery	**tacacs+ preemption-time** *value*	Optional, 0 by default，means no switching
TACACS+ configuration display	**show tacacs+**	required
Display the authentication configurations	**show muser**	optional

Note:

In the command of muser tacacs+ *[[author] [account] [command-account] [local]]*:

command-account means that all the command lines executed by the user are forwarded to the tacacs + server via the tacacs + account packet;

local means that the local authentication is attempted only when remote authentication fails.

2.3.5 Configuration Example for Remote Authentication

1.Network requirements

Here, only tacacs + is used as the authentication server. For the configuration of the RADIUS authentication server, refer to the 802.1x manual.

Before authentication, make sure that:
A. LOT can communicate with the authentication server.
B. The login user exists on the authentication server.
C. The parameters configured on the switch are the same as those of the authentication server, such as key and port number.

2. Configuration example
Configure the related parameters of tacacs + server
Configure the authentication type (the default type ascii is optional)
Switch(config)#tacacs+ authentication-type ascii

Configure the address and key of the master authentication server
Switch(config)#tacacs+ primary server 192.168.1.10 key 123456

Configure the address and key of slave authentication server (No configuration is required when there is no backup server)
Switch(config)#tacacs+ secondary server 192.168.1.11 key 123456

Configure to use the master server after the fault recovery of master server (No configuration is required when there is no backup server)
Switch(config)#tacacs+ preemption-time 20
display the information
Switch(config)#show tacacs+
Primary Server Configurations:
IP address: : 192.168.1.10
Connection port: : 49
Connection timeout: : 5
Key: : 123456

Secondary Server Configurations:
IP address: : 192.168.1.11
Connection port: : 49
Connection timeout: : 5
Key: : 123456

configure to use tacasc+ to perform remote authentication
Switch(config)#muser tacacs+

2.4 Manage IP Limit

2.4.1 Overview for IP Limit

IP limit restricts the users IP who log in to the switch, that is, only the users with specified IP can access the switch. The IP limit can improve system security.

2.4.2 Configure IP Limit

By default, there is no restriction, that is, any IP can access the switch, as long as the user name and password are correct. If you need to allow users with the specified IP to access the

switch, you must first configure them to not allow any IP access, and then configure the allowed IP addresses.

The configuration for telnet user access also applies to users via ssh.

Configure IP Limit

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Allow all IPs access	login-access-list [snmp | ssh | telnet | web] 0.0.0.0 [0.0.0.0 | 255.255.255.255]	optional default setting
Do not allow any IP access	no login-access-list [all |{ telnet | ssh | snmp | web} all]	required
Allows specified IP access	login-access-list [snmp | ssh | telnet | web] ***ip-address mask***	required
Display the configurations	show login-access-list	optional
Limit the number of users login through Telnet and entering privileged mode at the same time	login-access-list telnet-limit ***user-number***	optional 5 by default

2.4.3 Configuration Example

1.Network requirements

Switch IP=192.168.1.1/24, only allow 192.168.1.0/24 ip network segment through the Telnet login management switch.

2.Configuration steps

check default setting: all IPs can login switch through telnet.

Switch(config)#show login-access-list

sno	ipAddress	wildcard bits	terminal
1	0.0.0.0	255.255.255.255	snmp
2	0.0.0.0	255.255.255.255	web
3	**0.0.0.0**	**255.255.255.255**	**telnet**

Total [3] entry.

Do not allow any IP to login to the switch via telnet

Switch(config)#no login-access-list telnet all

Configure to allow the 192.168.1.0/24 network to access the switch via telnet

Switch(config)#login-access-list telnet 192.168.1.0 0.0.0.255

Switch(config)#show login-access-list

sno	ipAddress	wildcard bits	terminal
1	0.0.0.0	255.255.255.255	snmp
2	0.0.0.0	255.255.255.255	web
3	**192.168.1.0**	**0.0.0.255**	**telnet**

Total [3] entry.

2.5 Timeout Configuration

2.5.1 Timeout overview

The user who login via telnet, ssh, console terminal, if login for a long time, but not perform any operation, it will not only be very insecure, but also account for cpu process. Therefore, if the logged-in user does not operate for a long time, the system automatically forces it to exit, known as the timeout function.

2.5.2 Timeout Configuration

Timeout Configuration

Operation	Command	Remarks
Enter the privilege configuration mode	enable	required
Enable and configure the timeout value	timeout ***min***	Optional; Enabled by default, and the timeout value is 20m.
Disable the timeout function	no timeout	optional
Display the timeout configuration	show running-config oam	optional

---

Note:

Timeout configuration only takes the effects on the telnet, ssh, console terminal; web terminal timeout mechanism needs to be configured on the web.

---

3. Port Configuration

3.1 Port Basic Configuration

For switch devices, only Ethernet ports are supported, so the following configurations are for Ethernet ports.

3.1.1 Enter Port Configuration Mode

Enable port

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter port configuration mode	interface ethernet ***port-number***	-
Enter bulk port mode	interface range ethernet ***port-number to*** ethernet ***port-number***	

3.1.2 Enable the port

By default, all of the switch ports are enabled, that is to say, the port will be in linkup state if the switch is on-line. However, for security reasons, some certain ports will be disabled, and these ports will be in linkdown state even when the switch is on-line.

Enable the Port

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Enable the port	no shutdown	optional, default settings
disable the port	shutdown	optional
Display the port detailed configuration	show interface ethernet ***port-number***	optional
Display the port brief configuration	show interface brief ethernet [***port-number***]	optional

3.1.3 Interface Description

Interface Description

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-

Configure interface description	description *string*	optional
Delete interface description	no description	optional , default settings
Display interface description	show description interface [ethernet ***port-number***]	optional
Display the interface detailed configuration	show interface ethernet ***port-number***	optional
Display the interface brief configuration	show interface brief ethernet [***port-number***]	optional

3.1.4 Interface Rate

Interface Rate

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure interface rate	**s**peed { **1000 | 2500 | 10000 }**	optional
Restore the default rate	no speed	optional , 10000 by default
Display the interface detailed configuration	show interface ethernet ***port-number***	optional
Display the interface brief configuration	show interface brief ethernet [***port-number***]	optional

3.1.5 Rate Control Mode

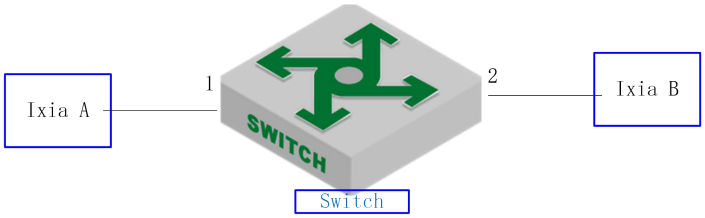
When two gigabit ports are connected and their rates are configured to be force mode, these two ports should be in master mode and slave mode respectively so as to make successful docking ports. Only a few devices need to be configured, the device which has no command means it does not need to be configured.

Rate Control Mode

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	required
Configure master mode	port-control mode master	required
Configure slave mode	port-control mode slave	required
Restore auto-negotiation mode	no port-control mode	optional，default setting
Display the configuration information	show port-control mode	optional

3.1.6 Configuration Example

1. Network requirements
Configure the description of port 1: test, modify the priority to be 7, configure the interface only accept tag frame, disable the ingress filtering.



schematic diagram of interface basic configuration

2.Configuration steps
Configure the description
Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#description test

#Modify the priority
Switch(config-if-ethernet-0/0/1)#priority 7

Configure the interface only accept tag frame
Switch(config-if-ethernet-0/0/1)#ingress acceptable-frame tagged

Disable the ingress filtering
Switch(config-if-ethernet-0/0/1)#no ingress filtering

Create vlan 100, only includes interface 2, and interface 2 adopts trunk mode.
Switch(config)#vlan 100
Switch(config-if-vlan)#switchport ethernet 0/0/2
Switch(config-if-vlan)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#switchport mode trunk

3.Result validation
（1）tester A forwards the message of vlan 100 to port 1, and the tester port can be able to receive the message of vlan 100；
（2）tester A forwards untag message to port 1, and port 1 discards the untag message.

3.2 Port Aggregation Configuration

3.2.1 Overview for Port Aggregation

Port aggregation is to aggregate multiple physical ports together to form an aggregation group to implement load balancing and redundant backup of links.

The basic configuration of the ports in an aggregation group must be consistent. The basic configuration includes STP, VLAN, port attributes and so forth.

STP configurations include STP enable / disable, STP priority, and STP cost.

VLAN configurations include: the port which is allowed to pass the VLAN, port PVID.

The configuration of the port attributes is as follows: port speed, duplex mode (must be full duplex), and link type (that is, trunk, hybrid, and access) should be corresponding.

On the same switch, if these attributes of a port in an aggregation group are modified, the remaining ports in the same aggregation group are automatically synchronized.

According to different aggregation modes, port aggregation can be classified into static aggregation and dynamic LACP aggregation.

There are three types of LACP protocol models:

Static mode (on): Does not run LACP protocol

Dynamic active mode: In active mode, the port automatically initiates LACP negotiation

Dynamic passive mode: In passive mode, a port only responds to LACP negotiation

When docking with another device, static can only dock with static, active can dock with active or passive, passive can only dock with activate.

3.2.2 Configure the Aggregation Group ID

The same port cannot join multiple ch-id at the same time. If a member exists in an aggregation group, you cannot directly delete the aggregation group. You must remove the member from the aggregation group firstly.

You can directly create an aggregation id under global configuration mode or automatically create when adding a port to an aggregation group.

Lacp id Configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the aggregation group ID	[no] channel-group ***ch-id***	optional

3.2.3 Configure the Aggregation Group

Port aggregation includes static and dynamic. “ON” refers to static aggregation while the “active” or “passive” refers to dynamic aggregation. Dynamic aggregation needs to negotiate with the other side via the lacp protocol while static aggregation does not.

A static aggregation group can have up to eight port members.

A dynamic LACP can contain up to 12 members, 8 of which are in bandl state, and the other four are in the backup state. Only members with bandl status will forward normal traffic. When the members of the *bandl* state linkdown, the backup member with the highest port priority becomes the bandl state.

Operation	Command	Remarks
Enter the interface configuration mode	interface ethernet ***port-num***	-
Add a port to an aggregation group	channel-group ***ch-id*** mode **{on | active | passive }**	required
Remove the port from the aggregation group	no channel-group ***ch-id***	optional
Display the information of the aggregation group	show lacp internal [***ch-id***]	optional

Display the neighbor information of the aggregation group	show lacp neighbor [***ch-id***]	optional

3.2.4 Configure the Load Balancing Policy

After the aggregation group takes effect, it will forward the service flow among LACP members according to certain policies. The default load balancing uses the src-mac and it can be modified. There is no command to view the load policy separately. You can find the configuration information via *show lacp internal [ch-id].*

Configure the load balancing policy

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the load balancing policy (all groups)	channel-group load-balance {src-mac | dst-mac | src-dst-mac | src-dst-ip | src-ip | dst-ip }	src-mac by default
Restore the default load policy	no channel-group load-balance	optional

3.2.5 Configure the System Priority

In dynamic LACP mode, the master switch and slave switch are selected according to the system ID. System ID is determined by the system priority and the local MAC address. When select the master switch and slave switch, it compares the priority firstly, the smaller value wins; the same priority, compared MAC, the smaller MAC value wins.

The default system priority is 32768.

Configure the System Priority

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the priority of the system	lacp system-priority ***value***	optional
Restore the default configuration	no lacp system-priority	optional
Display the system priority configuration	show lacp sys-id	optional

3.2.6 Configure the Port Priority

When a dynamic LACP is running, the master switch selects the logical port according to the port ID which consists of port priority and port number. The logical port in Lacp is used to forward protocol packets, such as stp.

When the master switch selects the logic port: Compare the priority firstly, the smaller value wins; if the priority is the same, then compare the port number, the smaller value wins. By default, ports priorities are the same, and the value is equal to 128. However, the port priority value can be modified. Moreover, the priority needs to be a multiple of 16, if the input value N is not an integer multiple of 16, it will automatically use the result of N divisible by 16. For example,

if you set the port priority as 17, the actual success is issued 17 \ 16 = 1.

The logical port on the switch does not need to be selected, and directly uses the LACP member port which connected to the master switch's logical port.

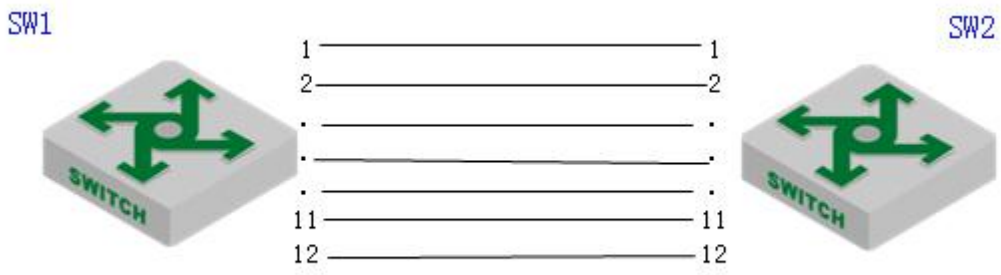
Configure the Port Priority

Operation	Command	Remarks
Enter interface configuration mode	interface ethernet ***port-num***	-
Configure interface priority	lacp port-priority ***value***	optional
Restore default priority	no lacp port-priority	optional

3.2.7 LACP Configuration Example

1. Network requirements

As shown below, the 1-12 ports of SW1 and SW2 respectively run dynamic LACP.



sketch map of LACP

2. Configuration steps

SW1 CONFIGURATION:
SW1(config)#channel-group 10
SW1(config)#interface range ethernet 0/0/1 to ethernet 0/0/12
SW1(config-if-range)#channel-group 10 mode active

#SW2 CONFIGURATION
SW2(config)#interface range ethernet 0/0/1 to ethernet 0/0/12
SW2(config-if-range)#channel-group 2 mode active

3. Result validation

After the dynamic LACP negotiation succeeds, the information is displayed
SW1(config)#show lacp internal
load balance: src-mac
Channel: 10, dynamic channel

Port	State	A-Key	O-Key	Priority	Logic-port	Actor-state
e0/0/1	bndl	11	11	128	1	10111100
e0/0/2	bndl	11	11	128	1	10111100
e0/0/3	bndl	11	11	128	1	10111100
e0/0/4	bndl	11	11	128	1	10111100
e0/0/5	bndl	11	11	128	1	10111100
e0/0/6	bndl	11	11	128	1	10111100
e0/0/7	bndl	11	11	128	1	10111100

e0/0/8	bndl	11	11	128	1	10111100
e0/0/9	standby	11	11	128	9	10110000
e0/0/10	standby	11	11	128	10	10110000
e0/0/11	standby	11	11	128	11	10110000
e0/0/12	standby	11	11	128	12	10110000

actor-state: activity/timeout/aggregation/synchronization
collecting/distributing/defaulted/expired

3.3 Port Isolation Configuration

3.3.1 Port Isolation

Through the port isolation feature, you can add the ports that need to be controlled into an isolation group, which can not only improve the security of the network by separating Layer 2 and Layer 3 among the ports in the isolation group according to the type of the isolation group, but also provide users with a flexible network solution.

If the port isolation group is configured, the downstream ports cannot communicate with each other while the upstream and downstream ports can communicate with each other.

Configure port isolation based on global configuration mode

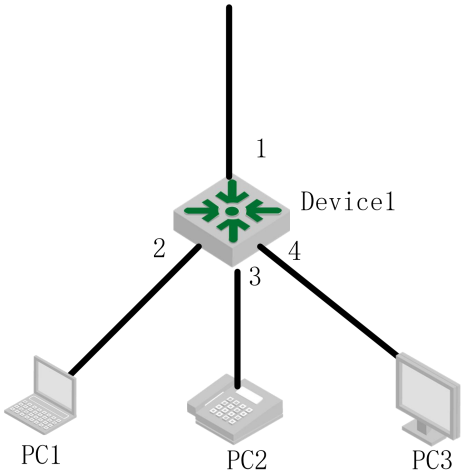
Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure port isolation	**port-isolation ethernet** *device/slot/port*	required
Delete port isolation	**no port-isolation {all | ethernet** *device/slot/port***}**	optional
Display the isolation port	**show port-isolation**	Any modes

In this configuration, the ports that are not configured as isolated ports are the uplink ports.

3.3.2 Configuration Example for Port Isolation

1. Network requirements

PC1, PC2 and PC3 are connected to ports 2, 3, and 4 of the switch. The switch is connected to the external network through port 1. PC1, PC2, and PC3 need to be isolated between Layer 2 and Layer 3. The networking diagram is as follows:



sketch map of port isolation

2. Configuration steps:

Configure ports 2, 3, and 4 as downlink ports and port 1 as uplink ports

Switch(config)#port-isolation ethernet 0/0/2 to e 0/0/4

Add port isolation downlink port successfully.

3. Result validation

Switch(config)#show port-isolation

Port isolation downlink port :

e0/0/2-e0/0/4.

At the moment, e0/0/2, e0/0/3 and e0/0/4 can not communicate with each other. They can only communicate with e0/0/1.

3.4 Loopback

3.4.1 Loopback Overview

A loopback test allows you to send and receive data from the same serial port to verify that the port is operational. To perform this test, you need to temporarily connect the proper pins to allow signals to be sent and received on the same port. During the test, the port cannot transmit the packets. Only when this test is finished can the packets be transmitted normally.

3.4.2 Configure Loopback

Configure Loopback

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Inner loop test	loopback internal	required
Outer loop test	loopback external	required
Enter port configuration mode	interface ethernet ***port-number***	optional
Loopback internal	loopback internal	optional
Loopback external	loopback external	optional

3.4.3 Loopback Configuration Example

1. Network requirements
No

2. Configuration steps
No

3.Result validation
(1) inner loop test
Switch(config)#interface interface range ethernet 0/0/1 to ethernet 0/0/2
Switch(config-if-range)#loopback internal
Port ethernet e0/0/1 internal looptest successfully
Port ethernet e0/0/2 internal looptest successfully

(2) outer loop test

Switch (config-if-range)#loopback external
Port ethernet e2/0/5 external looptest successfully
Port ethernet e2/0/6 external looptest successfully

3.5 VCT - Virtual Cable Test

3.5.1 VCT Overview

VCT is used to test whether the network is normal or open or short or impedance mismatch and so forth.

3.5.2 Configure VCT

The command of vct run should be performed under global configuration mode, and it is applied to all ports; if you run the command of vct run under the interface configuration mode, it can only test that interface only. It can find out the errors and the location of the errors.

System also supports VCT autotest. When the VCT is enabled, it will automatically perform the VCT if there are the ports in the link down state.

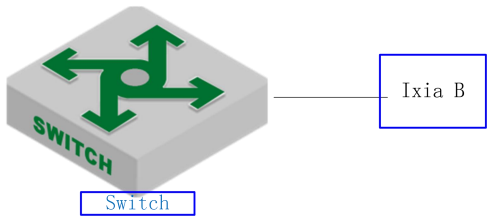
Configure VCT

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
vct by manually	vct run	required
vct by automatically	vct auto-run	optional
Display vct configuration	show vct auto-run	optional
Enter port configuration mode	interface ethernet ***port-number***	-

vct by manually	vct run	optional
vct by automatically	vct auto-run	optional

3.5.3 VCT Configuration Example

1. Network requirements
Connect the tester with the switch.



VCT schematic diagram

2. Configuration steps
No

3.Result validation
(1) Perform the vct on the port which is connected to cables:
Switch(config)#interface ethernet 0/0/1
Port ethernet 0/0/1 VCT result :

		TX pair	RX pair
status	:	NORMAL	NORMAL
locate	:	_	_

(2) Perform the vct on the port which is not connected to cables:
Switch(config-if-ethernet-0/0/1)#interface ethernet 0/0/12
Switch(config-if-ethernet-0/0/12)#vct run
Port ethernet 0/0/12 VCT result :

		TX pair	RX pair
status	:	OPEN	OPEN
locate	:	7	6

3.6 DDM- Digital Diagnostic Monitor

3.6.1 DDM Overview

DDM (Digital Diagnostic Monitor) is used to test SFP parameter, for example: Temperature, Vcc, Tx Bias Current, Tx Power, Rx Power.

3.6.2 Show DDM Test Information

Display DDM Test Information

Operation	Command	Remarks
Enter port configuration mode	configure terminal	-
Show DDM Test Information	show interface sfp [ethernet ***port-number***]	required

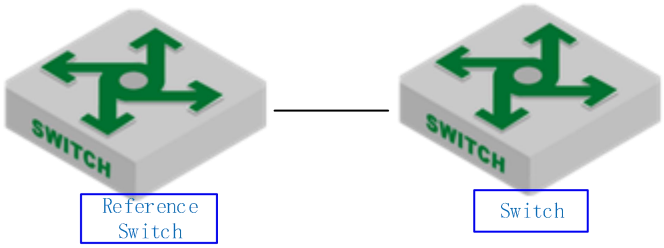
Note:

Insert the optical module, you can perform DDM detection

3.6.3 DDM Configuration Example

1.Network requirements

Switch and the reference device are connected with the SFP which supports DDM test to check the information of DDM test.



sketch map of DDM test

2.Configuration steps

No

3.Result validation

Display DDM test information on the switch

Switch(config)#show interface sfp ethernet 0/1/1

Port e0/1/1 :

Common information:

Transceiver Type :SFP

Compliance 10G BASE-LR

Connector Type :LC

WaveLength(nm) :1310

Transfer Distance(m) :10000(9um)

Digital Diagnostic Monitoring :YES

VendorName :WTD

Manufacture information:

Manu. Serial Number :BP132500260047

Manufacturing Date :2013-06-19

VendorName :WTD

Diagnostic information:

Temperature(℃) :28

Voltage(V) :3.3098

Bias Current(mA) :35.419

Bias High Threshold(mA) :70.00

Bias Low Threshold(mA) :15.00
RX Power(dBm) :-2.80
RX Power High Threshold(dBm) :0.00
RX Power Low Threshold(dBm) :-15.20
TX Power(dBm) :-3.10
TX Power High Threshold(dBm) :0.00
TX Power Low Threshold(dBm) :-8.20

3.7 Port-Statistic

The port will calculate the status of receive packet and transmit packet to make it easier for administrator to analyze failure causes.

3.7.1 Ordinary Port Packet- statistic

Generally speaking, statistic information includes the rate of receiving packet and transmitting packet, the errors of receive packet and transmit packet, classified statistic according to the byte, multicast& unicast& broadcast, packet loss statistic.

Please refer to the instance example for the detailed information.

Ordinary Port- statistic

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Display port-statistic information	show statistic interface ethernet [***port-number***]	required
Clear port-statistic	clear interface ethernet [***port-number***]	optional
Display interface real-time statistic information	show statistics dynamic interface	required
Display the port utilization	show utilization interface	required
Display the interface information	show interface [ethernet ***port-number***]	optional

3.7.2 CPU Port- statistic

This function is used to calculate the packet counts transmitted to cpu.

Ordinary Port- statistic

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Display cpu port-statistic information	show cpu-statistic [ethernet ***port-number***]	required
Clear cpu port-statistic	clear cpu-statistics	optional
Display cpu classified statistic information	show cpu-classification [interface ethernet ***port-number***]	optional
Clear cpu classified statistic information	clear cpu-classification [interface ethernet ***port-number***]	optional
Display the port utilization	show utilization interface	optional

Display cpu vacancy rate	show cpu-utilization	optional

3.7.3 5-minutes Port Rate Statistic

This function is used to calculate the average rate of receiving packet and transmitting packet in specified time. The default statistical cycle and largest statistical period are five minutes.

5-miniutes port rate statistic

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the port- rate statistic interval	[no] port-rate-statistics interval ***value***	optional 5 minutes by default “No command” is used to restore the default settings
Display port- rate statistic information	show statistics interface [ethernet ***port-number***]	optional

3.7.4 Port Statistic of Aggregation Group

This function is used to calculate the receive packet and transmit packet of the aggregation group.

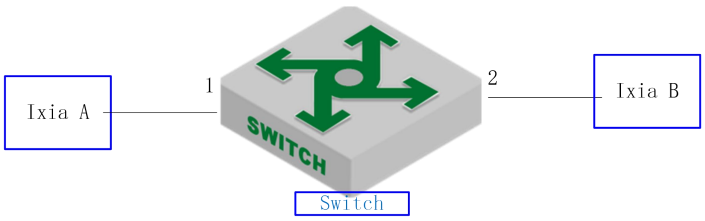
Port Statistic of Aggregation Group

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Display LACP statistic information	show statistics channel-group [***channel-id***]	required
Clear LACP statistic information	clear channel-group [***channel-id***]	required
Display LACP real-time statistic information	show statistics dynamic interface channel-group	required

3.7.5 Configuration Example of Interface Statistic

1. Network requirements

Tester A forwards packets to DUT with line speed to check interface statistic information.



sketch map of interface statistic

2.Configuration steps

No

3.Result validation

Display interface statistic information

Switch(config)#show statistics interface ethernet 0/0/1

Port number : e0/0/1

last 5 minutes input rate 6198600 bits/sec, 12106 packets/sec

last 5 minutes output rate 28256 bits/sec, 55 packets/sec

64 byte packets:4267810

65-127 byte packets:0

128-255 byte packets:0

256-511 byte packets:0

512-1023 byte packets:0

1024-1518 byte packets:0

4267707 packets input, 273132992 bytes , 1 discarded packets

4267707 unicasts, 0 multicasts, 0 broadcasts

1 input errors, 0 FCS error, 0 symbol error, 0 false carrier

1 runts, 0 giants

23763 packets output, 1520832 bytes, 0 discarded packets

0 unicasts, 23763 multicasts, 0 broadcasts

0 output errors, 0 deferred, 0 collisions

0 late collisions

Total entries: 1.

Switch(config)#show interface ethernet 0/0/1

Fast Ethernet e0/0/1 current state: enabled, port link is up

Time duration of linkup is 31 second

Hardware address is 00:0a:5a:00:04:1e

SetSpeed is auto, ActualSpeed is 100M, Duplex mode is full

Current port type: 100BASE-T

Priority is 0

Flow control is disabled

Broadcast storm control target rate is 50000pps

PVID is 1

Port mode: hybrid

Untagged VLAN ID : 1

Input : 5361414 packets, 343130240 bytes

0 broadcasts, 0 multicasts, 5361414 unicasts

Output : 23763 packets, 1520832 bytes

0 broadcasts, 23763 multicasts, 0 unicasts

3.8 Flow Control

3.8.1 Overview for Flow Control

When one-side switch and the other-side switch enable the flow control function, if the one-side switch is congested:

- The one-side switch sends a message to other-side switch to notify the other-side switch to stop sending packets or slow down the sending speed.
- After receiving the message, the other-side switch stops sending packets to the one-side or slows down the sending speed. This prevents packets from being lost and ensures normal operation of the network services.

3.8.2 Configure Flow Control

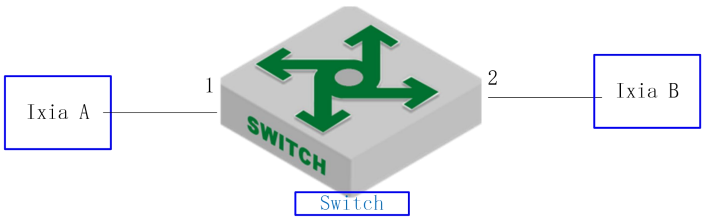
Configure Flow Control

Operation	Command	Remarks
Enter port configuration mode	configure terminal	-
Enable/disable flow control function	[no]flow-control	required; Disabled by default.
Display the flow control configuration information	show flow-control interface [ethernet ***port-number***]	optional

3.8.3 Configuration Example for Flow Control

1.Network requirements

Tester A, switch port1 and port2 enable flow control, port2 with export bandwidth 1M, tester A sends packets to ixia B, and then you can check whether the DUT issued flow control frame.



sketch map of flow control

2.Configuration steps

configure the flow control of switch port1 and port2
Switch(config)#interface range ethernet 0/0/1 ethernet 0/0/2
Switch(config-if-range)#flow-control
Switch(config-if-range)#exit

configure port2 with export bandwidth 1M
Switch(config)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#bandwidth egress 1024

3.Result validation
(1) Tester A enables the flow control, wire-speed forwarding; ------ Tester A received flow control frame issued by DUT, and the packet rate automatically adjusted to 1M.

4. VLAN Configuration

4.1 VLAN Overview

Virtual Local Area Network (VLAN) is the technology which realizing virtual work group through segmenting the LAN devices into every network segment logically but not segmenting the LAN devices into every network segment physically. IEEE issued the IEEE 802.1Q in 1999, which was intended to standardize VLAN implementation solutions.

Network managers can logically segment the physical LAN into different broadcast domains via VLAN technology. Each VLAN contains a group of computer workstation with the same demands. The workstations of a VLAN do not have to belong to the same physical LAN segment. With VLAN technology, the broadcast and unicast traffic within a VLAN will not be forwarded to other VLANs. Therefore, it is very helpful in traffic controlling, saving device investment, simplifying network management and improving security. The following are VLAN features:

Compared with the traditional Ethernet, VLAN enjoys the following advantages.

helpful in traffic controlling

In traditional network, mass broadcast data will be sent to all network devices directly regardless of whether it is necessary or not, leading to network jitter consequently. However, VLAN supports to configure the necessary communication device in each VLAN so as to reduce broadcast and then improve network efficiency.

providing higher security

Device can only communicate with another device under the condition that both of them belongs to the same VLAN. For example, it must be under the help of router device if the VLAN device of Research and Development Department needs to connect with the VLAN device of Product Department. In this way, these two departments cannot communicate directly so as to improve system security function.

reducing network configuration workload

VLAN can be used to group specific hosts. When the physical position of a host changes within the range of the VLAN, you need not change its network configuration.

4.1.1 Vlan Configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Create/ delete vlan	(no) vlan *vlan-list*	
Add vlan interface	switchport ethernet ***port-number***	
Specify vlan description	description *string*	

4.1.2 Interface Default vlan ID

Interface default vlan is also called pvid. When receiving a untagged packet, system will add a tag to the packet in which the VLAN ID is the default VLAN ID.

Configure interface default vlan ID

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure interface pvid	switchport default vlan ***vlan-id***	optional
Restore default pvid	no switchport default vlan	pvid=1 by default.
Display interface detailed configuration	show interface ethernet ***port-number***	optional
Display interface brief configuration	show interface brief ethernet [***port-number***]	optional

4.1.3 Interface Type

Interface type can be divided into three types according to the different process modes the interface performs on tag label:

Access interface: the interface only belongs to one vlan, and it usually is used to connect the terminal device.

Trunk: the interface can be able to receive and forward multiple vlans. When the message is forwarded, the default vlan message will not carry the tag whereas the other vlan will carry the tag, and the tag is applied to the switch interface.

Hybrid interface: the interface can be able to receive and forward multiple vlans, and it allows multiple vlans to carry the tag or not carry the tag.

interface type	Processing on receiving message		Processing on forwarding message
	Untag	Tag	

Access	Receive it and add a tag with VID being equal to PVID.	If VID of the packet is equal to the port permitted VID, receive it; if VID is different, discard it.	Strip the Tag and transmit the packet as the VID of the packet is equal to the port permitted VID
Hybrid			If VID of the packet is equal to the port permitted untag VID, remove the tag and transmit it; If VID of the packet is equal to the port permitted tag VID, keep the tag and transmit it.
Trunk			If VID of the packet is equal to the port permitted VID, keep the tag and transmit it.

Configure interface vlan mode

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure interface vlan mode	switchport mode **{ access | hybrid | trunk }**	Optional; Hybrid by default.

4.1.4 VLAN Attributes Based on Hybrid Interface

Enter port configuration mode	interface ethernet ***port-number***	-
Configure interface vlan mode	switchport mode **hybrid**	Optional; Hybrid by default
Allow the specified vlan to pass this hybrid port	switchport hybrid {tagged | untagged} vlan { *vlan-list* | *all* }	"tagged attribute" means that the vlan packet carries tag; "untagged attribute" means that the vlan packet does not carry tag;
Does not allow the specified vlan to pass this hybrid port	no switchport hybrid vlan *vlan-list*	

4.1.5 VLAN Attributes Based on Trunk Interface

Enter port configuration mode	interface ethernet ***port-number***	-
Configure interface vlan mode	switchport mode trunk	Optional; Hybrid by default;
Allow the specified vlan to pass this trunk port	switchport trunk allowed vlan { *vlan-list* | *all* }	
Do not allow the specified vlan to pass this trunk port	no switchport trunk allowed vlan { *vlan-list* | *all* }	

4.1.6 Configure port priority

When switch receives a untagged packet, system will add a vlan tag to the packet in which the vid value in the tag is the PVID value and the priority value is the interface priority value.

Configure the interface priority

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure port priority	priority ***value***	optional
Restore default priority	no priority	optional 0 by default
Display interface detailed configuration	show interface ethernet ***port-number***	optional
Display interface brief configuration	show interface brief ethernet [***port-number***]	optional

4.1.7 Ingress Filtering

By default, interface will check whether the receiving packet belongs to the vlan member, if it is, the interface will perform the forward processing or it will discard the packet. This process is called ingress filtering. Switch will enable this function by default and this function is allowed to be disabled.

Ingress Filtering

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure ingress filtering	[no] ingress filtering	optional Enabled by default
Display the configuration information	show ingress [interface ***port-number***]	optional

4.1.8 Configure Types of Interface acceptable-frame

By default, no matter tag packet or untag packet the switch receives, it allows modifying the packets to be the type that the interface can be received.

Configure Types of Interface acceptable-frame

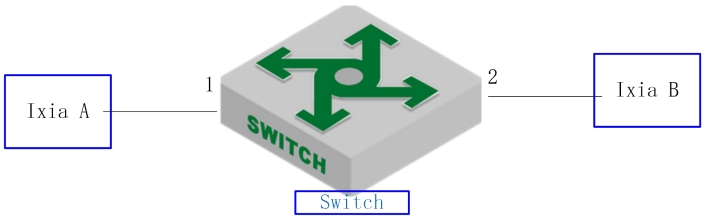
Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure interface priority	ingress acceptable-frame { **all** | **tagged** }	“all” means it can receive the tag packets and untag packets; “tagged” means it can only receive the tag packets.
Display the configuration	show ingress [interface ***port-number***]	optional

4.1.9 Configuration Example

Example 1

1.Network requirements

Create vlan 100, including member 1and 2, 1 is access port and 2 is trunk port.



sketch map of interface default vlan

2.Configuration steps

create vlan 100, and then add member 1 and member 2；

Switch(config)#vlan 100

Switch(config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/2

modify the vlan mode of port 1 and port 2, and then configure the pvid.

Switch(config)#interface ethernet 0/0/1

Switch(config-if-ethernet-0/0/1)#switchport mode access

Switch(config-if-ethernet-0/0/1)#switchport default vlan 100

Switch(config-if-ethernet-0/0/1)#interface ethernet 0/0/2

Switch(config-if-ethernet-0/0/2)#switchport mode trunk

Switch(config-if-ethernet-0/0/2)#switchport default vlan 100

Switch(config-if-ethernet-0/0/2)#exit

3.Result validation

display the information of port 1 and port 2

Switch(config)#show interface brief ethernet 0/0/1 ethernet 0/0/2

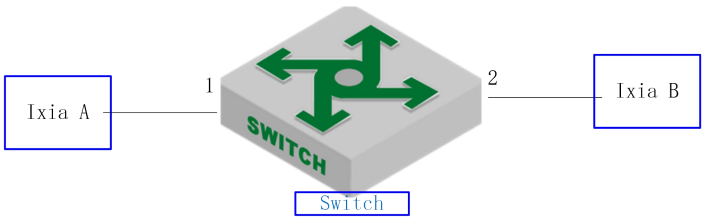
Port	Desc	Link	shutdn	Speed	Pri	PVID	Mode	TagVlan	UtVlan
e0/0/1		up	false	auto-f100	0	100	acc		100
e0/0/2		up	false	auto-f100	0	100	trk		100

Total entries: 2 .

Example 2

1.Network requirements

Configure port 1 to be access mode; configure port 2 to be trunk mode.



sketch map of interface vlan mode

2.Configuration steps

configure port 1 to be access mode;

Switch(config)#interface ethernet 0/0/1

Switch(config-if-ethernet-0/0/1)#switchport mode access

configure port 2 to be trunk mode;

Switch(config-if-ethernet-0/0/1)#interface ethernet 0/0/2

Switch(config-if-ethernet-0/0/2)#switchport mode trunk

Switch(config-if-ethernet-0/0/2)#exit

3.Result validation

display the information of port 1 and port 2:

Switch(config)#show interface brief ethernet 0/0/1 ethernet 0/0/2

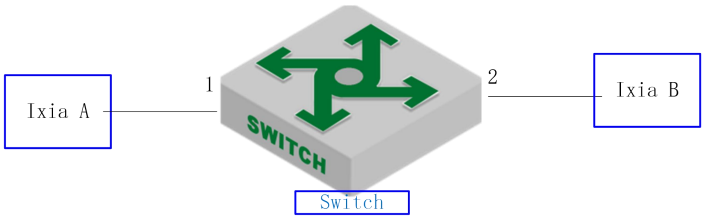
Port	Desc	Link	shutdn	Speed	Pri	PVID	Mode	TagVlan	UtVlan
e0/0/1		up	false	auto-f100	0	1	acc		1
e0/0/2		up	false	auto-f100	0	1	trk		1

Total entries: 2 .

Example 3

1.Network requirements

Create vlan 500, including member 1 and member 2; port 1 and port 2 are hybrid; configure the vlan 500 with tag in egress.



Tag attributes in hybrid egress

2.Configuration steps

configure vlan 500 and add member 1 and member 2;

Switch(config)#vlan 500

Switch(config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/2

Switch(config-if-vlan)#show vlan 500

show VLAN information
VLAN ID : 500
VLAN status : static
VLAN member : e0/0/1-e0/0/2.
Static tagged ports :
Static untagged Ports : e0/0/1-e0/0/2.
Dynamic tagged ports :
Total entries: 1 vlan.#

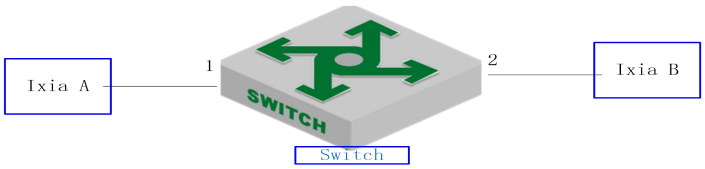
configure vlan 100 with tag in egress of port 1 and port 2;
Switch(config-if-vlan)#interface range ethernet 0/0/1 ethernet 0/0/2
Switch(config-if-range)#switchport hybrid tagged vlan 500
Switch(config-if-range)#show vlan 500
show VLAN information
VLAN ID : 500
VLAN status : static
VLAN member : e0/0/1-e0/0/2.
Static tagged ports : e0/0/1-e0/0/2.
Static untagged Ports :
Dynamic tagged ports :
Total entries: 1 vlan.

3.Result validation
(1) tester A forwards the unknown packet of vlan =500, ixia B can be able to receive the packet of vlan =500 with tag.

Example 4

1.Network requirements

Create vlan 100 and then add member 1 and member 2 ; create vlan 200 and then add member 1 and member 2.



sketch map of adding the port to vlan

2.Configuration steps

create vlan 100 and then add member 1and member 2
Switch(config)#vlan 100
Switch(config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/2
Switch(config-if-vlan)#show vlan 100
show VLAN information
VLAN ID : 100
VLAN status : static
VLAN member : e0/0/1-e0/0/2.
Static tagged ports :

Static untagged Ports : e0/0/1-e0/0/2.
Dynamic tagged ports :

create vlan 200 and then add member 1and member 2
Switch(config)#vlan 200
Switch(config-if-vlan)#exit
Switch(config-if-range)#interface range ethernet 0/0/1 ethernet 0/0/2
Switch(config-if-range)#switchport hybrid untagged vlan 200
Switch(config-if-range)#show vlan 200
show VLAN information
VLAN ID : 200
VLAN status : static
VLAN member : e0/0/1-e0/0/2.
Static tagged ports :
Static untagged Ports : e0/0/1-e0/0/2.
Dynamic tagged ports :
Total entries: 1 vlan.

4.2 Management vlan

4.2.1 Management vlan Overview

Management vlan is used to manage the equipment.

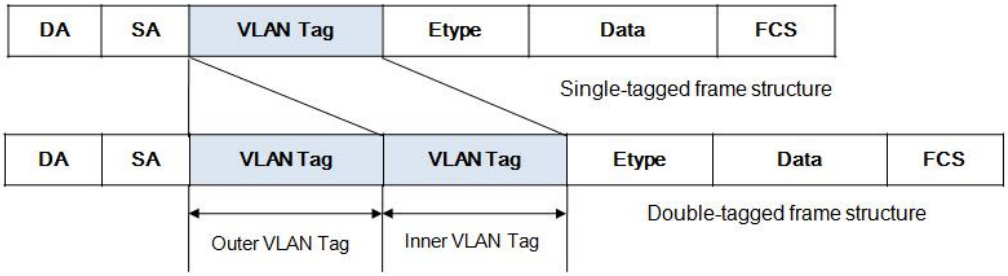
In layer -2 switch, only the user who belongs to management vlan can he login the device via telnet, ssh, web and snmp. That is, only users belonging to the management VLAN can communicate with the switch CPU. Therefore, the layer-2 switch should configure management vlan.

In layer -3 switch, there is no need to configure management vlan for its interface vlan is the management vlan.

4.3 QinQ Configuration

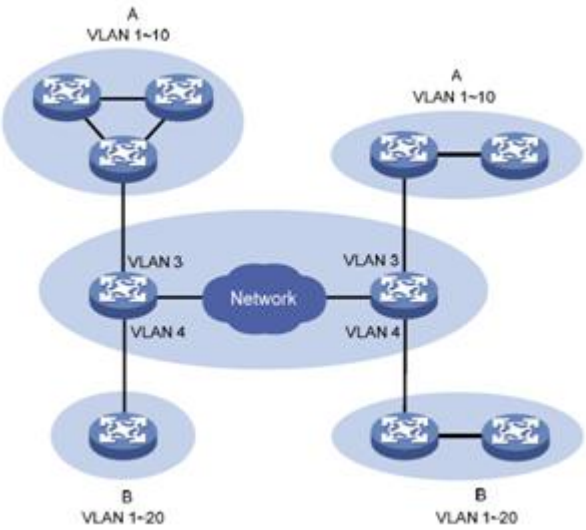
4.3.1 QINQ Overview

In the VLAN tag field defined in IEEE 802.1Q, only 12 bits are used for VLAN IDs, so a switch can support a maximum of 4,094 VLANs. In actual applications, however, a large number of VLANs are required to isolate users, especially in metropolitan area networks (MANs), and 4,094 VLANs are far from satisfying such requirements. The generation of QinQ technology solves the insufficient of VLAN users. The following shows the structure of 802.1Q-tagged and double-tagged Ethernet frames. The QinQ feature enables a device to support up to 4,094 x 4,094 VLANs to satisfy the requirement for the amount of VLANs.



QinQ Ethernet frame structure

The port QinQ feature is a flexible, easy-to-implement Layer 2 VPN technique, which enables the access point to encapsulate an outer VLAN tag in Ethernet frames from customer networks (private networks), so that the Ethernet frames will travel across the service provider's backbone network (public network) with double VLAN tags. The inner VLAN tag is the customer network VLAN tag while the outer one is the VLAN tag assigned by the service provider to the customer. In the public network, frames are forwarded based on the outer VLAN tag only, with the source MAC address learned as a MAC address table entry for the VLAN indicated by the outer tag, while the customer network VLAN tag is transmitted as part of the data in the frames.



QinQ application

A VLAN tag uses the tag protocol identifier (TPID) field to identify the protocol type of the tag. The value of this field, as defined in IEEE 802.1Q, is 0x8100. The device can identify whether there is corresponded VLAN Tag according to TPID. If configured TPID is the same as the corresponded field, packet is regarded as with VLAN Tag.

The TPID in an Ethernet frame has the same position with the protocol type field in a frame without a VLAN tag. To avoid problems in packet forwarding and handling in the network, you cannot set the TPID value to any of the values in the table below.

common protocol type values

Protocol type	value
ARP	0x0806
PUP	0x0200
RARP	0x8035
IP	0x0800
IPv6	0x86DD

PPPoE	0x8863/0x8864
MPLS	0x8847/0x8848
IPX/SPX	0x8137
IS-IS	0x8000
LACP	0x8809
802.1x	0x888E
GnLink	0x0765
GSTP	0X5524

In QinQ, there are four interface mode: none, customer, customer-no-static and uplink. After enabling QinQ, interface mode will be **none** mode by default.

Four interface modes are suitable for different applications:

1. None, this mode is suitable for device to connect PC directly, and the packet of this interface will try to identify outer VLAN tag and inner VLAN tag; the inner VLAN tag will be stripped before forwarding while whether the outer VLAN tag will be stripped or not is determined by the outer VLAN configuration.

2. Customer, the interface of this mode and the interface of the uplink mode form the static QinQ application. The packet which enter this interface will be considered not include outer VLAN tag and inner VLAN tag, so the packet will take the PVID or the other specified vlan ID as the packet outer VLAN ID; the outer VLAN tag will be stripped before forwarding regardless the outer VLAN configuration.
   Customer interface needs to connect with the interface of user-device. The single VLAN tag enters from this interface will carry 2 VLAN tags when it is forwarded from the uplink interface.

3. customer-no-static, the interface of this mode and the interface of the uplink mode will form the dynamic QinQ application. The packet which enters this interface will try to identify outer VLAN tag, and then match the dynamic QinQ rule according to the VLAN ID of outer VLAN tag. If the dynamic QinQ rule is matched successfully, the packet will be considered not include outer VLAN tag and inner VLAN tag, what is more, it will take the specified VLAN ID as the outer VLAN ID; the outer VLAN tag will be stripped before forwarding regardless the outer VLAN configuration.

4. uplink, the interface of this mode needs to connect with the interface of operator-device. The packet which enters this interface will try to identify outer VLAN tag, and the packet will carry the inner VLAN tag before being forwarded, moreover, whether the outer VLAN tag will be stripped or not is determined by the outer VLAN configuration. The uplink interface judges whether the packet carries tag or not according to the vlan protocol number and the outer-tpid value. If the vlan protocol number is the same as the outer-tpid value, it will be taken as carrying with tag; if the vlan protocol number is the different from the outer-tpid value, it will be taken as not carrying with tag.

📖 Note:

The customer mode interface cannot communicate with none mode interface.

4.3.2 Static QINQ Overview

After enabling dtag function, device will not run static QinQ by default. Only when you configure customer under interface mode can the static QinQ take effect. Under the condition that the interface does not configure insert rule, the packet will be processed according to static QinQ. The following is static QinQ handling process:

1. Customer port process flow
   Customer port will add vlan tag what is the same as interface pvid to the original packet and forward the packet in this vlan regardless whether the packet carries tag or not, whether the tpid is the same as configured outer-tpid/inner-tpid or not, whether the vid exists in the equipment or not.

2. Uplink port process flow
   Whether the uplink port will insert a tag or not is decided by whether the packet carries a tag or not. Whether the uplink port carries tag or not is decided by whether the vlan protocol number is the same as global outer-tpid value or not:
   a) If the vlan protocol number is the same as global outer-tpid value, it will be taken as carrying with the tag and then it will forward the packet in the carried vlan without inserting a tag; if the carried vlan does not exists in the system, the packet will not be forwarded；
   b) If the vlan protocol number is different from global outer-tpid value, it will insert a vlan tag according to the pvid firstly, and then add an outer-tpid according to the inner-tpid before forwarding；

3. TPID value processing
   When the packet forwarded from customer port or uplink port, TPID value is always consistent with the outer-tpid value.

Static QinQ Configuration

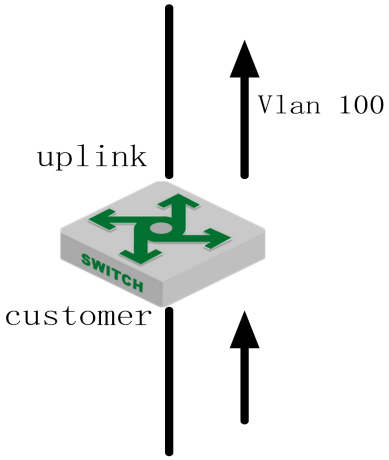
Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable(disable) global QinQ	**[no]dtag**	required
Enter port configuration mode	**interface ethernet** *device/slot/port*	-
Specify the port QinQ mode	**dtag mode {customer|uplink}**	required，customer and uplink interface form static qinq
Restore port default mode	**no dtag mode**	Optional none mode by default
Display QinQ configuration	**show dtag**	any mode

4.3.3 Configuration Example for Static QinQ

1.Network requirements:

It requires packet add outer vlan 100 tag when forwarding from uplink interface regardless whether the packet carries VLAN tag or not and what the tag it will be.

Networking is as follows:



2.Configuration steps:
enable global dtag
Switch(config)#dtag
enter port 1, configure the QINQ port mode to be customer
Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#dtag mode customer
configure the default vlan of port 1 to be 100
Switch(config-if-ethernet-0/0/1)#switchport default vlan 100

enter port 2, configure the QINQ port mode to be uplink
Switch(config)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#dtag mode uplink
configure port 2 to be trunk port
Switch(config-if-ethernet-0/0/2)#switchport mode trunk

3.Result validation:
No matter whether the receiving packet of customer interface carries the vlan tag or not as well as what the tag value it will be, the packet will add the tag head of vlan 100 when transmit from uplink port.

4.3.4 Dynamic QINQ Overview

Uplink port can be able to configure insert rule. If the packet tag conforms to insert rule, it will be dealt according to the corresponding rule regardless what the configured outer-tpid value it will be.

Customer-no-static port will not only check whether the packet carries tag or not but also

judge whether it conforms to the insert rule. If the packet carries tag and it conforms to the insert rule, it will be processed according to the dynamic QINQ, or it will be processed according to 802.1Q rule. The following is the detailed process:

1. Compared the ingress packet with the configured outer-tpid firstly. If the packet tpid is different from the configured outer-tpid, the packet will be processed according to 802.1Q rule, adding a tag according to the interface PVID；
2. If packet's tpid is the same as outer-tpid but the packet VID is not in the insert range, the packet will be processed according to 802.1Q rule, performing unvarnished transmission or discard；
3. If packet's tpid is the same as outer-tpid and the packet VID is in the insert range, the packet will be processed according to the dynamic QINQ:
   a) The packet will be processed according to insert rule regardless whether the packet VID exists or not, adding an outer tag according to the insert rule；
   b) If the outer vlan which corresponding to the insert rule does not exist, or the vlan exists but not includes the ingress, the packet will not perform any processing but discard it；
   c) If the outer vlan which corresponding to the insert rule do exist but the vlan not includes any ingress port, the packet will be processed according to insert rule firstly, that is, this packet will learn the MAC address according to the outer vlan, but this packet will be discarded later.

Dynamic QINQ Configuration

Operation	Command	Operation
Enter global configuration mode	**configure terminal**	-
Enable(disable) global QinQ	**[no]dtag**	required
Enter port configuration mode	**interface ethernet** *device/slot/port*	-
QinQ mode of specified port	**dtag mode {customer |uplink}**	required
Restore the QinQ default mode of the interface	**no dtag mode**	Optional None mode by default
Configure insert rule	**dtag insert** *start-vlan end-vlan service-vlan*	required
Delete insert rule	**no dtag insert {all |** *start-vlan end-vlan* **}**	optional
Display QinQ configuration	**show dtag**	any mode

Note:

Static QinQ and dynamic QinQ can be able to enable at the same time. However, when configuring the dynamic QinQ rule, the dynamic QinQ rule will be determined according to the outer VLAN ID of static QinQ. Therefore, dynamic QinQ interface mode is recommended to be configured as customer-no-static.

4.3.5 Configuration Example for Dynamic QINQ

1. Network requirements

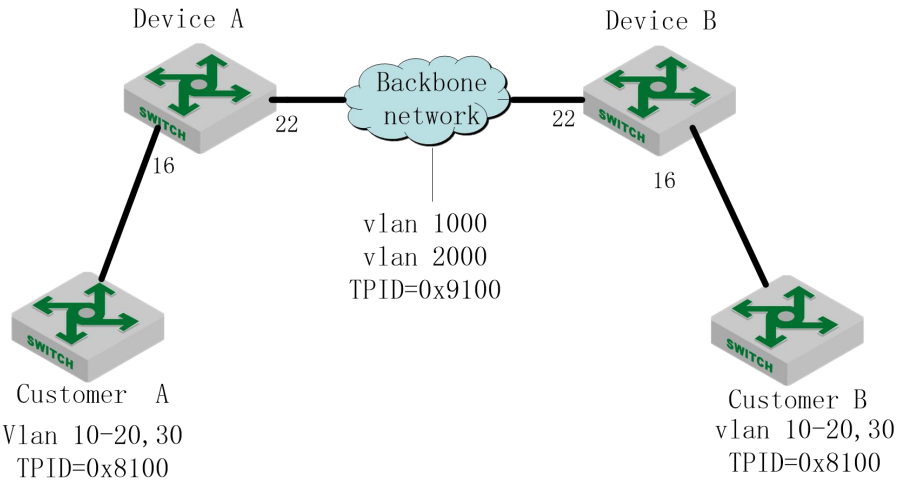
Device A and device B act as the operator s network to access the device. Customer A and customer B act as user ends network to access the device. Device A connects with device B via trunk port, allow vlan 1000 and vlan 2000 to pass through; use the equipment of other equipment manufacturers among device A and device B, TPID=0x9100;

It requires implementing the following requirements:

a) After forwarded by the operator's network vlan1000, Customer A vlan10-20 packet can be able to communicate with customer B vlan 10-20 packet;

b) After forwarded by the operator's network vlan2000, Customer A vlan30 packet can be able to communicate with customer B vlan 30 packet;

Network diagram is as follows:



2.Configuration steps

Device A configuration:

create vlan 1000 and vlan 2000， and then add them to user-port (port 16) and service-port (port 22)

Switch(config)#vlan 1000,2000

Switch(config-if-vlan)#switchport ethernet 0/0/16 ethernet 0/0/22

configure the service-port to be trunk port

Switch(config)#interface ethernet 0/0/22

Switch(config-if-ethernet-0/0/22)#switchport mode trunk

enable global qinq function

Switch(config)#dtag

Configure outer tag tpid to be 0x9100

Switch(config)#dtag outer-tpid 9100

configure user-port qinq mode

Switch(config)#interface ethernet 0/0/16

Switch(config-if-ethernet-0/0/16)# dtag mode customer

configure service-port qinq mode

Switch(config)#interface ethernet 0/0/22

Switch(config-if-ethernet-0/0/22)#dtag mode uplink

configure the user-port insert rule
Switch(config)#interface ethernet 0/0/16
Switch(config-if-ethernet-0/0/16)#dtag insert 10 20 1000
Set SVLAN successfully.
Switch(config-if-ethernet-0/0/16)#dtag insert 30 30 2000
Set SVLAN successfully.

The configuration of Device B is the same as Device A, so no repeat here.

3.Result validation
After Customer A vlan10-20 packet passed through device A, the packet will carry the tag of vlan=1000, tpid=0x9100 to forward.
After Customer A vlan30 packet passed through device A, the packet outer carry the tag of vlan=2000, tpid=0x9100 to forward.

4.4 Adjustable Function of VLAN Tag TPID Value

4.4.1 Configure TPID Value of VLAN Tag to be Adjustable

TPID value of VLAN Tag

Operation	Command	Operation
Enter global configuration mode	**configure terminal**	-
Configure inner tag TPID value	**dtag inner-tpid** *tpid*	Optional 0x8100 by default
Restore default inner-tpid value	**no dtag inner-tpid**	optional
Enter port configuration mode	**interface ethernet** *device/slot/port*	-
Configure the outer tag TPID value	**dtag outer-tpid** *tpid*	Optional 0x8100 by default
Restore default outer-tpid value	**no dtag outer-tpid**	optional
Display tpid value	**show dtag**	any mode

📖 Note:

1. When double tag packet communicates with CPU, packet inner vlan tpid information should be in line with equipment configuration. Otherwise, it can’t not perform any communication.
2. When forwarded from uplink interface, the packet TPID value will be in line with the egress outer-tpid value.

4.4.2 Configuration Example for TPID Value Adjustable

Modify the inner tag tpid to be 0x9100
Switch(config)#dtag inner-tpid 9100
Modify the inner tag tpid to be 0x9200
Switch(config-if-ethernet-0/0/1)#dtag outer-tpid 9200
Display TPID information
Switch(config)#show dtag
Current dtag status: enabled
inner-tpid : 0x9100
outer-tpid : 0x9200
cpu inner-tag : vid 1 priority 0
interface : dtag-mode

4.5 GVRP Configuration

4.5.1 GVRP Overview

GVRP is abbreviations of GARP VLAN Registration Protocol. It is a kind of application of GARP, based on GARP working mechanism to maintain VLAN dynamic register information in switch and transfer it to other switches. All switches that support GVRP can receive VLAN register information from other switches and dynamically upgrade local VLAN register information which includes: current VLAN members, and by which interface can reach VLAN members. And all switches supported GVRP can transfer local VLAN register information to other switches to make the consistency of the VLAN information of devices which support GVRP. VLAN register information transferred by GVRP includes static register information of local manual configuration and the dynamic register information of other switch.

4.5.2 Enable GVRP

GVRP has two switches, one is under global configuration mode, and the other is under interface configuration mode. If you want to enable GVRP, these two switches should be in enable state.

By default, global GVRP and interface GVRP are disabled. It is important to note that, the GVRP can only be enabled on trunk port.

Enable GVRP

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	required
Enable global configuration GVRP	**gvrp**	required
Disable global configuration GVRP	**No gvrp**	required
Enter port configuration mode	interface ethernet *interface-num*	

Enable GVRP	**(no) gvrp**	required, Use “no command” to disable GVRP

4.5.3 Configure the VLAN which Needs GVRP to Forward

The VLAN registration information transmitted by GVRP can be the local static VLAN or the VLAN learned via GVRP protocol. Moreover, this vlan can be transmitted via GVRP only after the administrator specified.

Configure the VLAN which Needs GVRP to forward

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure the VLAN which Needs GVRP to forward	**(no) garp permit vlan** *vlan-list*	required, Use “no command” to cancel the configuration

4.5.4 Configure the VLAN which Forbids the Port to Forward

Configure the VLAN which Forbids the Port to Forward

Operation	Command	Remarks
Enter port configuration mode	interface ethernet *port_num*	-
Configure the vlan which forbids port transmission	**garp forbid** vlan *id*	optional
Configure the vlan which allows port transmission	no **garp forbid** vlan *id*	optional

4.5.5 GVRP Display and Debugging

After finishing the above configurations, you can check the configuration via the following command.

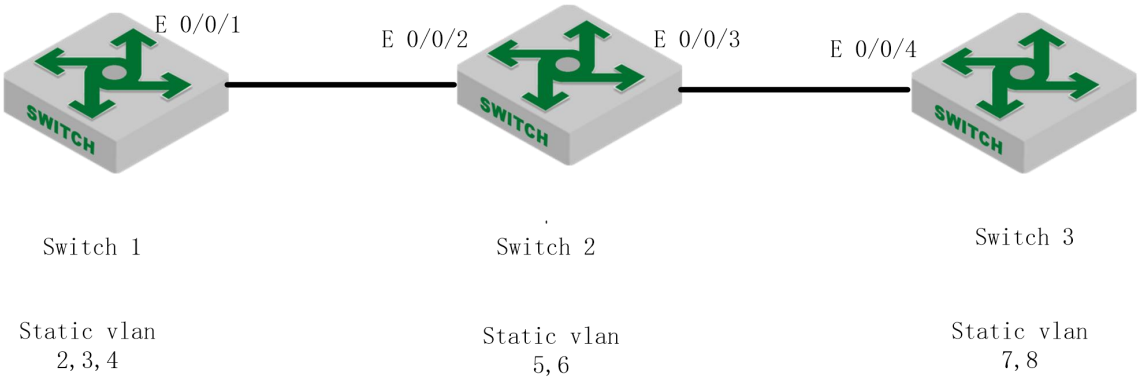
GVRP Display and Debugging

Operation	Command	Remarks
Display GVRP state (enable/disable)	**show gvrp**	It is executable in any modes.
Display the interface GVRP state (enable/disable)	**show gvrp interface [ethernet** device/slot/port**]**	
Display the VLAN which needs to be	**show garp permit vlan**	

transmitted by GVRP		

4.5.6 GVRP Configuration Example

1、Network requirements



Configuration example of GVRP service

As shown in the figure, switch 1 and switch 3 forward their static VLAN information to switch 2 via GVRP protocol packet. Switch 2 transmits itself static VLAN and the VLAN that it learns via GVRP. At last, the VLAN information in switch 1, switch 2 and switch 3 can be synchronized.

2、Configuration though:

1). Enable switch 1 GVRP, and then distribute VLAN information announcement
2). Enable switch 2 GVRP, and then distribute VLAN information announcement
3). Enable switch 3 GVRP, and then distribute VLAN information announcement

3、Configuration steps

Switch 1 configurations

Switch 1(config)# vlan 2-4

Switch 1(config-if-vlan)#switchport ethernet 0/0/1

Add VLAN port successfully.

Switch 1 (config-if-vlan)#interface e 0/0/1

Switch 1 (config-if-ethernet-0/0/1)#switchport mode trunk

Switch 1 (config-if-ethernet-0/0/1)#exit

Switch 1 (config)#gvrp // configure GVRP

Turn on GVRP successfully.

Switch 1 (config)#garp permit vlan 2-4

Switch 1 (config)#interface e 0/0/1

Switch 1 (config-if-ethernet-0/0/1)#gvrp

Switch 1 (config-if-ethernet-0/0/1)#exit

Switch 1 (config)#show gvrp // verify GVRP configuration

GVRP state : enable

Switch 1 (config)#show gvrp interface ethernet 0/0/1

port GVRP status fixed-vlan forbidden-vlan

e0/0/1 enable 1-4

Total entries: 1.

Switch 1 (config)#show garp permit vlan

VLAN 1 is Garp default permit VLAN

Other Garp permit VLAN :

2-4

Switch 2 configurations

Switch 2 (config)#vlan 5-6

Switch 2 (config-if-vlan)# switchport ethernet 0/0/2 to ethernet 0/0/3

Add VLAN port successfully..

Switch 2 (config-if-vlan)#exit

Switch 2 (config)#interface range ethernet 0/0/2 to ethernet 0/0/3

Switch 2 (config-if-range)# switchport mode trunk

Switch 2 (config-if-range)#exit

Switch 2 (config)#gvrp // configure gvrp

Turn on GVRP successfully

Switch 2 (config)#interface range ethernet 0/0/2 to ethernet 0/0/3

Switch 2 (config-if-range)#gvrp.

Switch 2 (config-if-range)#exit

Switch 2 (config)#garp permit vlan 5-6

Switch 2 (config)#show gvrp // verify gvrp configurations

GVRP state : enable

Switch 2 (config)#show gvrp interface ethernet 0/0/2 ethernet 0/0/3

port GVRP status fixed-vlan forbidden-vlan

e0/0/2 enable 1,5-6

e0/0/3 enable 1,5-6

Total entries: 2.

Switch 2 (config)#show garp permit vlan

VLAN 1 is Garp default permit VLAN

Other Garp permit VLAN : 5-6

Switch 3 configurations

Switch 3 (config)#vlan 7-8

Switch 3 (config-if-vlan)#switchport ethernet 0/0/4

Add VLAN port successfully.

Switch 3 (config-if-vlan)#interface e 0/0/4

Switch 3 (config-if-ethernet-0/0/4)#switchport mode trunk

Switch 3 (config-if-ethernet-0/0/4)#exit

Switch 3 (config)#gvrp // configure gvrp

Turn on GVRP successfully.

Switch 3 (config)#interface e 0/0/4

Switch 3 (config-if-ethernet-0/0/4)#gvrp

Switch 3 (config-if-ethernet-0/0/4)#exit

Switch 3 (config)#garp permit vlan 7-8

Switch 3 (config)#show gvrp // verify gvrp configurations

GVRP state : enable

Switch 3 (config)#show gvrp interface ethernet 0/0/4

port	GVRP status	fixed-vlan	forbidden-vlan
e0/0/4	enable	1,7-8	

Total entries: 1.

Switch 3 (config)#show garp permit vlan

VLAN 1 is Garp default permit VLAN

Other Garp permit VLAN : 7-8

After finishing the configuration, you can use the command of "**show vlan**" to check the VLAN registration information learned via GVRP.

Check the vlan information on switch 1, you can find that the vlan 5-8 is learned via GVRP.

Switch 1 (config)#show vlan

show VLAN information

VLAN ID : 1

VLAN status : static

VLAN member : e0/0/1-e0/2/2

Static tagged ports : e0/0/1

Static untagged Ports : e0/0/2-e0/2/2

Dynamic tagged ports :

show VLAN information

VLAN ID : 2

VLAN status : static

VLAN member : e0/0/1.

Static tagged ports : e0/0/1.

Static untagged Ports :

Dynamic tagged ports :

show VLAN information

VLAN ID : 3

VLAN status : static

VLAN member : e0/0/1.

Static tagged ports : e0/0/1.

Static untagged Ports :

Dynamic tagged ports :

show VLAN information

VLAN ID : 4

VLAN status : static

VLAN member : e0/0/1.

Static tagged ports : e0/0/1.

Static untagged Ports :

Dynamic tagged ports :

show VLAN information

VLAN ID : 5

VLAN status : dynamic

VLAN member : e0/0/1

Static tagged ports :

Static untagged Ports :

Dynamic tagged ports : e0/0/1

show VLAN information

VLAN ID : 6

VLAN status : dynamic

VLAN member : e0/0/1

Static tagged ports :

Static untagged Ports :

Dynamic tagged ports : e0/0/1

show VLAN information

VLAN ID : 7

VLAN status : dynamic

VLAN member : e0/0/1

Static tagged ports :

Static untagged Ports :

Dynamic tagged ports : e0/0/1

show VLAN information

VLAN ID : 8

VLAN status : dynamic

VLAN member : e0/0/1

Static tagged ports :

Static untagged Ports :

Dynamic tagged ports : e0/0/1

Total entries: 8 vlan.

4.6 N:1 VLAN Translation

4.6.1 N:1 VLAN Translate Overview

There are two types of N: 1 vlan translate, one is 1:1 vlan translate, and the other is N: 1 vlan translate. Definitions are as follows:

a) 1:1 VLAN translate: modify the VLAN tag from one specified VLAN packet to be a new VLAN tag.

b) N: 1 VLAN translate: modify the different VLAN tags from two or multiple specified VLAN packets to be the same VLAN tag.

N: 1 vlan translate function can be realized via vlan-swap or vlan-translate.

4.6.2 Vlan Translate Configuration

Configure vlan-translate

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable/disable vlan-translation	**[no]vlan-translation**	required

function		
Configure ingress vlan translate table	**vlan-translation-table interface ethernet** *device/slot/port old-vid new-vid priority*	required
Delete ingress vlan translate table	**no vlan-translation-table [interface ethernet** *device/slot/port old-vid* **]**	optional
Display vlan translate table	**show vlan-translation-table [interface ethernet** *device/slot/port old-vid* **]**	any mode

4.6.3 Configure vlan-swap (N: 1)

Configure vlan-swap

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**interface ethernet** *device/slot/port*	-
Configure vlan translate table	**vlan-swap** *start-vid end-vid swap-vid*	required
Delete vlan translate table	**no vlan-swap [all |** *start-vid end-vid*]	optional
Display vlan translate table	**show vlan-swap interface [ethernet** *device/slot/port*]	any mode

Note:

vlan-swap only supports ingress translate, and the ingress port should add new_vlan but it does not need to add old_vlan.

4.6.4 Configuration Example for N: 1 vlan-swap

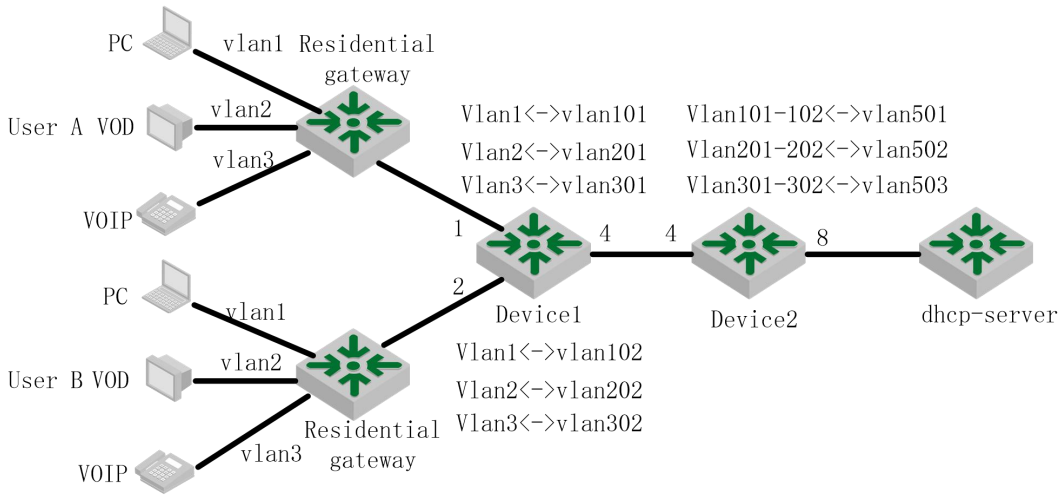
1.Network requirements

In a cell network, the service provider provides three types of data application services: PC, VoD and VoIP, which are connected to the home network. Each user accesses the switch through the home gateway and automatically obtains an IP address via DHCP. When distributing the home gateway to the user, the service provider performs unified configurations on the home gateway: The PC service is divided into VLAN 1, the VoD service is divided into VLAN 2, and the VoIP service is divided into VLAN 3.

In the corridor switch (Device1), it marked each service of every user with a separate VLAN in order to distinguish the same services of different users as well as to prevent information leakage and malicious attacks among users.

In the park switch (Device2), you need to classify data according to the service type to save VLAN resources. Among them: PC traffic is transmitted through VLAN 501, VoD service is transmitted through VLAN 502, and VoIP service is transmitted through VLAN 503.

Network diagram is as follows:



sketch map of N:1-swap

2.Configuration steps
Configure Device 1
Create a vlan, and add the port to the corresponding vlan
Switch(config)#vlan 1,2,3,101,201,301,102,202,302
Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#switchport hybrid tagged vlan 101,201,301
Switch(config-if-ethernet-0/0/1)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#switchport hybrid tagged vlan 102,202,302
Switch(config-if-ethernet-0/0/2)#interface ethernet 0/0/4
Switch(config-if-ethernet-0/0/4)#switchport hybrid tagged vlan 101,201,301,102,202,302

configure the vlan swap table on user ports e0 / 0/1 and e0 /0/ 2
Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#vlan-swap 1 1 101
Switch(config-if-ethernet-0/0/1)#vlan-swap 2 2 201
Switch(config-if-ethernet-0/0/1)#vlan-swap 3 3 301
Switch(config)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#vlan-swap 1 1 102
Switch(config-if-ethernet-0/0/2)#vlan-swap 2 2 202
Switch(config-if-ethernet-0/0/2)#vlan-swap 3 3 302

Configure Device2
Create the translation vlan (old_vlan and new_vlan) and the translation vlan includes the uplink and downlink ports
Switch(config)#vlan 101,201,301,102,202,302,501,502,503
Switch(config)#interface ethernet 0/0/4
Switch(config-if-ethernet-0/0/4)#switchport hybrid tagged vlan 101,102,201,202,301,302
Switch(config-if-ethernet-0/0/4)#interface ethernet 0/0/8
Switch(config-if-ethernet-0/0/8)#switchport hybrid tagged vlan 501,502,503

configure vlan-translate egress table under global configuration mode; enable the vlan-translate function on egress.
Switch(config)# vlan-translation

Switch(config)# vlan-translation-table interface ethernet 0/0/4 101 501 7
Switch(config)# vlan-translation-table interface ethernet 0/0/4 102 501 7
Switch(config)# vlan-translation-table interface ethernet 0/0/4 201 502 7
Switch(config)# vlan-translation-table interface ethernet 0/0/4 202 502 7
Switch(config)# vlan-translation-table interface ethernet 0/0/4 301 503 7
Switch(config)# vlan-translation-table interface ethernet 0/0/4 302 503 7

3.Result validation

Capture the packets on Device1 egress e0/4, you can capture the packet with tag of vlan 101, vlan201, vlan301, vlan202, vlan302;

Capture the packets on Device2 egress e0/8, you can capture the packet with tag of vlan 501, vlan502, vlan503;

4.7 MAC-Based VLAN Configuration

4.7.1 Overview for MAC-Based VLAN

As noted earlier, a single port in the campus network has multiple services, and each service belongs to different VLANs. So the flexible configuration of VLAN under the switch port to identify different services has become a key issue of the campus network management.

In order to solve the above-mentioned problems, the MAC-based VLAN is proposed. MAC (Media Access Control) address is burnt on a Network Interface Card (NIC), also known as the hardware address. It's composed of 48 bits long (6 bytes), 16 hex digits.

MAC-based VLAN is another way to distinguish VLAN that tag of VLAN is added to packet according to the source MAC address. This is often in combination with security technologies (such as 802.1X) to achieve the purpose of the terminal's safety and flexible access.

4.7.2 Configure MAC-Based VLAN

Users should bind the terminal MAC address with VLAN via the command line, and the device will generate a corresponding MAC VLAN table.

Configure MAC-Based VLAN

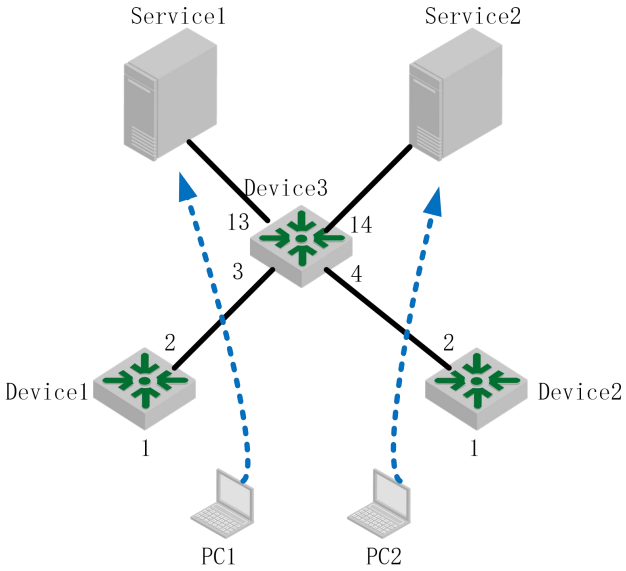
Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure static vlan-mac table	**vlan-mac-table** *mac-address vid priority*	required
Delete vlan-mac table	**no vlan-mac-table [***mac-address***]**	optional
Display vlan-mac table	**show vlan-mac-table[***mac-address***]**	any mode

4.7.3 Configuration Example for MAC-Based VLAN

1. Application request

As shown below, port 1 of Device1 and Device2 connects to two meeting rooms respectively; PC1 and PC2 are the laptops which will be used during the meeting.
PC1 and PC2 respectively belong to two departments, and these two departments are isolated by VLAN 100 and VLAN 200. The requirement is that no matter these two laptops are used in which meeting room; they can only access the servers of their own departments, which are server 1 and server 2. The Mac address of PC1 is 00:00:00:00:11:22 and the Mac address of PC2 is 00:00:00:00:11:33.

Network diagram is as follows:



Network diagram for MAC-Based VLAN

2. Configuration steps

(1) Configuration of Device1
create VLAN 100 and VLAN 200, and then configure the port 2 to be trunk port to allow the packet of VLAN 100 and VLAN 200 to pass through.
Switch>enable
Switch#configure terminal
Switch(config)#
Switch(config)#vlan 100,200
Switch(config-if-vlan)#exit
Switch(config)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#switchport mode trunk
Switch(config-if-ethernet-0/0/2)#switchport trunk allowed vlan 100,200

configure port 1 to be hybrid port, and remove the vlan tag when it forwards the packet of VLAN100 and VLAN200.
Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#switchport mode hybrid
Switch(config-if-ethernet-0/0/1)#switchport hybrid untagged vlan 100,200

create the MAC address of PC1 associates with VLAN100, create the MAC address of PC2 associates with VLAN200, enable MAC-VLAN function.
Switch(config)#vlan-mac-table 00:00:00:00:11:22 100 0
Switch(config)#vlan-mac-table 00:00:00:00:11:33 200 0

(2) Configuration of Device2

The configuration of device 2 is totally same as the configuration of device 1, so that won't be covered again here.

Configuration of Device3
create vlan 100 and vlan 200, and then add port 3 and port 4 to these two vlan.
Switch(config)#vlan 100,200
Switch(config-if-vlan)#switchport ethernet 0/0/3 ethernet 0/0/4

configure port 13 and port 14 to be trunk port to allow the packet of VLAN 100 and VLAN 200 to pass through.
Switch(config)#interface range ethernet 0/0/13 ethernet 0/0/14
Switch(config-if-range)#switchport mode trunk
Switch(config-if-range)#switchport trunk allowed vlan 100,200

3. Result validation

No matter these two laptops are used in which meeting rooms, they can only access the servers of their own departments

4.8 Protocol-Based VLAN Configuration

4.8.1 Overview for Protocol-Based VLAN

Protocol-based VLAN: the packet distributes different VLAN ID according to the receiving protocol types and encapsulation formats. “Protocol types + encapsulation formats” is also called model agreement. One protocol vlan can be able to bind multiple model agreements. Different model agreements can be distinguished by the vlan-protocol table index. Agreement template is referenced to the port, and then you can modify the packet vlan according to the model agreements.

Untagged packet processing (no vlan tag):

1. If the packet protocol types and encapsulation formats are conform to the model agreements, it will be tagged with the protocol vlan-id.
2. If the packet protocol types and encapsulation formats are not conforming to the model agreements, it will be tagged with the port default VLAN ID.

Tagged packet processing (has vlan tag):

1. If the packet protocol types and encapsulation formats are conform to the model agreements, the outer vlan information will be modified to be the protocol vlan-id.
2. If the packet protocol types and encapsulation formats are not conform to the model agreements, the processing mode will be the same as the port-based vlan.

This feature is mainly applied to bind the service type with VLAN, providing convenient management and maintenance.

There are two types’ configuration modes of protocol-based VLAN. Please choose the suitable one according to the equipment type.

4.8.2 Configure Protocol-Based VLAN

Configure Protocol-Based VLAN

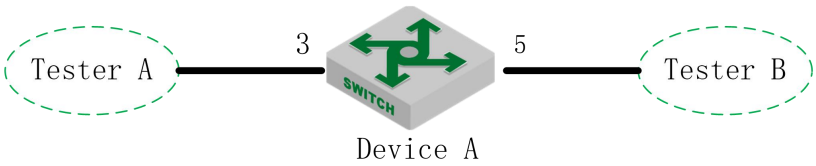
Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure protocol model	**vlan-protocol frametype {8023-llc-snap|8023-llc|ethernet2}** *ethertype* **interface ethernet** *device/slot/port vlan-id*	required
Delete protocol model	**no vlan-protocol [frametype {8023-llc-snap|8023-llc|ethernet2}** *ethertype* **interface ethernet** *device/slot/port* **]**	optional
Display the configuration of protocol model	**show vlan-protocol [frametype {8023-llc-snap|8023-llc|ethernet2}** *ethertype* **interface ethernet** *device/slot/port***]**	any mode

4.8.3 Example for Protocol-Based VLAN

1. Network requirements

Create vlan 10, and then configure the protocol model, the model index value is 1, protocol type is 0x0800, with ethernetv2 encapsulation.
It requires the encapsulated IP data flow of ethernetv2 from port 3 add the tag of vlan 10.

Network diagram is as follows:



Network diagram for Protocol-Based VLAN

2. Configuration steps

create protocol vlan 10 and then add it to all ports.
switch(config)#vlan 10
switch(config-if-vlan)#switchport all
Add VLAN port successfully.

configure vlan 10 of port 5 to be tag attribute transmission.
switch(config)#interface ethernet 0/0/5
switch(config-if-ethernet-0/0/5)#switchport hybrid tagged vlan 10
switch(config-if-ethernet-0/0/5)#exit

create protocol model, protocol type to be 0x0800 with ethernetv2 encapsulation
switch(config)#vlan-protocol table index 1 ethertype 0800 protocol ethernetv2

configure the ingress enables the vlan protocol function firstly. Next, bind protocol template index and configure protocol vlan10.
switch(config)#interface ethernet 0/0/3
switch(config-if-ethernet-0/0/3)#vlan-protocol
switch(config-if-ethernet-0/0/3)#vlan-protocol table index 1 vlan 10

3． Result display and verification:
switch(config)#show vlan-protocol table
index ethertype protocol
1 0x0800 EthernetV2
switch(config)#show vlan-protocol interface ethernet 0/0/3
e0/0/3: : enable
global protocol-vlan table index 1 vlan 10
result: all the ethernetv2 IP data flow entering from port 3 should add vlan 10 tag before transmitting.

4.9 IP-subnet VLAN

4.9.1 Overview for IP Subnet-Based VLAN

IP subnet-based vlan is divided according to packet source IP address and subnet mask. After device received packets from the interface, it will confirm the packets belonging to which VLAN and then automatically divide these packets to specified VLAN to transmit.

This feature is mainly used for the specified IP address or network segment message transmission in the specified VLAN. Currently, our company S5300 BCM series, S5330 BCM series and S5900-24S - BCM possess this function. Please refer to the corresponding products for more details.

4.9.2 Configure IP Subnet-Based VLAN

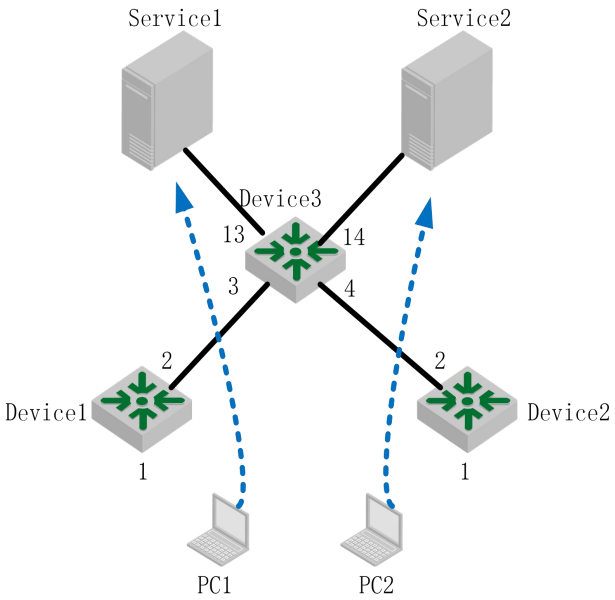
IP Subnet-Based VLAN

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable (disable) the VLAN based on IP subnet	**[no]vlan-subnet precede**	required
Configure the table of the VLAN based on IP subnet	**ip-subnet-vlan** *ipaddress mask vlan-id priority*	required
Delete IP subnet table	**no ip-subnet-vlan** *ipaddress mask*	optional
Display IP subnet table	**show ip-subnet-vlan [***ipaddress*| *mask***]**	any mode

4.9.3 Configuration Example

1.Network requirements

An enterprise network allocates IP subnet according to service type. The requirement is that different subnet users adopt different transmission paths to access upstream server.
As shown below:



Network diagram of IP Subnet-Based VLAN

The packets of device1 include data, IPTV, voice and so on. Their IP addresses are different from each other. Configure the IP Subnet-Based VLAN in device 1. After received the service packets, device will automatically divide these packet to specified VLAN according to different source IP. Moreover, device will forward these packets to the upper server.

2. Configuration steps

create VLAN and it should include the interfaces.
Switch(config)#vlan 100,200,300
Switch(config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/2 ethernet 0/0/3

enable the vlan based on IP subnet, and then configure the table of IP subnet
Switch(config)#vlan-subnet precede
Switch(config)#ip-subnet-vlan 192.168.1.1 255.255.255.0 100 0
Switch(config)#ip-subnet-vlan 192.168.1.2 255.255.255.0 200 0
Switch(config)#ip-subnet-vlan 192.168.1.3 255.255.255.0 300 0
Switch(config)#
note: please ensure the uplink interface vlan100、vlan200、vlan300 with the tag.

3. Result validation
Switch(config)#show run garp
![GARP]
vlan-subnet precede
ip-subnet-vlan 192.168.1.1 255.255.255.0 100 0
ip-subnet-vlan 192.168.1.2 255.255.255.0 200 0
ip-subnet-vlan 192.168.1.3 255.255.255.0 300 0
Upon testing, the service message can only be transmitted to the specified server.

4.10 VLAN-trunking Configuration

4.10.1 VLAN-trunking Overview

Vlan-trunking is used to transmit transparently the unknown message. Currently, only S5650-MVL/S2600-MVL chip possesses this function. Explanation:

1. Each switch only configures a set of vlan-trunking.
2. Each vlan-trunking includes two interfaces: unlink and downlink.
3. If transmits the unknown message and do not perform any change.
4. It does not learn the mac address of the unknown message.
5. The mechanism of the known message is the same as 802.1Q VLAN.

4.10.2 Configure Vlan-trunking

Configure vlan-trunking

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure vlan-trunking mode	**vlan-trunk-mode [auto | manual]**	required
Enter infterface configuration mode	**interface ethernet** *device/slot/port*	
(Disable) Enable vlan-trunking	**(no) vlan-trunking [allow pass vlan all]**	required
Display vlan-trunking	**show vlan-trunking**	any mode

5. MAC Address Table Configuration

5.1 Configure MAC Address Table

The system maintains a MAC address table for forwarding packets. The entry in this table contains the device MAC address, VLAN ID, and Switch port number which the packets entering. When a packet enters the Switch, the Switch looks up the MAC address table based on the destination MAC address of the packet and the VLAN ID of the packet. If the packet is found, the Switch sends the packets to the specified ports. Otherwise, Switch broadcast the packets in this VLAN.

The system can be able to learn MAC address table. If the source MAC address of a received packet does not exist in the MAC address table, the system will add the source MAC address, VLAN ID, and port number of the received packet as a new entry to the MAC address table.

You can manually configure MAC address entries. The administrator can configure the MAC address table based on the actual network condition, that is, the administrator can add or modify static entries, permanent entries, blackhole entries, dynamic entries.

System provides MAC address aging function. If a device does not send any packets within a certain period of time, the system deletes the MAC address entries associated with the device. MAC address aging only takes effect on the learned MAC address or the MAC address entries which can be aged (the dynamic MAC address entries which are configured by the user).

5.1.1 Configure MAC Address Table Aging Time

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
	mac-address-table age-time {*second*| disable}	Aging time unit is second; *disable* means mac address will not be aged
Display the MAC address aging time	show mac-address-table age-time	

5.1.2 Add MAC Address Table by Manual

In addition to dynamically learned entries, the MAC address table can be manually added.

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
	mac-address-table { static | permanent | dynamic } H:H:H:H:H:H interface Ethernet *interface-num* vlan *vlan-id*	
Display mac table	show mac-address-table { static| permanent| dynamic| blackhole|vlan| interface ethernet ***port-number*** }	

static: static mac address, and it will not be aged
permanent: permanent mac address. It will be aged. If you save the configurations, the entries will still exist after the device is powered down.
dynamic: dynamic mac address, and it will be aged

5.1.3 Add Blackhole MAC Address

The Switch can configure a MAC address entry as a blackhole entry. When the source address or destination address of the packet is a blackhole MAC address, the Switch discards the packet.

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
	mac-address-table blackhole H:H:H:H:H:H vlan *vlan-id*	

5.1.4 Enable/disable MAC Address Learning

You can configure whether the device learns MAC addresses dynamically or not.

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable/disable MAC address learning	(no) mac-address-table learning	Enabled by default
Enter port configuration mode	interface ethernet ***port-number***	
Enable/disable MAC address learning	(no) mac-address-table learning	
Display the status of MAC address learning(enable/disable)	show mac-address learning [interface [interface-num]]	

Note:
If MAC address learning is disabled under global configuration mode, all ports cannot learn MAC address; If you want to disable mac address learning on some ports, just enable MAC address learning under global configuration mode and disable MAC address learning on the port will be OK.

5.1.5 Quantity Limitation on MAC Address Learning Table

Under port configuration mode, you can configure the maximum number of learned MAC addresses on a port. By default, the number of MAC addresses learning table are unlimited.

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter interface configuration mode	interface ethernet *port-number*	
Enable quantity limitation on mac address learning table	mac-address-table max-mac-count *integer*	It is effective for a single port
Disable quantity limitation on mac address learning table	no mac-address-table max-mac-count	
Display the maximum number of MAC address learning on a port	show mac-address max-mac-count [interface [interface-num]]	

5.2 Local-switch Function

Normally, the Switch does not forward the incoming packets from the port. However, you may need to forward the packets coming from the port sometimes. In this case, you can use the local-switch.

5.2.1 Configure Local-switch

Configure Local-switch

Operation	Command	Remarks
Enter interface configuration mode	interface ethernet ***port-num***	-
Configure the local forwarding function	**[**no**]** Local-switch	optional
Display the configuration	show local-switch **[**interface ethernet ***port-num*** **]**	optional

5.2.2 Configuration Example for Local-switch

1.Network requirements
Enable local forwarding on port 0/0/1.

2.Configuration steps
Switch(config)#show local-switch interface ethernet 0/0/1
port local-switch-state
e0/0/1 disable
Total entries: 1 .

Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#local-switch
Setting successfully! local-switch is enable

Switch(config)#interface ethernet 0/0/1
Switch(config-if-ethernet-0/0/1)#show local-switch interface ethernet 0/0/1
port local-switch-state
e0/0/1 enable
Total entries: 1 .

Switch(config-if-ethernet-0/0/1)#no local-switch
Setting successfully! local-switch is disable

Switch(config-if-ethernet-0/0/1)#show local-switch interface ethernet 0/0/1
port local-switch-state
e0/0/1 disable
Total entries: 1 .

5.3 Slf-control Function

Whether the Switch forwards source unknown packets or not needs to be managed by the network administrator according to the security policy. By default, Switch forwards the source unknown packets. You can disable the source unknown packet forwarding function by using the specified command, in that case, when the device receives a packet, it checks whether the source MAC exists in the MAC table or not. If it does not exist, it discards the packet, that is, it can only forward the packet whose source address is being known.

5.3.1 Configure Slf-control

Configure Slf-control

Operation	Command	Remarks
Enter port configuration mode	**interface ethernet** *port-num*	required
Disable the forwarding function that source mac	**[no]src_dlf_forward**	required

address is being unknown		
Enable the forwarding function that source mac address is being unknown	**src_dlf_forward**	
Display the configuration	**show src_dlf_forward interface** [**ethernet** ***port-num***]	optional

This function is usually combined with the port MAC address learning function or port MAC address limit function.

5.3.2 Configuration Example for Slf-control

1.Network requirements
Disable the forwarding function of source unknown packets on the port0/0/9.

2.Configuration steps
Switch(config)#show src_dlf_forward interface ethernet 0/0/9
Port src_dlf_forward status
0/0/9 enable

Switch(config)#interface ethernet 0/0/9
Switch(config-if-ethernet-0/0/9)#no src_dlf_forward
Switch(config-if-ethernet-0/0/9)#no mac-address-table learning
Switch(config-if-ethernet-0/0/9)#show src_dlf_forward interface ethernet 0/0/9
Port src_dlf_forward status
0/0/9 disable

5.4 DLF-control Overview

Unknown packets are classified into unknown unicast packets and unknown multicast packets

Unknown unicast packets are packets that cannot find the destination MAC address of the packets in the MAC table.

Unknown multicast packets are packets that cannot find the destination MAC address of the multicast packets in the multicast MAC table.

5.4.1 Configure DLF-control

If enable based on the global configuration, this command will take effect on egress packets of all ports;

If enable based on the interface configuration, this command will take effect on egress

packets of this port.

By default, unknown packets are allowed to be forwarded.

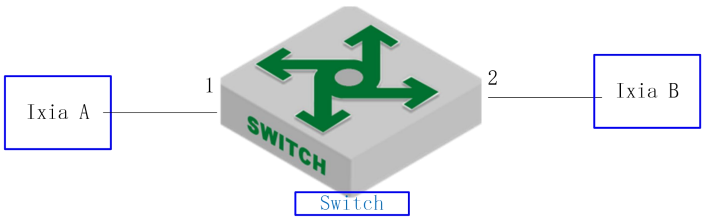
Configure dlf-forward

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable the forwarding function of unknown unicast packets	[no]dlf-forward unicast	Optional. Enabled by default.
Enable the forwarding function of unknown multicast packets	[no]dlf-forward multicast	Optional. Enabled by default.
Enter interface configuration mode	interface ethernet ***port-number***	-
Enable the forwarding function of unknown unicast packets	[no]dlf-forward unicast	Optional. Enabled by default.
Enable the forwarding function of unknown multicast packets	[no]dlf-forward multicast	Optional. Enabled by default.
Display dlf-forward configuration	show dlf-forward interface [ethernet ***port-number***]	optional

5.4.2 Configuration Example for DLF-control

1.Network requirements

Configure the port 2 egress not to forward unknown unicast packets.



sketch map for Dlf-control

2.Configuration steps

Disable the port 2 forwarding function of unknown unicast packets
Switch(config-if-ethernet-0/0/1)#no dlf-forward unicast

Display the configurations
Switch(config-if-ethernet-0/0/1)#show dlf-forward interface ethernet 0/0/2
Forwarding unknown unicast packets global status: enable
Forwarding unknown multicast packets global status: enable
Port Forwarding Unknown Unicast Forwarding Unknown Multicast
e0/0/1 disable enable

3.Result validation

(1) The tester A sends an unknown packet at line speed, and the tester B does not receive the packet.

(2) The tester A sends a known packet at line speed, and the tester B receives the packet.

6. Multicast Configuration

6.1 IGMP-Snooping Configuration

6.1.1 Overview for IGMP-Snooping

IGMP (Internet Group Management Protocol) is a part of IP protocol which is used to support and manage the IP multicast between host and multicast router. IP multicast allows transferring IP data to a host collection formed by multicast group. The relationship of multicast group member is dynamic and host can dynamically add or exit this group to reduce network load to the minimum to realize the effective data transmission in network.

IGMP Snooping is used to monitor IGMP packet between host and routers. It can dynamically create, maintain, and delete multicast address table according to the adding and leaving of the group members. At that time, multicast frame can transfer packet according to its own multicast address table.

6.1.2 Enable igmp-snooping

Configure igmp-snooping

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable IGMP Snooping	**igmp-snooping**	required
Disable IGMP Snooping	no **igmp-snooping**	optional

6.1.3 Configure igmp-snooping Timer

Configure igmp-snooping timer

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the aging time of dynamic multicast members	igmp-snooping **host-aging-time** *time*	Optional; By default, the aging time of dynamic multicast member is 300s
Disable the aging time for dynamic multicast member	**no** igmp-snooping **host-aging-time**	optional

Igmp Snooping queries the maximum response time configuration	igmp-snooping **max-response-time** *time*	optional Configure the maximum waiting time for group ports to be removed after receiving leave messages. The default setting is 10 seconds.
Disable the maximum response time configuration for the Igmp snooping query	no igmp-snooping **max-response-time**	optional

6.1.4 Configure fast-leave

Generally, after receiving an IGMP leave message, IGMP-Snooping does not delete the port directly from the multicast group. Instead, it waits for a period of time before deleting the port from the multicast group.

If enables the fast-leave, IGMP Snooping removes the port directly from the multicast group when receiving an IGMP Leave message. When there is only one user under the port, the fast-leave can save the bandwidth.

Configure fast-leave

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter interface configuration mode	interface ethernet *port-num*	
Configure fast-leave	igmp-snooping **fast-leave**	optional By default, the fast-leave function is disabled.
Cancel the fast-leave function	no igmp-snooping **fast-leave**	optional

6.1.5 Configure the Maximum Learning Number of Multicast Groups

Use the following commands to configure the maximum learning number of multicast groups.

Configure the maximum learning number of multicast groups

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter port configuration mode	interface ethernet *port-num*	
Configure the	igmp-snooping **group-limit** *number*	optional

maximum learning number of multicast groups		
Disable the configuration of the maximum number of multicast packets learning	no igmp-snooping **group-limit**	optional
Configure the action when the port is full of multicast groups	igmp-snooping group-limit **action { replace | drop}**	optional

Note:

Igmp-snooping group-limit not only refers to the maximum number of multicast which the port can learn, but also refers to the maximum number of multicast which the machine can learn. The maximum number of each product may be different.

6.1.6 Configure igmp-snooping Multicast Learning Strategy

After a multicast learning strategy is configured, the administrator can control the router to learn only a specific multicast group. If a multicast group is added to the blacklist, the router will not learn the multicast group; on the contrary, if a multicast group is added to the whitelist, the router will learn the multicast group.

configure igmp-snooping multicast learning strategy

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the default learning rule for multicast groups that are not in the blacklist or whitelist	igmp-snooping { permit | deny } { group all | vlan *vid* }	optional By default, the learning rule of a multicast group that is not in the blacklist or whitelist is to learn all multicast groups.
Enter interface configuration mode	interface ethernet *port-num*	
Configure the blacklist and the whitelist of the multicast group	igmp-snooping { permit | deny } group-range *MAC* multi-count *num* vlan *vid*	optional
	igmp-snooping { permit | deny } group *MAC* vlan *vid*	optional By default, no multicast group will be added to the blacklist or whitelist.

6.1.7 Configure IGMP Snooping Querier

In a multicast network running IGMP, a multicast router or a Layer 3 multicast router is responsible for sending IGMP query messages.

However, the IGMP function is not supported on the Layer 2 Switch. Therefore, the querier function cannot be implemented and the common-group query cannot be sent. You can

configure an IGMP Snooping querier to enable the Layer 2 Switch to actively send a general query message at the data link layer to establish and maintain multicast forwarding entries.

You can also configure the VLAN, source address, maximum response time, and query interval for IGMP Snooping queriers to send general query messages.

Configure igmp-snooping Querier

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable IGMP-Snooping querier	igmp-snooping **querier**	required
Disable IGMP-Snooping querier	no igmp-snooping querier	optional
Configure the version of the query	igmp-snooping **querier** version *id*	Optional; Version2 by default.
Configure VLANs for general query packets	igmp-snooping **querier-vlan** *vid*	optional
Remove the VLAN configurations of the general query packets	no igmp-snooping querier-vlan *vid*	optional
Configure the interval for sending general query messages	igmp-snooping **query-interval** *interval*	optional
Remove the interval of sending general query messages	no igmp-snooping **query-interval**	optional
Configure the maximum response time for general query messages	igmp-snooping **query-max-respond** *time*	optional
Disable the maximum response time configuration for general query messages	no igmp-snooping **query-max-respond**	optional
Configure the source IP address for sending general query messages	igmp-snooping **general-query source-ip** *ip*	optional
Disable the source IP address for sending general query messages	no igmp-snooping **general-query source-ip**	optional

6.1.8 Configure the Routing Port

You can add a route port to the dynamic multicast learned by IGMP Snooping automatically so that the routing port can forward multicast traffic.

When an Switch receives a membership report from a host, the Switch forwards the report to the routing port.

Configure the Routing Port

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the hybrid routing port function	igmp-snooping **route-port forward**	optional
Disable the hybrid routing port function	no igmp-snooping **route-port forward**	optional
Configure the aging time of the dynamic routing port	igmp-snooping **router-port-age** { on | off | *age-time* }	optional

Restore the aging time of the dynamic routing port	no igmp-snooping **router-port-age**	Aged by default, and the value is 300s;
Configure a static route port	igmp-snooping **route-port** vlan *vid* interface { all | * thernet *interface-num* }	optional
Disable the static routing port	no igmp-snooping **route-port** vlan *vid* interface { all | * thernet *interface-num* }	optional

6.1.9 Configure Multicast VLAN

After the multicast VLAN function is enabled on a port, the Switch changes the IGMP packet to a multicast VLAN regardless of the VLAN to which the received IGMP packet belongs.

Configure Multicast VLAN

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter interface configuration mode	interface ethernet *port-num*	
Configure a multicast VLAN for the port	igmp-snooping **multicast vlan** *vid*	optional
Disable the multicast VLAN for the port	**no** igmp-snooping **multicast vlan**	optional

6.1.10 Configure the Port to Record the Host MAC Address

When this function is enabled on the port, the Switch records the source mac address of the igmp report packet.

Configure the Port to Record the Host MAC Address

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter interface configuration mode	interface ethernet *port-num*	
Configure the port to record the host MAC address	igmp-snooping **record-host**	optional
Disable the port to record the host MAC address	no igmp-snooping **record-host**	optional

6.1.11 Configure the Suppression Multicast Report

After enabling igmp-snooping report-suppression:

1） Each group will only send a report to the mroute port(When the first report is received, the source mac is replaced with the mac from the Switch and sent to the mroute port), the report will not forward to client report. If receive the report of the same group later, only the local member or timer information will be updated and the report will not send to the mroute port.

2） After receiving the general query, the Switch encapsulates all the packets in the report packet to the mroute port, and then forwards the query to all clients. When receiving a specific query,

the Switch encapsulates the specified group into a report packet and sends it to the mroute port. The Switch then forwards the query to the specified client. If the Switch does not learn the specified group, it will directly discard the query;

3）；After receiving leave report, if there are other members in the group, only delete the member, do not send leave report to the mroute port; If it is the last member to leave, just replace the source mac with the Switch mac and then send the it to the mroute port;

Configure the Suppression Multicast Report

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the suppression multicast report	igmp-snooping **report-suppression**	optional
Disable the suppression multicast report	no igmp-snooping **report-suppression**	optional

6.1.12 Configure Whether to Drop Query / Report Packets

When this function is enabled on the port, the device discards IGMP query / report packets. The default port receives all IGMP messages.

Configure Whether to Drop Query / Report Packets

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter interface configuration mode	interface ethernet *port-num*	
Configure the port to drop query / report packets	**igmp-snooping drop {query|report}**	optional
Configure the port to receive query / report packets	no **igmp-snooping drop** {query|report}	optional

6.1.13 Configure the Multicast Preview Function

IGMP Snooping provides multicast preview function. You can configure the multicast preview channel. You can also configure the single preview duration, preview interval, preview reset duration, and allowable preview times of multicast.

Configure the Multicast Preview Function

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the multicast preview function	igmp-snooping **preview**	-optional
Disable the multicast preview function	**no** igmp-snooping **preview**	optional
Configure the multicast preview channel	igmp-snooping **preview group-ip** *IP* vlan *vid* interface ethernet *interface-num*	optional
Disable the multicast preview channel	no igmp-snooping **preview group-ip** *IP* vlan *vid* interface ethernet *interface-num*	optional
Configure the single preview duration,	igmp-snooping **preview** { time-once *time-once* time-interval *time-interval* time-reset *time-reset*	optional

preview interval, preview reset duration, and allowable preview times	permit-times *preview-times* }	
Disable the single preview duration, preview interval, preview reset duration, and allowable preview times	no igmp-snooping **preview** { time-once *time-once* time-interval *time-interval* time-reset *time-reset* permit-times *preview-times* }	optional

6.1.14 Configure the Profile Black and White List

IGMP snooping provides profile blacklist and whitelist function. It creates several profiles under the global configuration mode, and then configures the profile list referenced by the port under the interface configuration mode. You can configure the type and range of the IGMP Snooping profile, where type is permit / deny, and the range can be configured to use the multicast IP address or MAC address. IGMP snooping profile takes effect only when it is referenced by a port. To configure a port to reference a profile, you must specify the same type for multiple ports. That is, a port can reference only one type profile (permit or deny). When a port references the permit profile, you can only learn the multicast group defined by the profile. When a port references a deny profile, you can learn all the multicast groups except the profile definition. If the port does not reference any profile, learn the multicast group as usual.

Configure the Profile Black and White List

Operation	Command	Remarks
Enter global configuration mode	configure terminal	--
Create a profile and enter profile configuration mode	igmp-snooping **profile** *profile-id*	--
Disable profile configuration	no igmp-snooping **profile** *profile-id*	
Configure profile type	profile limit { permit | deny }	optional
Configure the range of profile IP	ip range *start-ip end-ip* [vlan *vlan-id*]	optional
Configure the range of profile mac	mac range *start-mac end-mac* [vlan *vlan-id*]	optional
Enter interface configuration mode	interface ethernet *interface-num*	--
Configure profile reference of the port	igmp-snooping **profile refer** *profile-list*	optional
Disable profile reference of the port	no igmp-snooping **profile refer** *profile-list*	optional

6.1.15 Igmp-snooping Display and Maintenance

After you completed the above configuration, you can use the following commands to view the configurations.

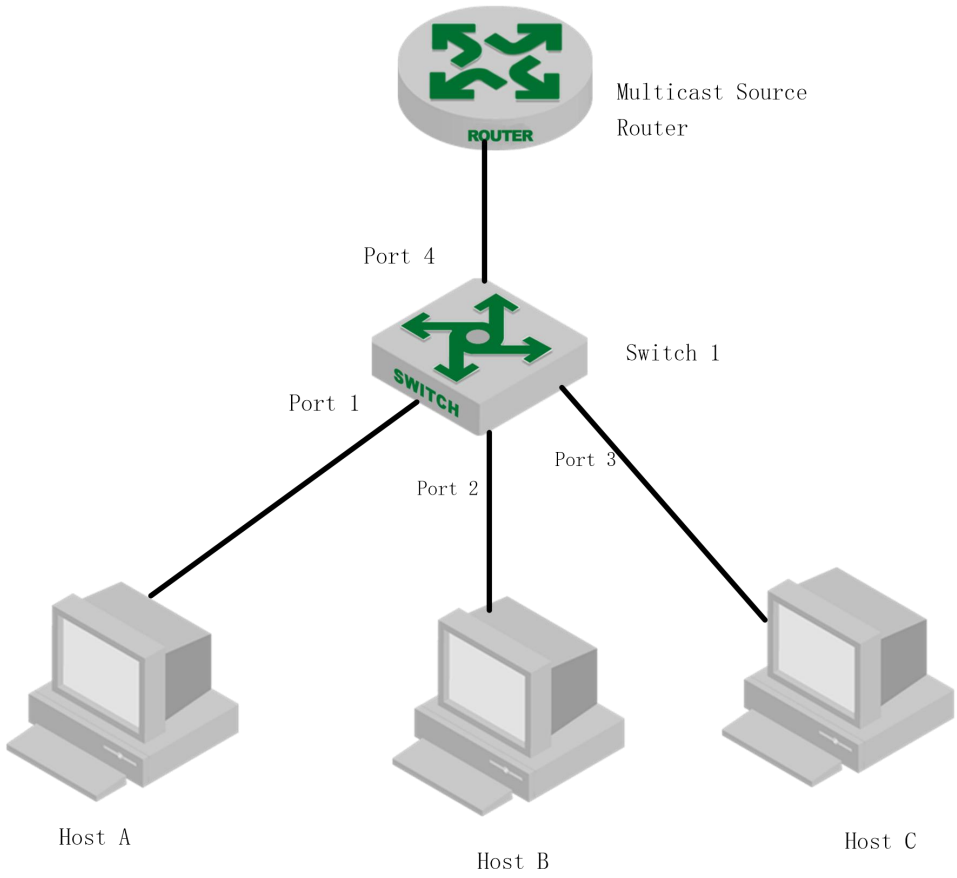
Display the related configurations of igmp-snooping

Operation	Command	Remarks
Display the related configurations of IGMP snooping	show igmp-snooping	All modes are executable
Display the dynamic routing port	show igmp-snooping router-dynamic	
Display the routing port of static configuration	show igmp-snooping router-static	

Display the statistics of igmp-snooping packets	#show igmp-snooping statistics{ interface| *vlan*}	
Display the host MAC of the record	show igmp-snooping record-host [interface]	
Display the multicast preview information	show igmp-snooping preview	
Display the current multicast preview channel status	show igmp-snooping preview status	
Display the configuration of the profile	show igmp-snooping profile	
Display the (brief) information of multicast table	show multicast	
Display the (detailed) information of multicast table	show multicast igmp-snooping interface	

6.1.16 Configuration Example

1、Network requirements



igmp-snooping configuration example

As shown above, Host-A, Host-B, and Host-C belong to VLAN 2, VLAN 3, and VLAN 4 respectively. The three hosts are configured to receive the multicast group data with the group addresses 224.0.1.1 to 224.0.1.3 respectively.

2、Configuration steps:
a. Enable igmp-snooping;
b. Add different ports to different VLANs
c. The host sends a report packet to the Switch and the Switch learns the multicast group
d. The multicast source router sends a query packet to the Switch and the Switch learns the routing port entries
e. The multicast source router sends the multicast service data flow to the Switch, and the Switch distributes them to the corresponding host.

3、Result validation:
Enable igmp snooping
Switch (config)#igmp-snooping

Configure VLAN2, VLAN3 and VLAN4, and then add Ethernet0 / 1, Ethernet0 / 2, and Ethernet0 / 3 to VLAN 2, VLAN 3 and VLAN 4 respectively.
Switch (config)#vlan 2
Switch (config-if-vlan)#switchport ethernet 0/0/1
Switch (config-if-vlan)#exit
Switch (config)#vlan 3
Switch (config-if-vlan)#switchport ethernet 0/0/2
Switch (config-if-vlan)#exit
Switch (config)#vlan 4
Switch (config-if-vlan)#switchport ethernet 0/0/3
Switch (config-if-vlan)#exit

When Host-A, Host-B, and Host-C send igmp reports to the Switch, the Switch will learn the corresponding multicast group entry. When the multicast source router sends igmp query packets to the Switch, the Switch will learn the corresponding routing port entries.

Display the multicast groups learned by the Switch
Switch (config)#show multicast
show multicast table information
MAC Address : 01:00:5e:00:01:01
VLAN ID : 2
Static port list :.
IGMP port list : e0/0/1
Dynamic port list :

MAC Address : 01:00:5e:00:01:02
VLAN ID : 3
Static port list :.
IGMP port list : e0/0/2
Dynamic port list :

MAC Address : 01:00:5e:00:01:03
VLAN ID : 4
Static port list :
IGMP port list : e0/0/3.
Dynamic port list :

Total entries: 3 .
Switch (config)#show igmp-snooping router-dynamic
Port VID Age Type

e0/0/4 2 284 { STATIC }
e0/0/4 3 284 { STATIC }s
e0/0/4 4 284 { STATIC }
Total Record: 3

When the Multicast Source Router sends multicast traffic of 224.0.1.1~ 224.0.1.3, the Switch will distribute the corresponding traffic flow to Host-A, Host-B and Host-C.

6.2 MLD Snooping Configuration

6.2.1 MLD Snooping Overview

MLD (Multicast Listener Discovery) is part of the IPv6 protocol, using to support and manage the IP multicast between the host and the multicast router. IP multicast allows IP datagrams to be transmitted to a set of hosts that make up a multicast group. The relationships among multicast group members are dynamic, that is, Hosts can dynamically join or leave groups to minimize network load so as to achieve the effective data transmission.

MLD snooping is used to monitor the MLD packets between the host and the router. The MLD snooping dynamically creates, maintains, and deletes the multicast address table based on the joining and leaving of the multicast group members. In this case, multicast frames are forwarded according to the multicast address table.

6.2.2 Enable MLD Snooping

Enable MLD Snooping

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable MLD Snooping	**mld-snooping**	required

6.2.3 Configure MLD Snooping Timer

Configure MLD Snooping Timer

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the aging time of dynamic multicast member ports	**mld-snooping host-aging-time** *time*	Optional; By default, the aging time

		of dynamic multicast member ports is 300 seconds
Configure the maximum response time of the leave packets	**mld-snooping max-response-time** *time*	Optional; By default, the maximum response time is 10 seconds

6.2.4 fast-leave

Normally, when receiving an MLD leave message, MLD-Snooping will not delete the port from the multicast group directly. Instead, it waits for a period of time to remove the port from the multicast group.

After fast-leave is enabled, MLD-Snooping removes the port from the multicast group directly when received the MLD leave packet. When there is only one user under the port, the fast-leave can save the bandwidth.

Configure fast-leave

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter port configuration mode	**interface ethernet** *interface-num*	-
Configure fast-leave	**mld-snooping fast-leave**	optional By default, fast-leave is disabled.

6.2.5 Configure the Maximum Number of Multicast Groups

You can use the following commands to set the maximum number of multicast groups that can be learned on each port.

Configure the Maximum Number of Multicast Groups

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter port configuration	**interface ethernet** *interface-num*	-

mode		
Configure the maximum number of multicast groups	**mld-snooping group-limit** *number*	optional By default, the maximum number of multicast groups is NUM_MULTICAST_GROUPS

Note:

NUM_MULTICAST_GROUPS refers to the largest number of multicast packets that the machine can be able learn. The NUM_MULTICAST_GROUPS of different products may be different.

Theoretically, the maximum number of multicast packets is NUM_MULTICAST_GROUPS, but it also means that the number of multicast packets learned by other ports will be occupied. That is,

All ports will share multicast group resource of this NUM_MULTICAST_GROUPS.

6.2.6 Configure the Multicast Learning Strategy of MLD Snooping

After a multicast learning strategy is configured, the administrator can control the router to learn only one specific multicast group. If a multicast group is added to the blacklist, the router will not learn the multicast group; on the contrary, the multicast group router in the white list can be learned.

Configure the Multicast Learning Strategy of MLD Snooping

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the default learning rule for multicast groups that are not in the blacklist or whitelist	**mld-snooping { permit | deny } { group all | vlan** *vid* **}**	Optional; By default, the learning rule of a multicast group that is not in the blacklist or whitelist is to learn all multicast groups.
Enter interface configuration mode	**interface ethernet** *interface-num*	-
Configure blacklist and whitelist	**mld-snooping { permit | deny } group-range** *MAC* **multi-count** *num* **vlan** *vid*	optional

for port multicast	**mld-snooping { permit | deny } group** *MAC* **vlan** *vid*	Optional; By default, no multicast group will be added to the blacklist and whitelist

6.2.7 Configure the MLD-Snooping Querier

In a multicast network running the MLD protocol, a multicast router or a Layer 3 multicast router with full-time query is responsible for sending MLD queries.

However, because Layer 2 Switch does not support MLD, there is no ways to implement the querier function and cannot send general query messages. You can configure the MLD-Snooping querier so that the Layer 2 Switch can actively send a general-purpose group query message at the data link layer to establish and maintain a multicast forwarding entry.

Users can also configure the MLD snooping querier to forward the source address, maximum response time, and query interval for sending general query messages.

Configure the MLD-Snooping Querier

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable MLD-Snooping querier	**mld-snooping querier**	required
Configure the interval for sending general query messages	**mld-snooping query-interval** *interval*	optional
Configure the maximum response time for general query messages	**mld-snooping query-max-respond** *time*	optional

6.2.8 Configuring the Routing Port

The route port can be added to the dynamic multicast learned by MLD Snooping automatically. In that case, the routing port can forward multicast traffic.

When an Switch receives a membership report from a host, the Switch forwards the report to the routing port.

Configuring the Routing Port

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the hybrid routing	**mld-snooping route-port forward**	optional

port function		
Configure the aging time of the dynamic routing port	**mld-snooping router-port-age { on | off |** *age-time* **}**	optional
Configure a static route port	**mld-snooping route-port vlan** *vid* **interface { all | ethernet** *interface-num* **}**	optional

6.2.9 Configure a Multicast VLAN

After enabling the multicast VLAN function on a port, the Switch changes them to a multicast VLAN regardless of the VLAN to which the MLD messages belong.

Configure a Multicast VLAN

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**interface ethernet** *interface-num*	-
Configure a multicast VLAN	**mld-snooping multicast vlan** *vid*	optional

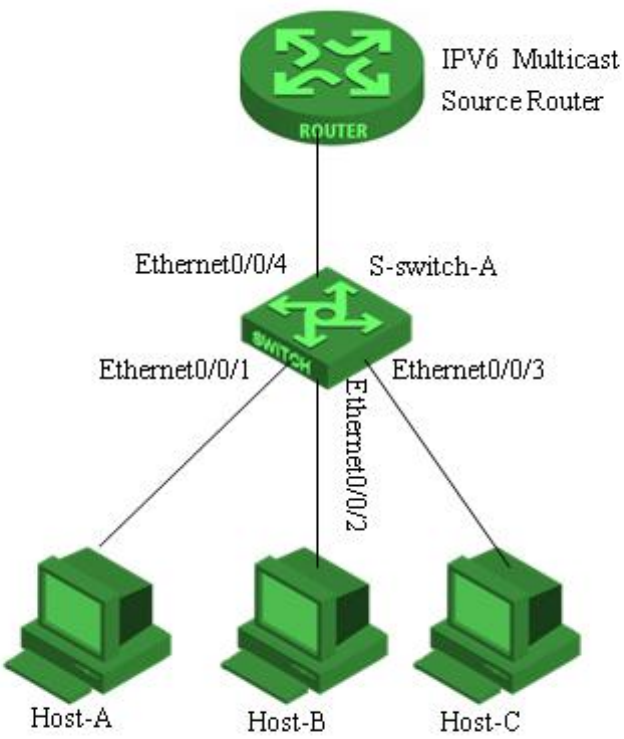
6.2.10 MLD Snooping Display and Maintenance

After you completed the above configurations, you can use the following command to view the configurations.

MLD Snooping Display and Maintenance

Operation	Command	Remarks
Display the MLD snooping-related configuration	**show mld-snooping**	All modes are executable
Display the dynamic routing port	**show mld-snooping router-dynamic**	
Display the routing port of static configuration	**show mld-snooping router-static**	
Display the multicast group	show multicast mld-snooping	

6.2.11 Configuration Example for MLD Snooping



1. Network requirements

As shown in Figure 1, Host-A, Host-B and Host-C hosts belong to VLAN 2, VLAN 3, and VLAN 4 respectively. The hosts are configured to receive the data of the multicast group with the address FF02::01::0101, FF02::01::0102 and FF02::01::0103 respectively.

2. Configuration steps

Configure S-switch-A

Configure VLAN 2, VLAN3 and VLAN4, and then add Ethernet 0/0/1, Ethernet 0/0/2 and Ethernet 0/0/3 to VLAN 2, VLAN 3, and VLAN 4 respectively.

S-switch-A(config)#vlan 2

S-switch-A(config-if-vlan)#switchport ethernet 0/0/1

S-switch-A(config-if-vlan)#exit

S-switch-A(config)#vlan 3

S-switch-A(config-if-vlan)#switchport ethernet 0/0/2

S-switch-A(config-if-vlan)#exit

S-switch-A(config)#vlan 4

S-switch-A(config-if-vlan)#switchport ethernet 0/0/3

S-switch-A(config-if-vlan)#exit

enable mld snooping

S-switch-A(config)#mld-snooping

When Host-A, Host-B and Host-C send mld report packets toS-switch-A, S-switch-A will learn the corresponding multicast group entries. When IPV6 Multicast Source Router sends mld query packets to S-switch-A, S-switch-A will learn the corresponding routing port entries.

Display the multicast groups learned by the S-switch-A

S-switch-A(config)#show mld-snooping group

show multicast table information

MAC Address : 33:33:00:01:00:01

VLAN ID : 2

port list : e0/0/1.

MAC Address : 33:33:00:01:00:02

VLAN ID : 3

port list : e0/0/2.

MAC Address : 33:33:00:01:00:03

VLAN ID : 4

port list : e0/0/2.

Total entries: 3 .

S-switch-A(config)#show mld-snooping router-dynamic

Port	VID	Age	Type
e0/0/4	2	284	{ QUERY }
e0/0/4	3	284	{ QUERY }
e0/0/4	4	284	{ QUERY }

Total Record: 3

When the Multicast Source Router sends the multicast data stream FF02::01::0101, FF02::01::0102 and FF02::01::0103, the S-switch-A will distribute the corresponding data stream to Host-A, Host-B and Host-C.

6.3 GMRP

6.3.1 GMRP Overview

GARP (GARP Multicast Registration Protocol) is an application of GARP (Generic Attribute Registration Protocol). It bases on the working mechanism of GARP and maintains the dynamic multicast registration information in the router. All routers that support GMRP feature can receive multicast registration information from other routers and dynamically update the local multicast registration information. At the same time, the router can also send local multicast registration information to other routers so that the multicast information of all GMRP-enabled devices in the same switching network can be consistent.

When a host wants to join an IP multicast group, it needs to send an IGMP join message, which is derived from the GMRP join message. When a host wants to join an IP multicast group, it needs to send an IGMP join message, which generates a GMRP join message. Upon receiving the GMRP join message, the Switch adds the port receiving the information to the appropriate multicast group. The Switch sends the GMRP join information to all the other hosts in the VLAN, with one host acting as the multicast source. When a multicast source sends multicast information, the Switch sends the multicast information only through the port that was previously added to the multicast group. In addition, the Switch periodically sends a GMRP query. If the host wants to stay in the multicast group, it will respond to the GMRP query. In this case, the Switch does not take any action.

If a host does not want to stay in a multicast group, it can either send a leave message or not respond to a periodic GMRP query. Once the Switch receives a leave message or does not receive a response during a leave all timer setting, it deletes the host from the multicast group.

6.3.2 Enable/disable GMRP

GMRP can be enabled in global configuration mode or in port configuration mode. By default, GMRP is disabled.

Enable GMRP

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enable GMRP in global mode	**gmrp**	required
Enter port configuration mode	interface ethernet *interface-num*	required
Enable GMRP in interface mode	**gmrp**	Required; if enable GMRP under port configuration mode, the port needs to be in trunk mode.
Disable GMRP	**no gmrp**	GMRP can be disabled in global configuration mode or in port configuration mode.

6.3.3 Configure the Multicast Released by GMRP

After GMRP is enabled, the system automatically propagates the multicast groups learned through GMRP, but if GMRP is required to propagate static multicast groups of local configuration, you need to perform the following configurations:

Configure the Multicast Released by GMRP

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure the multicast released by GMRP	(no) **garp permit multicast mac-address** *mac* **vlan** *vid*	Required; the corresponding static multicast should be created firstly; using the *no* command to delete the configuration;

Note: In GMRP, if it is not the default vlan, the corresponding vlan must be combined with gvrp to take effect. All of the features are generally used together with the gvrp. For the configuration of gvrp, see section 4.5；

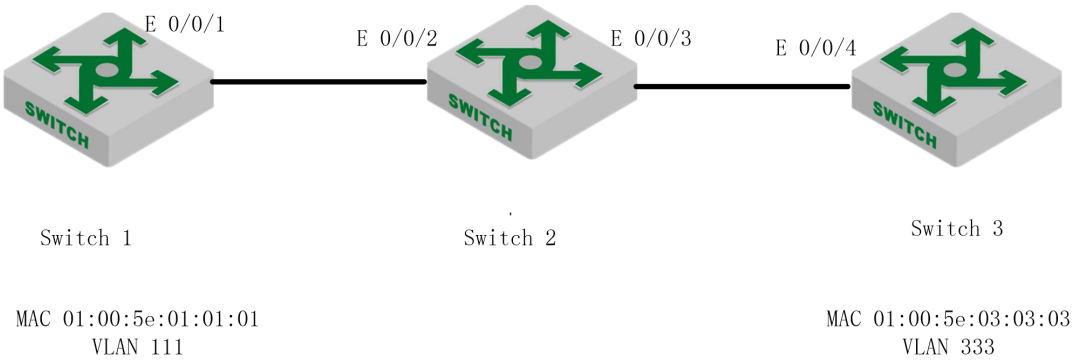
6.3.4 GMRP Display and Maintenance

After completing the above configuration, you can use the following command to view the configuration.

GMRP Display and Maintenance

Operation	Command	Remarks
Display the global GVRP status	**show gmrp**	All modes are executable
Display the status of GMRP	**Show gmrp interface [ethernet** *interface-num***]**	
Display GMRP	**show garp permit multicast**	
Display the local multicast groups (including static multicast groups and multicast groups learned via GMRP)	**show multicast**	

6.3.5 GMRP Configuration Example



Configuration Example for GMRP Service

1、Network requirements

In the network shown above, Switch 1 and Switch 3 advertise their static multicast information to Switch 2 via GMRP packets, and Switch 2 advertises the multicast information learned through GMRP. Eventually, the multicast information on Switch 1, Switch 2, and Switch 3 is synchronized.

2、Configuration Roadmap

1. Enable GMRP on Switch 1 and advertise the multicast information
2. Enable GMRP on Switch 2 and advertise the multicast information
3. Enable GMRP on Switch 3 and advertise the multicast information

3、Configuration steps

Switch1 Configuration

```
Switch1(config)#vlan 111,333
Switch1(config-if-vlan)#switchport ethernet 0/0/1 to ethernet 0/0/10
Add VLAN port successfully.
Switch1(config)#multicast mac-address 01:00:5e:01:01:01 vlan 111
adding multicast group successfully !
Switch1(config)#multicast mac-address 01:00:5e:01:01:01 vlan 111 interface ethernet 0/0/1 to ethernet 0/0/10
adding multicast group port successfully !
Switch1(config-if-vlan)#interface e 0/0/1
Switch1(config-if-ethernet-0/0/1)#switchport mode trunk
Switch1(config-if-ethernet-0/0/1)#exit
Switch1(config)#gvrp
Turn on GVRP successfully.
Switch1(config)#gmrp                    //Configure GMRP
Turn on GMRP successfully.
Switch1(config)#garp permit vlan 111,333
Switch1(config)#garp permit multicast mac-address 01:00:5e:01:01:01 vlan 111
Switch1(config)#interface e 0/0/1
Switch1(config-if-ethernet-0/0/1)#gvrp
Switch1(config-if-ethernet-0/0/1)#gmrp
Switch1(config-if-ethernet-0/0/1)#exit

 Switch1(config)#show gmrp          //Verify the GMRP configuration
GMRP status : enable
Switch1(config)#show gmrp interface ethernet 0/0/1
```

port GMRP status
e0/0/1 enable
Total entries: 1.
Switch1(config)#show garp permit multicast
GARP permit multicast:
vlan 111, mac 01:00:5e:01:01:01

Switch2 Configuration:
Switch2(config)#interface range ethernet 0/0/2 to ethernet 0/0/3
Switch2(config-if-range)#switchport mode trunk
Switch2(config-if-range)#exit
Switch2(config)#gvrp
Turn on GVRP successfully
Switch2(config)#gmrp //Configure GMRP
Turn on GMRP successfully.
Switch2(config)#interface range ethernet 0/0/2 to ethernet 0/0/3
Switch2(config-if-range)#gvrp
Switch2(config-if-range)#gmrp
Switch2(config-if-range)#exit

Switch2(config)#show gmrp //Verify the GMRP configuration
GMRP state : enable
Switch2(config)#show gmrp interface ethernet 0/0/2 ethernet 0/0/3
port GMRP status
e0/0/2 enable
e0/0/3 enable
Total entries: 2.

Switch3 Configuration
Switch3(config)#vlan 111,333
Switch3(config-if-vlan)#switchport ethernet 0/0/1 to ethernet 0/0/10
Add VLAN port successfully.
Switch3(config)#multicast mac-address 01:00:5e:03:03:03 vlan 333
adding multicast group successfully !
Switch3(config)#multicast mac-address 01:00:5e:03:03:03 vlan 333 interface ethernet 0/0/1 to ethernet 0/0/10
adding multicast group port successfully !
Switch3(config-if-vlan)#interface e 0/0/4
Switch3(config-if-ethernet-0/0/4)#switchport mode trunk
Switch3(config-if-ethernet-0/0/4)#exit
Switch3(config)#gvrp
Turn on GVRP successfully.
Switch3(config)#gmrp //Configure GMRP
Turn on GMRP successfully.
Switch3(config)#garp permit vlan 111,333
Switch3(config)#garp permit multicast mac-address 01:00:5e:03:03:03 vlan 333
Switch3(config)#interface e 0/0/4
Switch3(config-if-ethernet-0/0/4)#gvrp
Switch3(config-if-ethernet-0/0/4)#gmrp
Switch3(config-if-ethernet-0/0/4)#exit

Switch3(config)#show gmrp // //Verify the GMRP configuration
GMRP status : enable

Switch3(config)#show gmrp interface ethernet 0/0/4
port GMRP status
e0/0/4 enable
Total entries: 1.
Switch3(config)#show garp permit multicast
GARP permit multicast:
vlan 333, mac 01:00:5e:03:03:03

After the configuration is completed, you can use the *show multicast* command to view the multicast registration information learned by the GMRP function.

The multicast information on the Switch2 shows that 01: 00: 5e: 03: 03: 03 is the multicast learned through GMRP.

Switch1(config)#show multicast
show multicast table information
MAC Address : 01:00:5e:01:01:01
VLAN ID : 111
Static port list : e0/0/1-e0/0/10.
IGMP port list :
Dynamic port list :

MAC Address : 01:00:5e:03:03:03
VLAN ID : 333
Static port list :
IGMP port list :
Dynamic port list : e0/0/1.

Total entries: 2 .

The multicast information on the Switch2 shows that the 01: 00: 5e: 01: 01: 01 and 01: 00: 5e: 03: 03: 03 multicast messages are learned through GMRP.

Switch2(config)#show multicast
show multicast table information
MAC Address : 01:00:5e:01:01:01
VLAN ID : 111
Static port list :
IGMP port list :
Dynamic port list : e0/0/2.

MAC Address : 01:00:5e:03:03:03
VLAN ID : 333
Static port list :
IGMP port list :
Dynamic port list : e0/0/3.

Total entries: 2 .

The multicast information on the Switch3 shows that 01: 00: 5e: 01: 01: 01 are learned multicast packets via GMRP.

Switch3(config)#show multicast
show multicast table information
MAC Address : 01:00:5e:01:01:01
VLAN ID : 111
Static port list :
IGMP port list :

Dynamic port list : e0/0/4.

MAC Address : 01:00:5e:03:03:03
VLAN ID : 333
Static port list : e0/0/1-e0/0/10.
IGMP port list :
Dynamic port list :

Total entries: 2 .

6.4 Configure Static Multicast Table

6.4.1 Overview for Static Multicast Tables

In addition to dynamic learning, multicast tables can be manually configured, and a manually configured multicast table is a static multicast table. The static multicast MAC table will not be aged and it cannot be lost after being saved.

At present, only the corresponding multicast table of ipv4 can be static configured, and ipv6 multicast table cannot be static configured.

6.4.2 Create a Static Multicast Group

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Create a static multicast group	multicast {mac-address mac | ip-address ip } vlan vlan-id	required

The parameter mac refers to the mac address of the multicast group. It is required to use the multicast address format, for example: 01: 00: 5e: **: **: **, ip refers to multicast ip, for example, 224.0.1.1, vlan-id refers to VLAN ID, with the range of 1 to 4094. It must be an existed VLAN. When the static multicast group does not exist, the multicast group fails to be added.

For example:

！ Create a multicast group with the MAC address of 01: 00: 5e: 01: 02: 03 and the VLAN ID of 1

switch(config)#multicast mac-address 01:00:5e:01:02:03 vlan 1

！ Create a multicast group with the IP address of 224.0.1.1 and VLAN ID of 1

switch(config)#multicast ip-address 224.0.1.1 vlan 1

6.4.3 Add a Port to the Multicast Group

Add a Port to the Multicast Group

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Add member ports to a static multicast group	multicast {mac-address mac | ip-address ip } vlan vlan-id interface { all | interface-list }	required

For example:

！ Add the Ethernet ports 2, 3, 4 and 8 to the created multicast

switch(config)#multicast mac-address 01:00:5e:01:02:03 vlan 1 interface ethernet 0/0/2 to ethernet 0/0/4 ethernet 0/0/8

6.4.4 Configure the Proxy Port

When a Switch is configured with a static multicast table, if the Switch is configured with a proxy port, the Switch can send the multicast report to the multicast source to advertise the multicast member information.

Configure the Proxy Port

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Create a proxy port for a static multicast group	multicast {mac-address mac | ip-address ip } vlan vlan-id proxy-port ethernet interface-list	required
The interval at which the Switch sends report packets to the multicast source through the proxy port	multicas proxy-interval second	

6.5 IGMP Configuration

6.5.1 IGMP Overview

IGMP (Internet Group Management Protocol) is used to manage IP multicast group member as well as to establish and maintain the relationship between the IP host and multicast router.

Currently， there are three versions of IGMP: IGMPv1 (RFC 1112), IGMPv2 (RFC 2236) and IGMPv3 (RFC 3376). The most widely used is IGMPv2 version.

IGMPv1 defines two types of message: General Query and Group Membership Report. It manages the multicast group members based on query mechanism and response mechanism.

IGMPv2defines three types of message: Membership Query (including General Query and Group-Specific Query), Group Membership Report and Group Membership-Leave. Compared with IGMPV1, IGMPV12 added querier election mechanism and leave group mechanism.

IGMPv3 added source filter mechanism on the basis of v2, enhancing the function of query and report. Moreover, it presents the clear requirements to accept or reject the multicast message from some certain multicast source when the host adds certain multicast group.

All versions support ASM mode. Only IGMPv3 supports SSM mode. IGMPv1 and IGMPv2 can be able to apply to SSM mode under the help of IGMP SSM Mapping technology.

6.5.2 Enable Multicast Routing Protocol

You should enable multicast routing before the features of configuring IGMP protocol. Only if you enable the multicast protocol can relative configurations take effect.

Enable Multicast Routing Protocol

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable Multicast Routing Protocol	**ip multicast-routing**	Required. System disables multicast routing protocol.

6.5.3 Enable IGMP Protocol

Enable the IGMP protocol on interface to make Switch forward multicast message. Please perform the configuration under interface configuration mode (including VLAN interface and SuperVlan interface).

Enable IGMP Protocol

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	
Enable IGMP	**ip igmp**	required

Note:

You should enable multicast protocol before the interface enabling IGMP. Moreover, if it needs to work with PIM, you should configure the PIM protocol on this interface at the same time. Please refer to《PIM-DM/SM configuration manual》for more details.

6.5.4 Configure IGMP Version

Due to different IGMP version, different message construction and different types, it asks to configure the same IGMP version for all the routers in the same network segment. Or IGMP cannot be able to run normally. Please perform the configuration under interface configuration mode (including VLAN interface and SuperVlan interface).

Configure IGMP version

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	
Configure the interface to run IGMP version	**ip igmp version** *version-number*	required By default, the IGMP version is IGMPv2

6.5.5 Configure Static Multicast Group

Please perform the configuration under interface configuration mode (including VLAN interface and SuperVlan interface). When under the SuperVlan interface mode, you should specify the sub-VLAN.

Configure Static Multicast Group

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Add port into static multicast group	**ip igmp static-group** *groups-address* **{** *all* **|** *port-list* **} sourcelist {** * **|** *sourcelist* **}**	Required. By default, the port does not join any multicast group or multicast source in a static way.

6.5.6 Establish Static IP Multicast Table

Create a static IP multicast entry to realize the forwarding of multicast message. You can create (S, G) and (*, G) entries. If a static multicast member exists (created through the ip igmp static-group command), the port where the static member port is added is automatically added to the egress port of the corresponding static entry.

Establish Static IP Multicast Table

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Create static IP multicast table	**ip igmp create-group groups-address source { * | source-address }**	required There is no static multicast

		table by default.

Note:

1. This command is mainly used together with the command of ip igmp static-group.
2. Static multicast table can only be used for static port members.
3. As the ip igmp static-group command only creates static member ports for multicast groups, while this command creates a static multicast entry. If there is a static port member, it will automatically add the outgoing port to the corresponding static multicast entry. Similarly, when a static member port is created, if the corresponding static multicast entry exists, the outgoing port will be automatically added.

6.5.7 Configure Multicast Group Filter Function

The Switch sends IGMP query messages to confirm which multicast group contains the local group members directly connected to the Switch. If you do not want hosts on the network segment join certain multicast groups, you can configure ACL rules on the interface. The interface filters the received IGMP report messages according to the rule, and maintains the membership of the group only for the multicast groups allowed by the rule.

Configure Multicast Group Filter Function

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Configure filter function of multicast group	**ip igmp access-group access-list-number { all | ethernet port-number }**	Required By default, the host who under this interface can be able to add any legal multicast group.

6.5.8 Configure the Number of the Multicast Group Allowed Learning

It makes convenient and flexible for user to control the number of the multicast group allowed learning. If it exceeds the maximum quantity, Switch will not deal with the IGMP message.

Configure the Number of the Multicast Group Allowed Learning

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Configure the number of the multicast group allowed learning	**ip igmp limit-group** *num*	Required By default, the maximum multicast number is the maximum number of the multicast group allowed learning

Note:

1. This configuration only limits the number of dynamic multicast group, without limitation on the number of static multicast group.

2. If the IGMP group exceeds configuration value, the former IGMP group will not be deleted.

6.5.9 Configure IGMP General Query Interval

Switch forwards **Membership Query Message** periodically to check the existed multicast group. The interval is decided by Query Interval timer. Users can modify Query Interval of IGMP host via Query Interval timer.

Configure IGMP General Query Interval

Operation	Command	Remarks

Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Configure IGMP general query interval	**ip igmp query-interval** *seconds*	Required 125 seconds by default.

6.5.10 Configure IGMP Maximum Query Response Time

When a host receives a query from the Switch, it starts a delay timer for each multicast group it joins, using a random number (0, Max Response Time) as the initial value. Thereinto, the Max Response Time is the maximum response time specified by the query message. The maximum response time for IGMP Version 1 queries is Max(The maximum query response time for IGMP Version 1 is fixed at 10 seconds). The host should inform the Switch of the multicast group members before the timer timed out. If the Switch does not receive any group membership report after the maximum query response time times out, the Switch considers that it has no local group members and it will no longer transmit the received multicast packets to the network which it connected.

Configure IGMP Maximum Query Response Time

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Configure the maximum query response time of IGMP	**ip igmp query-max-response-time** *seconds*	Required By default, the maximum response time is 10 seconds

Note:

1. You can use this command only when IGMP V2 / V3 is running.

2. This command controls the interval for the host to respond to host member queries. The time interval is small, which enables the Switch to quickly learn the status of group members. If the host does not respond quickly to host member queries, it may be removed from the multicast group even if the user does not wish to delete them. Therefore, the user should ensure that the set time interval is greater than the host shortest response time.

6.5.11 Configure Last-Member-Query-Interval

After receiving leave-message, switch will forward specified group query message to know whether there are other group members in multicast group. User can be able to modify the interval value of specified group query message.

Configure Last-Member-Query-Interval

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Configure last-member-query-interval	**ip igmp last-member-query-interval** *seconds*	Required 1 second by default.

Note:

1. Only if IGMP V2/V3 is running can this command take effect.

2. Last-Member-Query is used to check how many multicast member are there in the network, so the interval should not be too long or it will lose corresponding meaning.

6.5.12 Configure Robustness Variable of IGMP Querier

Robustness variable is a very important parameter to reflect IGMP protocol performance, mainly applied to control message forwarding frequency so as to enhance the robustness of network protocol operation. In addition, robustness variable coefficient is a very important parameter to calculate other variables.

Configure Robustness Variable of IGMP Querier

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Configure robustness variable of IGMP querier	**ip igmp robustness-varible** *num*	Required 2 by default.

6.5.13 Configure IGMP Proxy

After enabling IGMP proxy, Switch acts as a host forwards the multicast group information via report message. When the multicast router receives the message, it transmits the multicast traffic to Switch and then Switch will transmit the multicast traffic to the downlink user. If a certain multicast has no host, Switch will forward leave message to multicast routing, and then multicast routing will stop forwarding multicast data to Switch. This function is mainly apply to network peripheral Switches, effectively save Switch resources because Switches can complete the multicast forwarding without enabling the multicast routing protocols.

Configure IGMP Proxy

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface vlan-interface** *vlan-id*	Currently, it does not support to configure on super interface.
Configure IGMP proxy	**igmp-proxy**	Required Disabled by default.

6.5.14 Configure IGMP SSM Mapping

In the SSM network, some recipient hosts only run IGMPv1 or IGMPv2 due to the variety of possible restrictions. You can configure the IGMP SSM Mapping function in router so as to offer SSM service to those recipient hosts of IGMPv1 or IGMPv2.

Enable SSM-Mapping

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface}** *vlan-id*	
Enable SSM -Mapping	**ip igmp ssm-mapping**	Required; Disabled by default.
Enter IGMP global configuration mode	**mroute igmp**	
Configure the SSM-Mapping static group address mapping rule	**ssm-mapping static access-control-list source-address**	Required; By default, no static group address mapping rule is configured.

Note:

1. You should enable SSM Mapping before configuring the mapping rules of SSM source/ group address, otherwise IGMP does not support SSM Mapping function.

2. When SSM Mapping router receives the report message of IGMPv1/v2, it will obtain the source address S via group address G and then form the（S,G）channel.

3. SSM mapping only needs to be enabled on the device connected to the receiving host.

4. To ensure that hosts on any IGMP version can receive SSM services, it is recommended to run IGMPv3 on the interface on the network segment.

5. The mapping from the same multicast group to multiple multicast sources in specified SSM multicast group can be realized via multiple configurations. Because static SSM mapping can be configured multiple times, the source-address parameter of multiple ACLs can be used as the mapping source for group G if group G belongs to multiple Permit entries of multiple ACLs at the same time. The maximum number of static SSM mappings can be set to 8.

SSM-Mapping static group address mapping rule

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter IGMP and then enter global configuration mode	**mroute igmp**	
Configure the SSM-Mapping static group address mapping rule	**ssm-mapping static {** *access-control-list source-address* **}**	Required By default, no static group address mapping rule is configured.

6.5.15 IGMP Display and Maintenance

After completing the above configuration, can use the following command to display configuration.

IGMP display and maintenance

Operation	Command	Remarks
Display IGMP interface information	**show ip igmp interface [{ vlan-interface** *vid* **} | { supervlan-interface** *number* **}]**	-
Display static configuration and the IGMP multicast group information	**show ip igmp groups {** dynamic |static | *multicast-ip* **}**	
Display IGMP proxy	**show igmp-proxy**	
Display SSM-Mapping mapping rule	**show ip igmp ssm-mapping [** *multicast-ip* **]**	
Debug IP IGMP	**debug ip igmp**	

6.5.16 Configuration Example for IGMP Basic Function

1. Network requirement and network construction

1） User-PC receives VOD information via multicast；

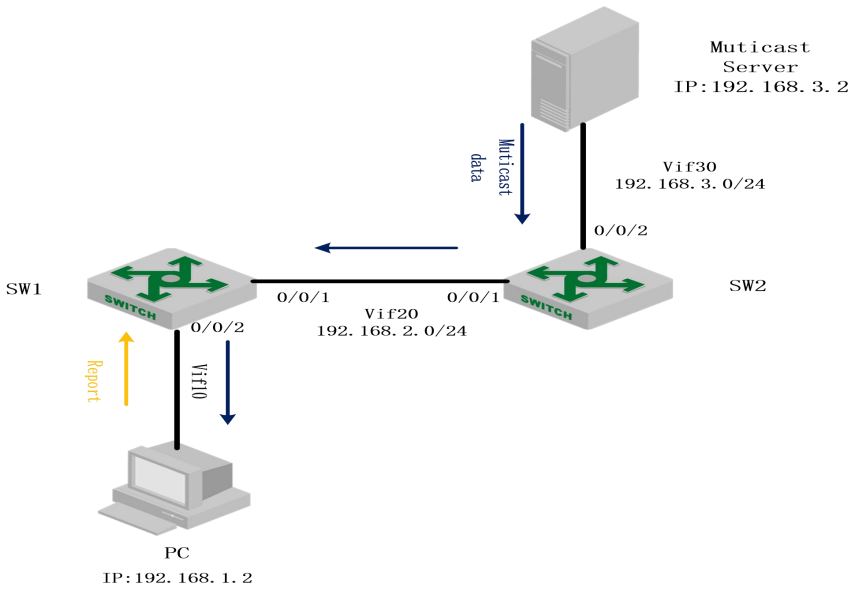
2） User-PC adopts VLC Media Player as multicast receiver client; Multicast server adopts VLC Media Player as multicast source to offer multicast video services, too；

3） Run IGMPv2 protocol between SW1 and SW2；

4） Through the configuration, PC can only be added into 224.1.1.1 to view the video of 224.1.1.1.

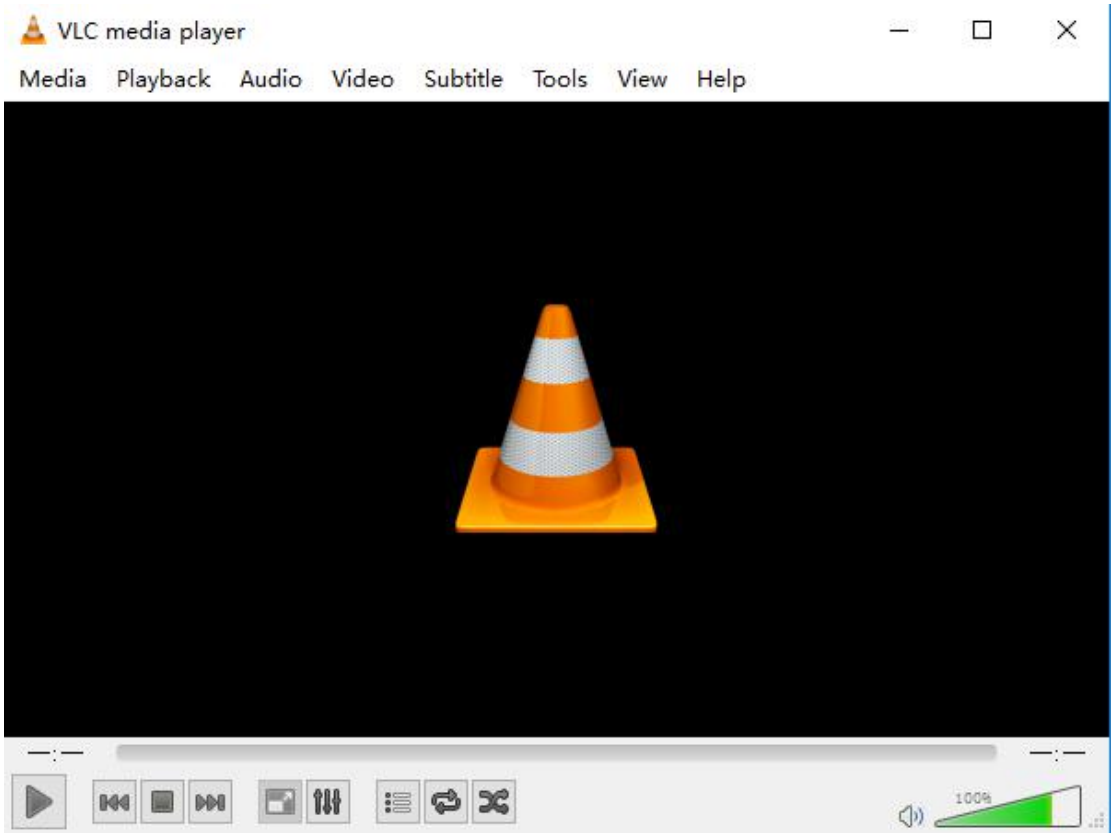
Network diagram is shown as follow:

configurations example of IGMP basic function



2.Configuration steps

1) install VLC media player on both user PC and multicast server



2) Relative configurations of SW1 and SW2. This configuration adopts PIM-DM protocol. Please refer to <PIM configuration> for detailed information.

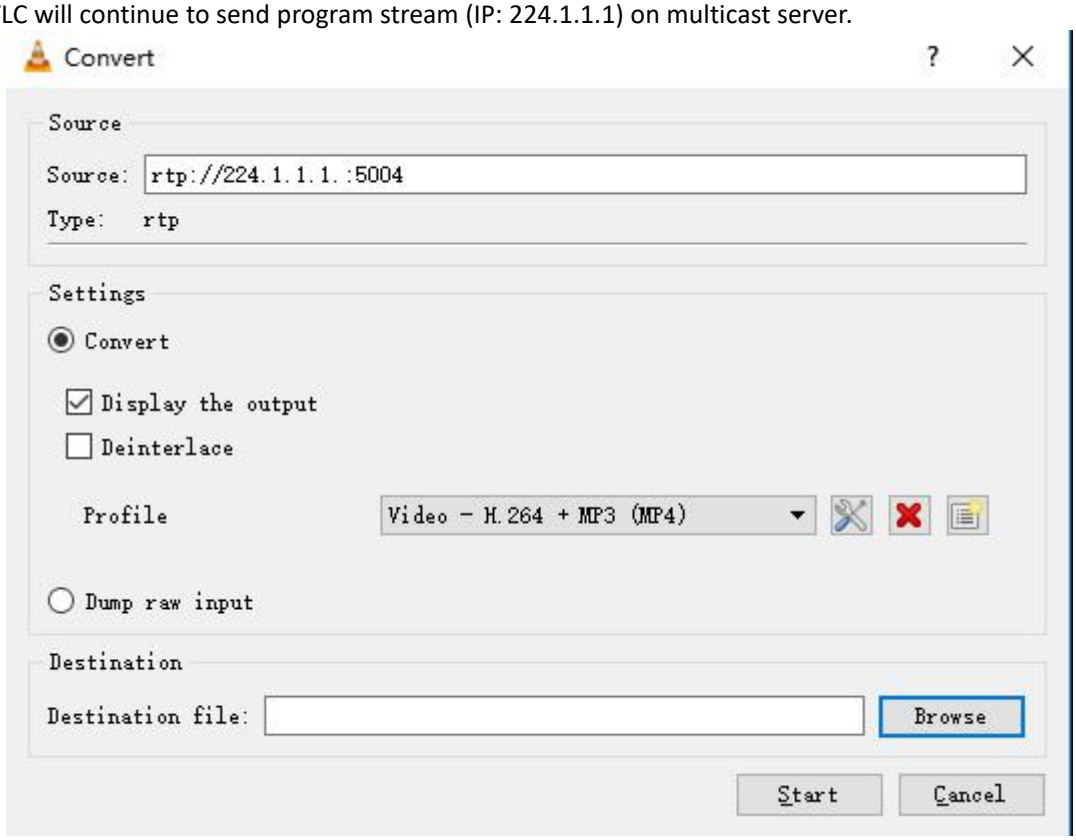
#SW1 configuration
SW1(config)#vlan 20
SW1(config-if-vlan)#vlan 10
SW1(config-if-vlan)#interface ethernet 0/0/1
SW1(config-if-ethernet-0/0/1)#switchport default vlan 20
SW1(config-if-ethernet-0/0/1)#interface ethernet 0/0/2
SW1(config-if-ethernet-0/0/2)#switchport default vlan 10
SW1(config-if-ethernet-0/0/2)#interface vlan-interface 10
SW1(config-if-vlanInterface-10)#ip address 192.168.1.1 255.255.255.0
SW1(config-if-vlanInterface-10)#interface vlan-interface 20
SW1(config-if-vlanInterface-20)#ip address 192.168.2.1 255.255.255.0
SW1(config-if-vlanInterface-20)#exit
SW1(config)#ip multicast-routing //enable multicast routing protocol
SW1(config)#interface vlan-interface 10
SW1(config-if-vlanInterface-10)#ip igmp // enable interface IGMP
SW1(config-if-vlanInterface-10)#ip pim dense-mode
SW1(config-if-vlanInterface-10)#interface vlan-interface 20
SW1(config-if-vlanInterface-20)#ip pim dense-mode // run interface PIM-DM multicast routing protocol
SW1(config)#access-list 99 permit 22.4.1.1 0
SW1(config)#access-list 99 deny any
SW1(config)#interface vlan-interface 10
SW1(config-if-vlanInterface-10)#ip igmp access-group 99 ethernet 0/0/2 //configure multicast group filter
SW1(config)#router ospf
SW1(config-router-ospf)#network 192.168.1.1 0.0.0.255 area 0

SW1(config-router-ospf)#network 192.168.2.1 0.0.0.255 area 0
SW1(config-router-ospf)#

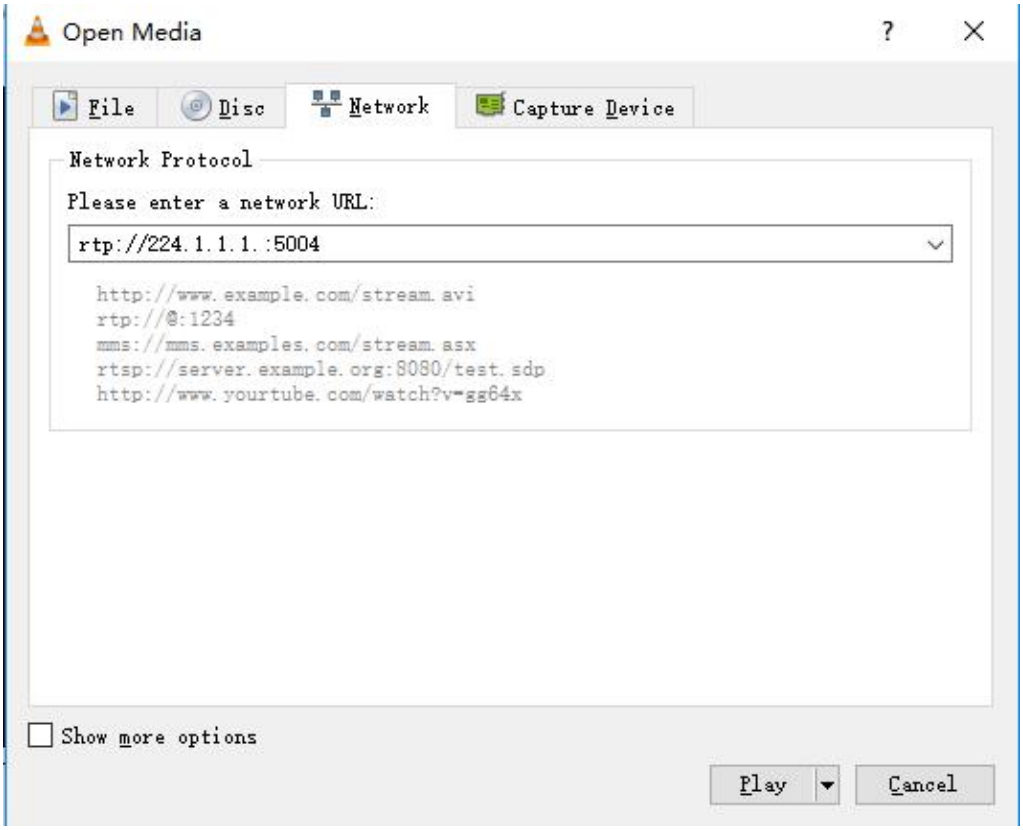
#SW2 configuration
SW2(config)#vlan 20
SW2(config-if-vlan)#vlan 20
SW2(config-if-vlan)#vlan 30
SW2(config-if-vlan)#interface ethernet 0/0/1
SW2(config-if-ethernet-0/0/4)#switchport default vlan 20
SW2(config-if-ethernet-0/0/4)#interface ethernet 0/0/2
SW2(config-if-ethernet-0/0/2)#switchport default vlan 30
SW2(config-if-ethernet-0/0/2)#interface vlan-interface 20
SW2(config-if-vlanInterface-20)#ip address 192.168.2.2 255.255.255.0
SW2(config-if-vlanInterface-20)#interface vlan-interface 30
SW2(config-if-vlanInterface-30)#ip address 192.168.3.1 255.255.255.0
SW2(config-if-vlanInterface-30)#exit
SW2(config)#ip multicast-routing
SW2(config)#interface vlan-interface 20
SW2(config-if-vlanInterface-20)#ip pim dense-mode
SW2(config-if-vlanInterface-20)#exit
SW2(config)#interface vlan-interface 30
SW2(config-if-vlanInterface-30)#ip pim dense-mode
SW2(config)#router ospf
SW2(config-router-ospf)#network 192.168.2.2 0.0.0.255 area 0
SW2(config-router-ospf)#network 192.168.3.1 0.0.0.255 area 0

3) Result validation

VLC will continue to send program stream (IP: 224.1.1.1) on multicast server.



By this time, PC can be able to receive 224.1.1.1 program via VLC client side.



Display the learning multicast group from 224.1.1.1 on SW1:

SW1(config)#show ip igmp groups
IGMP Connected Group Membership

Group Address: 224.1.1.1
Vlan: 10, port: 0/0/2, Uptime: 00:00:07
Expires: 00:04:13, Last Reporter: 192.168.1.2
V1 Expires: 00:00:00, V2 Expires: 00:04:13, Self: False
FilterMode: EXCLUDE, Static: False
SourceList(0):
Current State(IGMP_MS_NORMAL2)

Total Groups: 1, Total group members: 1

The command of “show ip mroute” can be able to display SW1 and SW2 forwarding items of multicast routing:

SW1(config)#show ip mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
J - Join SPT, O - SRM originator, H - Hardware switched
K - Static entry

Timers: Uptime/Expires, Interface state: State/Mode

(192.168.3.2, 224.1.1.1), 00:05:04/00:03:21, flags: DCTH
  Incoming interface: VLAN-IF20, RPF nbr: 192.168.2.2
  Outgoing interface list:
VLAN-IF10, 0/0/2, Forward/Dense, 00:05:04/stopped

Total dynamic entries 1. Total static entries 0.

Total ip multicast entries 1.

SW2(config)#show ip mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
       P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
       J - Join SPT, O - SRM originator, H - Hardware switched
       K - Static entry
Timers: Uptime/Expires, Interface state: State/Mode

6.6 PIM Configuration

PIM is short for Protocol Independent Multicast. PIM means multicast router has nothing to do with the unicast routing protocol as long as the unicast router protocol can be able to perform RPF check on multicast message and generate corresponding multicast routing table.

According to different realization mechanisms, PIM is divided into PIM -DM （Protocol Independent Multicast-Dense Mode, PIM -SM（Protocol Independent Multicast-Sparse Mode and PIM-SSM（Protocol Independent Multicast Source-Specific Multicast.

Several important concepts in IP multicast

- Multicast Distribution Tree: it is the path which transmitting to the receiver after IP multicast data generated by the source, and this path is just like a bifurcate tree.
- SPT (Shortest Path Tree): it is the shortest path which transmitting from the source to the receiver.
- RPT （Rendezvous Point Tree）: Rendezvous Point is the root of the shared tree, datagram is sent from source to RP and forwarded along the shared tree, and the corresponding tree is called RPT.
- RPFC（ Reversed Path Forwarding Check）: The unicast Reverse Path Forwarding (RPF) check ensures that an IP packet that enters a router uses the correct inbound interface. This feature supports unicast RPF check on the spoke-side interfaces. Because different VRFs are used for downstream and upstream forwarding, the RPF mechanism ensures that source address checks occur in the downstream VRF.
- Multicast forwarding-table: Similar to unicast routing table, applying to record the state of distribution tree and guide the transmitting of multicast data.

6.6.1 PIM -DM Overview

Protocol Independent Multicast-Dense Mode (PIM -DM) is a dense-mode multicast routing protocol, which is applicable to small-sized networks and using “Push mode” to transmit multicast data. In a PIM -DM network, members of a multicast group are densely distributed.

6.6.2 PIM -DM Working Mechanism

The above PIM - DM process can be summarized as: neighbor discovery mechanism, flooding mechanism & pruning mechanism, grafting mechanism. In addition, it ensures the normal work of PIM -DM via assertion mechanism and state refresh mechanism.

- Neighbor discovery

Upon startup, a PIM-DM router needs to discover neighbors by sending Hello packets. The relationships between PIM-DM capable network nodes are maintained through exchange of Hello packets. In PIM-DM, Hello packets are sent periodically.

- Flooding mechanism & Pruning mechanism

PIM-DM assumes that all the hosts on a network are ready to receive multicast data. A packet is transmitted from multicast source S to multicast group G. After receiving this multicast packet, the router performs an RPF check based on the unicast routing table and creates an (S, G) entry if the RPF check is successful. Then the router floods the packet to all the downstream PIM-DM nodes in the network. The router discards the packet if the RPF check fails (the multicast packet is from an incorrect interface). In the flooding process, an (S, G) entry will be created in the PIM-DM multicast domain.

If no downstream node is a multicast group member, the router sends a Prune message to notify the upstream node that data should not be sent to downstream nodes any more. After receiving the Prune message, the upstream node removes the interface that sends the multicast packet from the outbound interface list matching the (S, G) entry. Eventually, a Shortest Path Tree (SPT) with S as the root is created. The prune process is initiated by a leaf router.

The whole process is called the flooding&prune process. A timeout mechanism is made available on a pruned router so that the router may initiate a flooding&prune process again if the prune process times out. The flooding&prune mechanism of PIM-DM operates periodically over and over again.

In the flooding&prune process, PIM-DM performs RPF check and builds a multicast forwarding tree with the data source as the root based on the current unicast routing tables. When a multicast packet arrives, the router first judges whether the path of the multicast packet is correct. If the interface where the packet arrives is what specified in the unicast route, the path is considered correct. Otherwise, the multicast packet is discarded as a redundant packet and will not be forwarded in multicast mode. The unicast route may be discovered by any unicast routing protocol such as RIP and OSPF instead of a specific routing protocol.

- Grafting mechanism

When the pruned downstream node needs to enter the forwarding state again, it sends a Graft message to the upstream node. Before configuring the features of IGMP, you must enable the multicast routing function.

- Assertion mechanism

As shown in the following figure, multicast routers A and B are on the same LAN segment and they have their respective paths to multicast source S. After receiving a multicast packet

from S, both of them will forward the packet on the LAN. As a result, the downstream multicast router C will receive two identical multicast packets.

An upstream router uses the Assert mechanism to select the only forwarder. The upstream router sends Assert messages to select the best route. If two or more paths have the same priority and metric value, the router with the largest IP address is selected as the upstream neighbor of the (S,G) entry and is responsible for forwarding the (S,G) multicast packet.

- State Refresh Mechanism

To avoid repeated flooding&prune actions, the State refresh mechanism is added to new protocol standards. The router in direct connection with the multicast source sends state update packets periodically. After receiving a state update packet, the PIM-capable router refreshes the prune state.

6.6.3 PIM -SM Overview

PIM -SM is short for protocol independent multicast-sparse mode, using pull mode to transmit the multicast data. It is usually applied to the network that the multicast group distribution is relatively scattered or a wide range of large and medium-sized network.

6.6.4 PIM-SM Working Mechanism

The operation of Protocol Independent Multicast-Sparse Mode (PIM-SM) can be understood as neighbor discovery, DR selection, RP selection, rendezvous point tree (RPT) generation, multicast source registration, and SPT switch. The neighbor discovery of PIM-SM is the same as that of PIM-DM.

- DR election

You can also elect a DR for a shared network (such as Ethernet) through a Hello packet. The DR serves will act as the only forwarder of multicast data in the shared network.

The network connected to the multicast source or the network connected to the receiver needs to elect the DR, and the DR at the receiver side sends the join message to the RP.

- RP election

RP is a core device in a PIM-SM domain and aggregates the join / prune requests of multicast receivers and the multicast data of a multicast source. RPs can be specified statically or multiple C-RPs (Candidate-RPs) can be configured in the PIM-SM domain to dynamically elect the RPs through the bootstrap mechanism so that different RPs can serve different multicast groups.

The BSR (Bootstrap Router) collects the Advertisement Message (C-RP) from the C-RP, which carries the address and priority of the C-RP and the group range of the C-RP. This information is

summarized as RP-Set encapsulation in the bootstrap message spread to the entire network. The routers in the network use the same rules to select their corresponding RPs from specific C-RPs according to the information provided by the RP-Set.

- Build RPT

When the host joins a multicast group G, the leaf router directly connected to the host learns the receivers of the multicast group G through the IGMP message, calculates the corresponding convergence point RP for the multicast group G. And then sends a join message to the upper-level node in the RP direction (join message). Each router from the leaf router to the RP will generate a (*, G) entry in the forwarding table, indicating that no matter which source sends it to the multicast group G, it applies to the entry. When the RP receives a packet destined for multicast group G, the packet arrives at the leaf router along the established path to reach the host. This generates an RPT rooted at RP.

- Multicast source registration

When the multicast source S sends a multicast packet to multicast group G, the PIM-SM multicast router directly connected to S is responsible for encapsulating the received multicast packet into a register packet and then forwarded the packet to the corresponding RP in unicast mode. If there are multiple PIM-SM multicast routers on a network segment, the multicast router sends the multicast packet to the Designated Router (DR).

When the RP receives the packet, it de-encapsulates the register message and forwards the encapsulated multicast packet along the RPT to the receiver. On the other hand, the RP sends the multicast packet to the multicast source hop-by-hop (S, G) Message. In this way, the routers from the RP to the multicast source form the SPT, and these routers generate the (S, G) entries in their forwarding tables. SPT takes multicast source as root and RP as leaf.

The multicast data sent by the multicast source reaches the RP along the established SPT, and then the RP forwards the multicast data along the RPT to the receivers. When the RP receives the multicast data forwarded along the SPT, it sends a Register-Stop Message to the DR through unicast. The multicast source registration process ends.

- RPT switch to SPT

When the DR at the receiver side finds that the multicast data rate from the RP to the multicast group G exceeds a certain threshold, it initiates the switch from the RPT to the SPT. The process is as follows:

(1) First, the receiver-side DR sends the (S, G) join message hop by hop to the multicast source S, and then sends the packet to the multicast source-side DR. All the routers along the route generate the (S, G)), thus establishing the SPT branch；

(2) Then, the receiver-side DR sends a prune message containing RP bits hop-by-hop to the RP. Upon receiving this packet, the RP sends a prune message to the multicast source (assuming that there is only this receiver) ,thus ultimately switching from RPT to SPT.

After the RPT is switched to the SPT, the multicast data is sent directly from the multicast source to the receiver. By switching from RPT to SPT, PIM-SM can establish SPT in a more cost-effective manner than PIM-DM.

6.6.5 PIM-SSM Overview

PIM SSM is short for Protocol Independent Multicast ---- Source Specific Multicast. Generally speaking, IP multicast and SSM can coexist on a single router. What is more, both of them can be realized by using PIM - SM protocol. PIM SSM should be used in concert with IGMPv3.

Usually IGMPv3 is deployed on the host to establish and maintain multicast group memberships. Compared with IGMPv2, IGMPv3 is designed with the source-based filtering function. This function allows a host to receive only the data from a specific group and even from a specific source in this group. Based on a received IS_IN packet of IGMPv3, the SSM-enabled router learns that a host on the network connected with the interface receiving the IS_IN packet wants to receive (S, G) packets. This router unicasts a PIM (S,G) Join message to the next-hop router of the multicast source hop by hop and thereby an SPT can be established between the multicast source and the last-hop router. When the multicast source is sending multicast data, the data reaches the receiver along the SPT.

If a host supports only IGMPv1/IGMPv2, you can configure SSM mapping on the router connected with the host to convert the (*, G) Join messages of IGMPv1/IGMPv2 into (S, G) Join messages.

6.6.6 Enable Multicast Routing

You should enable multicast routing before configure PIM -DM protocol. Only if you enable multicast protocol can relative configurations take effect.

Enable PIM -DM protocol

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable multicast routing	**ip multicast-routing**	Required. System disables the multicast routing protocol by default.

6.6.7 Enable PIM -DM Protocol

PIM -DM protocol needs to be started respectively on each interface. After configuring PIM -DM on interface, PIM -DM will regularly send Hello message of PIM protocol. And it will handle the protocol message sent by PIM neighbor.

Enable PIM -DM Protocol

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {[vlan-interface| supervlan-interface]}** *vlan-id*	
Enable PIM -DM protocol	**ip pim dense-mode**	required

6.6.8 PIM –DM Advanced Configuration

6.6.9 Configure the Transmission Interval of Hello Packets

After starting PIM protocol on interface, it will regularly send Hello message. You can make suitable revises on the interval of Hello message according to the network bandwidth and types.

Configure the Transmission Interval of Hello Packets

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {[vlan-interface| supervlan-interface]}** *vlan-id*	
Configure the transmission interval of hello packets.	**ip pim query-interval** *seconds*	Optional 30S by default.

6.6.10 Configure PIM Neighbor Filtering

You can configure the PIM neighbor filtering via basic access control list.

Configure PIM Neighbor Filtering

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure ACL	**access-list** *access-list-number* permit *a.b.c.d 0*	used to match pim neighbor
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	
Configure PIM neighbor filtering	**ip pim neighbor-poliy** *access-list-number*	optional

6.6.11 Configure the Maximum PIM Neighbors for an Interface

Vast PIM neighbor relations will drain router memory and then lead to router fault. In this case, you can limit on PIM neighbor number of router interface so the neighbors cannot be added in if the PIM routing amount exceeds the limits. Moreover, the total number of router PIM neighbor is limited by system so users can not modify it via the command.

Configure the Maximum PIM Neighbors for an Interface

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	
Configure the Maximum PIM Neighbors for an Interface.	**ip pim neighbor-limit** *limit*	optional

6.6.12 Configure Multicast Source (Group)-Based Filtering

You can filter according to the source address of the multicast data so as to improve the security of the network.

Configure Multicast Source (Group)-Based Filtering

Operation	Command	Remarks
Enter global	**configure terminal**	-

configuration mode		
Configure ACL	**access-list** *access-list-number* permit *a.b.c.d 0*	used to match pim neighbor
Enter PIM configuration mode	**mroute pim**	
Configure multicast source filter	**source-policy** *access-list-number*	optional

6.6.13 PIM -DM Display and Maintenance

PIM -DM display and maintenance

Operation	Command	Remarks
Display the information of PIM interface	**show ip pim interface [vlan-interface** *vid* **]**	-
Display the information of PIM neighbor	**show ip pim neighbor**	
Display PIM multicast routing table	**show ip mroute{** *group-address* **static | dynamic** }	
Debug PIM	**debug pim**	

6.6.14 Enable PIM -SM Protocol

PIM -DM protocol needs to be started respectively on each interface. After configuring PIM -SM on interface, PIM -SM will regularly send Hello message of PIM protocol. And it will handle the protocol message sent by PIM neighbor.

Enable PIM -SM Protocol

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {[vlan-interface| supervlan-interface]}** *vlan-id*	
Enable PIM -SM Protocol	**ip pim sparse-mode**	required

6.6.15 Configure Static RP

The entire network only relies on a RP to perform multicast forwarding information. You can specify routing RP location in the PIM -SM domain so as to improve the robustness of the network.

Configure Static RP

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter PIM configuration mode	**mroute pim**	
Configure Static RP	**static-rp** *address*	optional

Note:

If adopts static RP, all the routers in the PIM domain should use the same configuration. If the static RP address is the interface address of some “up” state interface, the host will be static RP. The static RP interface does not need to enable PIM protocol.

If the RP which selected in BSR mechanism valid, the static RP takes no effect; if it fails to obtain dynamic RP, the static RP will take effect.

6.6.16 Specify a Candidate BSR

In a PIM -SM domain，there should be one unique BSR to support the regular work of PIM -SM network devices (such as routers, Ethernet Switch, etc.). BSR is responsible for collecting RP information and publish the RP information. The unique BSR is selected from multiple C-BSR via bootstrap message. Before receiving BSR message, C-BSR regards himself as BSR, and it will regularly broadcast bootstrap message in PIM -SM domain. Bootstrap message includes C-BSR address and C-BSR priority, and PIM selects the BSR from C-BSR address and C-BSR priority. The foundation that C-BSR becomes BSR is: the C-BSR who has the highest priority will become BSR; if all C-BSR has one same priority, the C- BSR who has biggest IP address will become BSR.

Specify a Candidate BSR

Operation	Command	Remarks

Enter global configuration mode	**configure terminal**	-
Enter PIM configure mode	**mroute pim**	
Specify a Candidate BSR	**bsr-candidate** *interface-type interface-number hash-mask-length* **[** *priority* **]**	optional

Note:

Usually, there is only one C - BSR and one C-RP in the network configuration. In a general way, the Switch or the router is the networking core.

6.6.17 Configure the Candidate RP

After the election of the BSR, all C-RPs send C-RP Advertisements to BSR periodically. The BSR aggregates and advertises the RP information to the entire network. (There may be multiple RPs in the network. They each have different groups services), so that all the Ethernet Switches can gain RP information.

When configuring C-RP, you can specify the scope of the RP service, which can serve all multicast groups or only certain multicast groups.

Configure the candidate RP

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter PIM configuration mode	**mroute pim**	
Configure the candidate RP	**rp-candidate** *interface-type interface-number* **group-list** *acl-number priority*	optional

6.6.18 Configure BSR Border

Configure the Switch interface as the BSR domain boundary of PIM. All bootstrap messages cannot pass through the domain boundary when the PIM domain boundary is set on this interface. However, other PIM messages can pass through this domain boundary. This allows the users to effectively split the network that runs PIM-SM into multiple domains, each with a different Bootstrap Router.

Configure BSR Border

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**Interface {[vlan-interface| supervlan-interface]}** *vlan-id*	
Configure BSR Border	**ip pim bsr-border**	optional

6.6.19 Configure the SPT Switching Threshold

In PIM-SM mode, receiving host commonly take the initiative to join the RP and then obtain the multicast packet via RP. Generally speaking, the path in RPT is not the shortest path from receiving host to the multicast source. In this case, DR optional will join SPT so as to avoid packet broadcast delay. Currently, it supports two fixed threshold, **immediately** and **infinity.**

Configure the SPT Switching Threshold

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter PIM configuration mode	**mroute pim**	
Configure the SPT switching threshold	**spt-threshold {** *immediately* **|** *infinity* **}**	optional，*"immediately"* is default-value

6.6.20 Configure the Range of an SSM Multicast Group

PIM-SSM is used as a subset of PIM-SM, and PIM-SSM or PIM-SM model is used to transfer the information from the multicast source to the receiver, depending on whether the receiver subscribes channel (S, G) is in the range of SSM multicast group (232.0.0.0/8). All the interfaces enabled with PIM-SM will regard the multicast group in this range adopted PIM-SSM model.

Configure the Range of an SSM Multicast Group

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter PIM configuration mode	**mroute pim**	
Configure the Range of an SSM Multicast Group	**ssm {default | range** *access-list***}**	required

6.6.21 PIM -SM Display and Maintenance

PIM -SM Display and Maintenance

Operation	Command	Remarks
Display the information of PIM interface	**show ip pim interface [vlan-interface** *vid* **]**	-
Display the information of PIM neighbor	**show ip pim neighbor**	
Display multicast routing of PIM	**show ip mroute** *group-address* **[** *static* **|** *dynamic* **]**	
Display PIM RP information	**show ip pim rp-info** *group-address*	
Display the information of BSR	**show ip pim bsr**	
Display address range of SSM group	**show ip pim ssm range**	
Debug PIM	**debug pim**	

6.6.22 PIM Configuration Examples

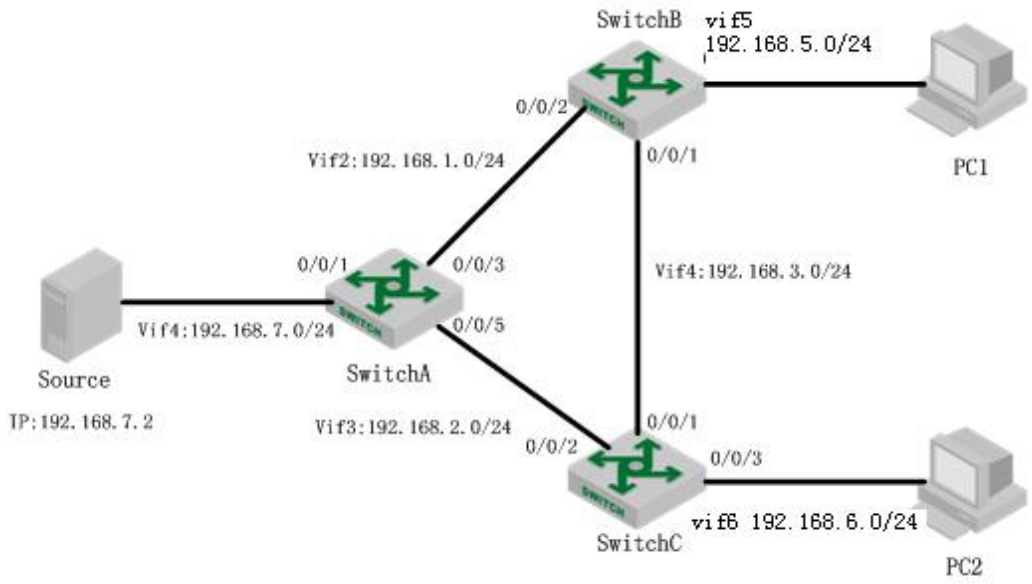
一、Requirement and Networking

Layer 3 Switch A, Switch B and Switch C adopt OSPF protocol to perform interconnection. Receiver receives VOD information via multicast. User-pc1 and User-PC2 receive different video information.

In this network, it adopts PIM -SM protocol to realize the allocation of multicast data. In addition, it dynamically elects the RP via bootstrapping mechanism. It configures the static RP to avoid communication outage due to the dynamic RP faults.

Network diagram is shown as follow:

Network Diagram of PIM –SM Configuration



二、Configuration steps

the configuration of switch A
(1) Enable multicast routing
SwitchA(config)#ip multicast-routing
(2) Configure each interface as well as the interface address, and then enable PIM -SM.

SwitchA(config)#vlan 2-4
SwitchA(config-if-vlan)#interface ethernet 0/0/1
SwitchA(config-if-ethernet-0/0/1)#switchport default vlan 4
SwitchA(config-if-ethernet-0/0/1)#interface ethernet 0/0/5
SwitchA(config-if-ethernet-0/0/5)#switchport default vlan 3
SwitchA(config-if-ethernet-0/0/5)#interface ethernet 0/0/3
SwitchA(config-if-ethernet-0/0/3)#switchport default vlan 2
SwitchA(config-if-ethernet-0/0/3)#exit
SwitchA(config)#interface vlan-interface 4

SwitchA(config-if-vlanInterface-4)#ip address 192.168.7.1 255.255.255.0
SwitchA(config-if-vlanInterface-4)#ip pim sparse-mode
SwitchA(config-if-vlanInterface-4)#exit
SwitchA(config)#interface vlan-interface 2
SwitchA(config-if-vlanInterface-2)#ip address 192.168.1.1 255.255.255.0
SwitchA(config-if-vlanInterface-2)#ip pim sparse-mode
SwitchA(config-if-vlanInterface-2)#exit
SwitchA(config)#interface vlan-interface 3
SwitchA(config-if-vlanInterface-3)#ip address 192.168.2.1 255.255.255.0
SwitchA(config-if-vlanInterface-3)# ip pim sparse-mode
SwitchA(config-if-vlanInterface-3)#exit

(3) Configure C-RP and static RP. Specify the IP address of 2 VLAN interface to be C-BSR and C-RP; specify the IP address of interface 3 to be static RP.
SwitchA(config)#mroute pim
SwitchA(config-router-pim)#bsr-candidate vlan-interface 2 8
SwitchA(config-router-pim)#rp-candidate vlan-interface 2
witchA(config-router-pim)#static-rp 192.168.2.2
SwitchA(config-router-pim)#

(4) Configure the unicast routing protocol so as to establish the correct unicast routing table.
SwitchA(config)#router ospf
SwitchA(config-router-ospf)#network 192.168.7.1 0.0.0.255 area 0
SwitchA(config-router-ospf)#network 192.168.1.1 0.0.0.255 area 0
SwitchA(config-router-ospf)#network 192.168.2.1 0.0.0.255 area 0

the configuration of switch B

(1) Enable multicast routing
SwitchA(config)#ip multicast-routing

(2) configure each interface as well as the interface address, and then enable PIM -SM. In addition, interface 5 enables IGMP protocol.
SwitchB#c t
SwitchB(config)#vlan 2
SwitchB(config-if-vlan)#vlan 4
SwitchB(config-if-vlan)#vlan 5
SwitchB(config-if-vlan)#exit
SwitchB(config)#interface ethernet 0/0/2
SwitchB(config-if-ethernet-0/0/2)#switchport default vlan 2
SwitchB(config-if-ethernet-0/0/2)#interface ethernet 0/0/3
SwitchB(config-if-vlan)#interface ethernet 0/0/3
SwitchB(config-if-ethernet-0/0/3)#switchport default vlan 4
SwitchB(config)#interface ethernet 0/0/6
SwitchB(config-if-ethernet-0/0/6)#switchport default vlan 5
SwitchB(config-if-ethernet-0/0/6)#
SwitchB(config)#interface vlan-interface 2
SwitchB(config-if-vlanInterface-2)#ip address 192.168.1.2 255.255.255.0
SwitchB(config-if-vlanInterface-2)#ip pim sparse-mode
SwitchB(config-if-vlanInterface-2)#interface vlan-interface 4
SwitchB(config-if-vlanInterface-4)#ip address 192.168.3.1 255.255.255.0
SwitchB(config-if-vlanInterface-4)#ip pim sparse-mode
SwitchB(config-if-vlanInterface-4)#interface vlan-interface 5

SwitchB(config-if-vlanInterface-5)#ip address 192.168.5.1 255.255.255.0
SwitchB(config-if-vlanInterface-5)#ip pim sparse-mode
SwitchB(config-if-vlanInterface-5)#ip igmp

(3) configure static RP. Specify the IP address of 13 VLAN interface to be the IP address of static RP; Specify the IP address of 4 VLAN interface to be C-BSR and C-RP; Specify the IP address of 3 VLAN interface in switch C to be the IP address of static RP.
SwitchB(config)#mroute pim
SwitchB(config-router-pim)#bsr-candidate vlan-interface 4 8
SwitchB(config-router-pim)#rp-candidate vlan-interface 4
SwitchB(config-router-pim)#static-rp 192.168.2.2

(4) configure the unicast routing protocol so as to establish the correct unicast routing table.

SwitchB(config)#router ospf
SwitchB(config-router-ospf)#network 192.168.1.2 0.0.0.255 area 0
SwitchB(config-router-ospf)#network 192.168.3.1 0.0.0.255 area 0
SwitchB(config-router-ospf)#network 192.168.5.1 0.0.0.255 area 0

the configuration of switch C

(1) Enable multicast routing
SwitchC(config)#ip multicast-routing
(2) Configure each interface as well as the interface address, and then enable PIM -SM. In addition, interface 6 enables IGMP protocol.
SwitchC(config)#vlan 3
SwitchC(config-if-vlan)#vlan 4
SwitchC(config-if-vlan)#vlan 6

SwitchC(config-if-vlan)#exit
SwitchC(config)#interface ethernet 0/0/2
SwitchC(config-if-ethernet-0/0/2)# witchport default vlan 3
SwitchC(config-if-ethernet-0/0/2)#interface ethernet 0/0/3
SwitchC(config-if-ethernet-0/0/3)#switchport default vlan 4
SwitchC(config-if-ethernet-0/0/3)#interface ethernet 0/0/6
SwitchC(config-if-ethernet-0/0/6)#switchport default vlan 6
SwitchC(config-if-ethernet-0/0/6)#exit
SwitchC(config)#interface vlan-interface 3
SwitchC(config-if-vlanInterface-3)#ip address 192.168.2.2 255.255.255.0
SwitchC(config-if-vlanInterface-3)#ip pim sparse-mode
SwitchC(config-if-vlanInterface-3)#interface vlan-interface 4
SwitchC(config-if-vlanInterface-4)#ip pim sparse-mode
SwitchC(config-if-vlanInterface-4)#ip address 192.168.3.2 255.255.255.0
SwitchC(config-if-vlanInterface-4)#interface vlan-interface 6
SwitchC(config-if-vlanInterface-6)#ip address 192.168.6.1 255.255.255.0
SwitchC(config-if-vlanInterface-6)#ip pim sparse-mode
SwitchC(config-if-vlanInterface-6)#ip igmp

(3) Configure static RP. Specify the IP address of 13 VLAN interface to be the IP address of static RP.
[SwitchC(config)#mroute pim
SwitchC(config-router-pim)#static-rp 192.168.2.2

(4) Configure the unicast routing protocol so as to establish the correct unicast routing table.

```
SwitchC(config)#router ospf
SwitchC(config-router-ospf)#network 192.168.2.2 0.0.0.255 area 0
SwitchC(config-router-ospf)#network 192.168.3.2 0.0.0.255 area 0
SwitchC(config-router-ospf)#network 192.168.6.1 0.0.0.255 area 0
```

三、Configuration Validation

Each pair of Switch A, Switch B and Switch C had established PIM neighbor relationship.

```
SwitchA(config)#show ip pim neighbor
Neighbor Address Interface                Uptime     Expires
192.168.1.2      VLAN-IF2                   00:03:57   00:01:22
192.168.2.2      VLAN-IF3                   00:03:33   00:01:42
Total Neighbors 2.
SwitchB(config)#show ip pim neighbor
Neighbor Address Interface                Uptime     Expires
192.168.1.1      VLAN-IF2                   00:05:19   00:01:26
192.168.3.2      VLAN-IF4                   00:10:17   00:01:28
Total Neighbors 2.


SwitchC(config)#show ip pim neighbor
Neighbor Address Interface                Uptime     Expires
192.168.2.1      VLAN-IF3                   00:01:24   00:01:21
192.168.3.1      VLAN-IF4                   00:09:54   00:01:18
Total Neighbors 2.
```

Display the RP information of PIM -SM domain via the command of “show ip pim rp-info”.

```
SwitchA(config)#show ip pim rp-info
GroupAddress      GroupMaskLen      RPAddress         ExpiryTime
224.0.0.0         4                 192.168.3.1       00:01:36
224.0.0.0         4                 192.168.1.1       00:01:36
```

Static RP is 192.168.2.2.

PC1 is added into225.0.1.2, SwitchB adds port6 to be the member interface of 225.0.1.2.

```
SwitchB(config)#show ip igmp groups
IGMP Connected Group Membership

Group Address: 225.0.1.2
   Vlan: 5, port: 0/0/6, Uptime: 00:00:08
   Expires: 00:04:12, Last Reporter: 192.168.5.2
   V1 Expires: 00:00:00, V2 Expires: 00:04:12, Self: False
   FilterMode: EXCLUDE, Static: False
   SourceList(0):
   Current State(IGMP_MS_NORMAL2)

Total Groups: 1, Total group members: 1
```

Multicast source server forwards the multicast video streams with the IP as 225.0.1.2 to check the PIM routing table information of Switch B. And then it will display the detailed information of（S，G） item, （*，G） item ,（S，G）item, PIM mode, inbound interface, upstream

neighbors, RPF neighbors, downstream interface, etc. At the same time, PC1 can be able to receive the video.

```
SwitchB(config)#show ip mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
            P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
            J - Join SPT, O - SRM originator, H - Hardware switched
            K - Static entry
Timers: Uptime/Expires, Interface state: State/Mode
(*, 225.0.1.2), 00:01:58/00:03:28, RP 192.168.2.2, flags: SCJ
   Incoming interface: VLAN-IF2, RPF nbr: 192.168.1.1
   Outgoing interface list:
VLAN-IF5, 0/0/6, Forward/Sparse, 00:01:58/00:01:33

(192.168.7.2, 225.0.1.2), 00:01:57/00:01:52, flags: SCTJ
   Incoming interface: VLAN-IF2, RPF nbr: 192.168.1.1
   Outgoing interface list:
VLAN-IF5, 0/0/6, Forward/Sparse, 00:01:57/00:01:33
Total dynamic entries 2. Total static entries 0.
Total ip multicast entries 2.

SwitchA(config)#SHOW IP mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
            P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
            J - Join SPT, O - SRM originator, H - Hardware switched
            K - Static entry
Timers: Uptime/Expires, Interface state: State/Mode
(*, 225.0.1.2), 00:00:47/00:03:33, RP 192.168.2.2, flags: SJ
   Incoming interface: VLAN-IF3, RPF nbr: 192.168.2.2
   Outgoing interface list:
VLAN-IF2, 0/0/3, Forward/Sparse, 00:00:47/00:02:43
(192.168.7.2, 225.0.1.2), 00:01:01/00:03:06, flags: SO
   Incoming interface: VLAN-IF4, RPF nbr: 0.0.0.0
   Outgoing interface list:
VLAN-IF2, 0/0/3, Forward/Sparse, 00:00:47/00:02:43
Total dynamic entries 2. Total static entries 0.
Total ip multicast entries 2.
```

7. IP Address Configuration

7.1 Layer2 Switch System IP Address

7.1.1 Overview for Layer2 Switch System IP Address

IP address is a unified address format provided by the IP protocol, assigning the logical address to each network and each host in order to shield the physical address differences. IP Address configures the IP for the device, which is easy to maintain and manage.

7.2 Layer3 Switch IP Address

7.2.1 Overview for Layer3 Switch IP Address

IP of L3 Switch can be used as management address or gateway. L3 Switch IP shall be configured at L3 interface, which is divided into VLAN interface and superVLAN interface. VLAN interface is to create an interface based on a certain VLAN, while superVLAN interface is based on superVLAN (superVlan is a virtual VLAN without any port). SuperVLAN includes multiple sub-VLANs (one sub-VLAN is a concrete VLAN).

7.2.2 Configure VLAN Interface

Configure VLAN Interface

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	
Create VLAN	vlan *vid*	optional
Add port into VLAN	**Switchport ethernet** *port*	optional
Create interface VLAN	interface **vlan-interface** *vid*	optional
Configure interface IP address	**ip address** {*ipaddress*| *primary*} *mask override*	optional
Delete interface IP address	**no ip address** *ipaddress mask*	optional
Configure IP access range control	**ip addres range** *start ipaddress end ipaddress*	optional
Delete IP access range control	**no ip address range** *start ipaddress end ipaddress*	optional

Note: Configure interface VLAN should be under L3 device. One interface can configure 32 IPs in different network.

IP access range control: every VLAN interface or superVLAN interface can be configured with up to eight access ranges. After the access range is configured, the user ARP must be within these ranges to learn, and thus limit the access of the user.

7.2.3 Configure SuperVLAN Interface

Configure SuperVLAN interface

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Create SuperVLAN interface	**interface supervlan-interface** id	optional
Add SuperVLAN sub-VLAN	**Subvlan** {vid| VLAN list }	optional
Configure SuperVLAN interface IP	**ip address** {ipaddress| **primary**} mask	optional
Delete SuperVLAN interface IP	no ip address {ipaddress| primary} mask	optional
Configure IP access range control	ip addres range start ipaddress end ipaddress	optional
Delete IP access range control	no ip address range start ipaddress end ipaddress	optional

Note: Configure interface superVLAN only under L3 device

7.2.4 Configure Override

When configuring the IP, add override command in the back, used to revise the ip in the same network segment.

Override configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Enter the interface mode	interface vlan-interface vid /supervlan-interface *id*	optional
Configuration	**ip address** ipaddress mask **overrid**	optional

7.2.5 Configure Loopback Interface

VLAN interface and SuperVLAN interface connect the ports directly while loopback interface connects the ports through VLAN interface and SuperVLAN interface. In the case, loopback interface won't be influenced by the port status, but always in linkup state. It will benefit a lot if loopback interface IP as the routing ID or source IP of sending packet.

Configure Loopback Interface

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Create loopback-interface	interface loopback-interface <0-1>	optional
Configure interface IP	ip address *ipaddress mask*	optional

7.2.6 Configure Interface Parameter

Configure system IP under VLAN interface for L3 device.

Configure system IP address under L3 device

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Enable the ICMP address mask reply packet	**ip icmp** mask-reply	optional
Enter interface VLAN mode	interface **vlan-interface** *vid*	required
Enable the sending of icmp destination unreachable packets	ip **icmp unreachable**	optional
Configure IP interface description	**Description** *interface-name*	optional
Delete IP interface description	**no description**	optional

7.2.7 Interface shutdown

You cannot manage a device after the interface is shut down.

Configure Interface shutdown

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Enter VLAN Interface	interface **vlan-interface** *vid*	required
Shutdown VLAN Interface	**shutdown**	optional
Enter SuperVLAN interface	interface **supervlan-interface** *vid*	required
Shutdown SuperVLAN interface	**Shutdown**	optional
Cancel shutdown interface	no **shutdown**	optional

7.2.8 IP Interface Display and Maintenance

After finishing the configuration above, you can use the following command to view the configuration.

Show IP Interface Configuration

Operation	Command	Remarks
Show IP interface configuration for L3 device	show ip interface {loopback-interface | supervlan-interface | vlan-interface }	In all modes

8. IPv6 Address Configuration

8.1 IPv6 Address Basics

IPv6 (Internet Protocol Version 6) is the second-generation standard network layer protocol, also known as IPng (IP Next Generation), which is designed by IETF (Internet Engineering Task Force), superior to IPv4. The biggest difference between IPv6 and IPv4 is that the length of IP address increase from 32 bits to 128 bits.

8.2 IPv6 address Pattern

IPv6 address is a series 16 bits hexadecimal number isolated by（:）. Each IPv6 address is divided in to 8 groups, and 4 hexadecimal numbers represent the 16 bits in each group. Two points（:）separate different groups, for example:

2001:0000:130F:0000:0000:09C0:876A:130B

In order to simplify the pattern, the "0" could be dealt with as below

- The front "0" could be omitted in each group. The above address could be shown as 2001:0:130F:0:0:9C0:876A:130B.
- If there is 0 in consecutively two or more than two groups, "::" double two points could replace it. For example: 2001:0:130F::9C0:876A:130B.

  There are two parts in the IPv6 address: address prefix and interface identification. Address prefix is similar to network number in IPv4 while interface identification is similar to mainframe number.

  Address prefix: IPv6 address/prefix length. IPv6 address could be in any forms listed above, but the prefix length is a decimal numeral, showing in what place from the left would be the prefix.

8.3 IPv6 Neighbor Discovery Protocol

IPv6 Neighbor Discovery Protocol adopts 5 types of ICMPv6 message to accomplish the activities below: address resolution, verifying neighbor reachability, duplicate address detection, router discovery/prefix discovery, address auto-configuration, redirection.

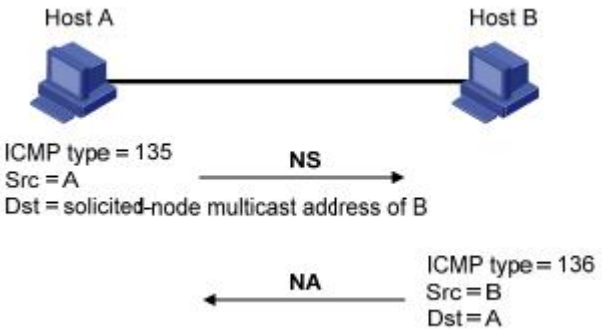
ICMPv6 message types and functions in neighbor discovery protocol are shown as below:

ICMPV6 Message	Funtion
Neighbor Solicitation (NS)	Get the link-layer address of a neighbor
	Neighbor reachability
	Duplicate address detection
Neighbor Advertisement(NA)	NS message for response
	When the link layer node address change sends NA messages, change information notice to the neighbors of this node
Router Solicitation(RS)	After startup, a host sends an RS message to request the router, request prefixes and other configuration information used to automatically configure the host
Router Advertisement(RA)	RS message for response
	In RA messages suppression disabled, the router sends an RA message containing the prefix and flag bits of news
Redirect	When certain conditions are satisfied, the default gateway sends a redirect message to the source host so that the host can reselect a correct next hop router to forward packets

Functions in neighbor discovery protocol:

- Address resolution

Get the link layer address of neighbor node on the same link (same as ARP function of IPv4) through NS and NA. As shown as chart 1-1, host A gets link layer address from host B.



(1) Host A sends NS by the means of multicast. The source address of NS is host A interface IPv6 address. Destination address is the node multicast address requested by host B and the message also includes link layer address of host A.

(2) When host B receives NS, judges whether the destination address matches the IPv6 requested node multicast address. If yes, the host B could learn the link layer address of host A and reply NA message by unicast mode, with the link layer address of itself.

(3) Host A receives NA and gets the link layer address of host B. Hence, host A could communicate with host B.

- Verify the neighbor reachability

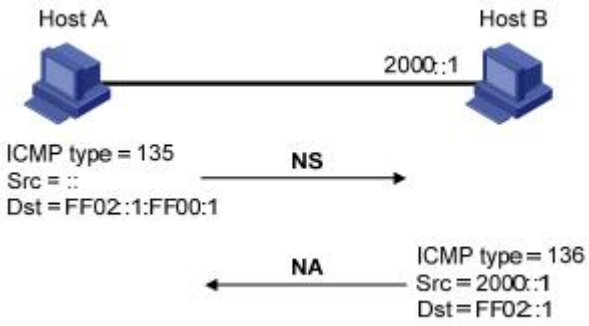
After getting the link layer address of neighbor node, the Neighbor Solicitation (NS) and Neighbor Advertisement (NA) messages can be used to verify that the neighbor node is reachable.

(1) Node sends NS, where the destination address is the IPv6 address of the neighbor node.

(2) If the confirmation message is received from the neighbor node, it considers the neighbor as reachable. Otherwise, it considers that the neighbor is unreachable..

- Duplicate address detection

When the node receives a IPv6 address, duplicate address detection would be reactive to ensure whether this address is occupied by other nodes by the means of NS and NA. (same as the IPv4 free ARP). Shown as below:



(1) Host A sends NS, NS message source address is unknown address::, and the destination address is the IPv6 address requested node multicast address to be detected, and message includes the IPv6 address to be detected.

(2) If host B already occupies this IPv6 address, NA message will be returned, and the message also includes the IPv6 address of its own.

(3) If host A receives NA from host B, that is to say the IPv6 address is taken. If not, then the address is available for the host A.

8.4 IPv6 Concrete Configuration

8.4.1 Configure Ipv6 Unicast Address

To access IPv6 network, IPv6 address must be configured, and we have to choose one from global unicast address, site local address, link local address.

Before access to IPv6 network, IPv6 unicast address must be configured.

- IPv6 site local address and global unicast address could be configured by the following 4 ways:

（（1） EUI-64: When adopt EUI-64 forming IPv6, the prefix of IPv6 address of the interface is the configured prefix. The interface identifier is translated from the link-layer address of the interface.

（2） Manual mode: Configure IPv6 site local address or global unicast address by manual mode.

（3）DHCP: It supports to get IPv6 site local address or global unicast address or some related information through DHCP server.

(4) Auto configuration: IPv6 and related information is automatically configured based on its own link layer address and the prefix information advertised by the router.

- Obtain IPv6 link local address by two ways:

(1) Automatic generation: automatically generates link local address for interface based on link local address prefix (FE80::/64) and link layer address of the interface.

(2) Manual assignment: Manually configure IPv6 link local address

Configure L2 device under global configuration mode, and L3 device under interface mode.

Layer3 Switch Ipv6 address Configuration

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enable ipv6 forwarding function	**Ipv6 {enable | disable}**	required
Enter vlan interface / super interface mode	**interface { vlan-interface** *vid* **| supervlan-interface** *interface-number* **}**	
Configure the site-local address and global unicast address in EUI-64 format	**ipv6 address** *ipv6-address/prefix-length* **eui-64**	required
Delete the site-local address and global unicast address configured in EUI-64 format	**no ipv6 address** *ipv6-address/prefix-length* **eui-64**	
Manually specify the site-local address and the global unicast address	**ipv6 address** *ipv6-address/prefix-length*	
Delete manually specified site-local addresses and global unicast addresses	**no ipv6 address** *ipv6-address/prefix-length*	
Automatically configure the site-local address and the global unicast address	**ipv6 address autoconfig**	
Delete the auto-configured site-local address and global unicast address	**no ipv6 address autoconfig**	
Specify the link-local address manually	**ipv6 address** *ipv6-address* **link-local**	By default, a link-local address is automatically formed
Delete the manually specified link-local address	**no ipv6 address** *ipv6-address* **link-local**	
Show ipv6 address status	**show ipv6 interface { vlan-interface** *vid* **| supervlan-interface** *interface-number* **}**	

8.4.2 Configure Static Neighbor List

Send NS&NA or configure static neighbor list by command line in order to resolve IPv6 address of neighbor node into link layer address.

Static neighbor list includes long static neighbor list or short static neighbor list.

long static neighbor list: Except to configure IPv6 address and MAC address, the VLAN and port of neighbor list must be configured too when configuring long static neighbor list. Long static neighbor list could be directly used for packet forwarding..

short static neighbor list: When configuring the short static neighbor list, only need to configure IPv6 address and MAC address. Short static neighbor list can't be used for packet forwarding. When it comes to short list static neighbor list, neighbor request is sent firstly. If the configured IPv6 address and MAC address are the same as the source IPv6 address and source MAC address of the response packet, then complete the ARP list and it could be used in forwarding IPv6 packet.

Operation	Command	Operation
Enter global configuration mode	**configure terminal**	
Configure long static neighbor list	**ipv6 neighbor** *ipv6-address mac-address vlan-id* device/slot/port	
Configure short static neighbor list	**ipv6 neighbor** *ipv6-address mac-address*	
Delete neighbor list	**no ipv6 neighbor {dynamic | static | all | }**	
Delete vlan interface neighbor list	**no ipv6 neighbor** *ipv6-address* **interface { vlan-interface** *vid* **| supervlan-interface** *interface-number* **}**	
Show neighbor list	**show ipv6 neighbors {** *ipv6-addres* **|all | dynamic| static|mac** *mac-address***}**	

8.4.3 Configure MAX Number of Neighbors

If the number of accessing neighbors is too big, it might influence the transmission performance. We can limit the MAX number of neighbors by configuration..

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure max neighbor number	**ipv6 neighbors max-learning-num** *number*	MAX neighbor number includes static and dynamic neighbor number, the default neighbor number is 64
Show limited max neighbor number	**show ipv6 neighbors max-learning-num**	

8.4.4 Configure the Number of Sending NS for Duplicate Address Detection

In order to verify whether there is a conflict among the address, the device sends NS for duplicate address detection. If there is no response in a certain time (through ipv6 nd ns retrans-timer command), then continue to send NS. If still no response after the times of sending message reached the setting number, the address is available.

表 1-1

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure the number of sending NS for duplicate address detection	**ipv6 nd dad attempts** *value*	optional By default, the number of sending NS in duplicate address detection is 1. When value is 0, it means that duplicate address detection is disabled
Recover **attemps** defaults	**no ipv6 nd dad attemps**	
Set the interval for sending NS	**ipv6 nd ns retrans-timer** *value*	Optional; units in seconds. By default, the interval for sending NS message is 1 second
Recover **retrans-time** default value	**no ipv6 nd ns retrans-time**	
Show the number of sending NS for duplicate address detection	**show ipv6 nd dad attemps**	
Show the interval for sending NS message	**show ipv6 nd ns retrans-time**	

8.4.5 Configure the Time to Keep the Neighbor Reachable State

When the neighbor reachability is determined by neighbor reachability detection, the device considers the neighbor reachable within the setting reachable time. After it runs over the setting time, If a packet needs to be sent to the neighbor, the neighbor is re-acknowledged.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure the time to keep the neighbor reachable state	**ipv6 nd reachable-time** *value*	Optional; Units in seconds. reachable-time is 30 seconds as default
Recover **reachable-time** default value	**no ipv6 nd reachable-time**	
show neighbor reachable-time	**show ipv6 nd reachable-time**	

8.4.6 Configure IPv6 Static Route

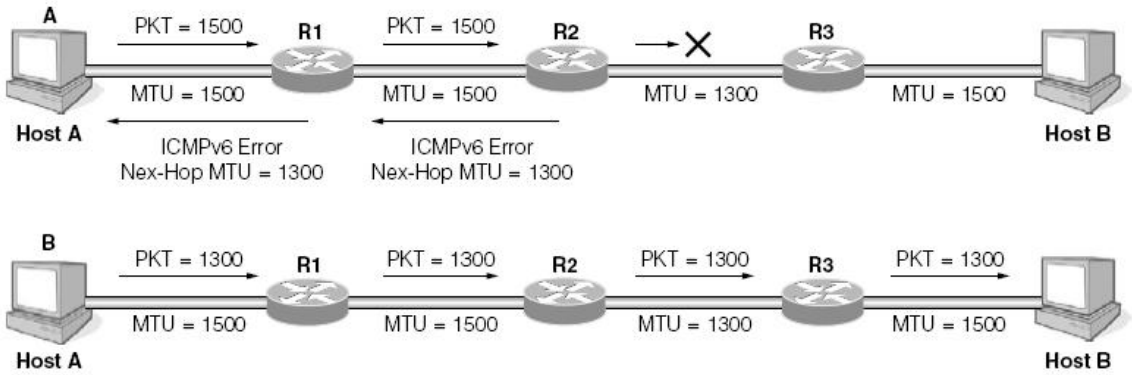
Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure IPv6 static route	**ipv6 route** [*ipv6-address mask|* *ipv6-address/prefix-length*] *nexthop-address*	required By default, no ipv6 static

		route is configured.
Delete IPv6 static route	**ipv6 route** [*ipv6-address mask|ipv6-address/prefix-length*] *nexthop-addres*	
Show ipv6 route list	**show ipv6 route**	

8.4.7 Configure Interface MAX Transmission Unit (MTU)

Ipv6 interface path MAX transmission unit can be set in the range of 1280-1510 bytes. The Ipv6 packet that smaller than 1280 bytes must be fragmented and encapsulated. The packet is bigger than setting MTU, interface will discard the packet.

- Steps for obtaining interface MTU
- (1) The sending node assumes that the path MTU is the forwarding egress link.
- (2) The sending node sends packet according to the assumed path MTU.
- (3) If router can't forward this packet due to the transmitting MTU is smaller than the assumed MTU, router discards the packet and return ICMPv6 overloaded packet to the sending node. The overloaded packet carries the failure transmission link MTU.
- (4) The sending node sets the path MTU to the MTU value in the ICMPv6 overloaded packet.



Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enter vlan interface / super interface mode	**interface { vlan-interface** *vid* | **supervlan-interface** *interface-number* **}**	
Configure the interface MTU	**ipv6 pathmtu** *value*	Optional, default value is 1500
Recover default	**no ipv6 pathmtu**	

8.4.8 Device receiving multicast Echo request responds Echo reply packet

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Response multicast	**ipv6 icmpv6 multicast-echo-reply enable**	

packet		
Do not response multicast packet	**no ipv6 icmpv6 multicast-echo-reply**	Default value is inactive

8.5 IPv6 Unicast Address Configuration Example

8.5.1 Networking Requirements

Two Switches are interconnected through Ethernet ports. Configure IPv6 address for both Switches, to verify their interconnection. SwitchA global unicast address is 2001::1/64, SwitchB global unicast address is 2001::2/64.

8.5.2 Networking Diagram

（1）configure SwitchA
SwitchA(config)#interface vlan-interface 1
SwitchA(config-if-vlanInterface-1)#ipv6 address 2001::1/64
SwitchA(config-if-vlanInterface-1)#ipv6 address fe80:12::1 link-local

（2）configure SwitchB
SwitchB(config)# interface vlan-interface 1
SwitchB(config-if-vlanInterface-1)#ipv6 address 2001::2/64
SwitchB(config-if-vlanInterface-1)#ipv6 address fe80:12::2 link-local

8.5.3 Verify the configuration

Show SwitchA ipv6 information
SwitchA(config)#show ipv6 interface vlan-interface 1
Show informations of ipv6 interface

VLAN-IF1:
sw0 Link type:Ethernet HWaddr 00:00:00:09:99:99 Queue:none
IPv6 forwarding is disabled
inet6 unicast 2001::1 prefixlen 64
inet6 unicast FE80::200:FF:FE09:9999%sw0 prefixlen 64 automatic
inet6 unicast 2001:: prefixlen 64 anycast
inet6 multicast FF02::1%sw0 prefixlen 16 automatic
inet6 multicast FF02::1:FF09:9999%sw0 prefixlen 16
inet6 multicast FF02::1:FF00:1%sw0 prefixlen 16
inet6 multicast FF02::1:FF00:0%sw0 prefixlen 16
UP RUNNING SIMPLEX BROADCAST MULTICAST PROMISC
MTU:1500 metric:1 VR:0 ifindex:2
RX packets:150 mcast:35 errors:0 dropped:0

TX packets:1216 mcast:33 errors:0
collisions:0 unsupported proto:0
RX bytes:13k TX bytes:55k

show SwitchB ipv6 information

SwitchA(config)#show ipv6 interface vlan-interface 1
Show informations of ipv6 interface

VLAN-IF1:
sw0 Link type:Ethernet HWaddr 00:01:7a:e9:68:58 Queue:none
IPv6 forwarding is disabled
inet6 unicast FE80:12::2%sw0 prefixlen 64
inet6 unicast 2001::2 prefixlen 64
inet6 unicast FE80::201:7AFF:FEE9:6858%sw0 prefixlen 64 automatic
inet6 multicast FF02::1%sw0 prefixlen 16 automatic
inet6 multicast FF02::1:FFE9:6858%sw0 prefixlen 16
inet6 multicast FF02::1:FF00:2%sw0 prefixlen 16
UP RUNNING SIMPLEX BROADCAST MULTICAST PROMISC
MTU:1500 metric:1 VR:0 ifindex:2
RX packets:21912 mcast:7990 errors:0 dropped:73
TX packets:8300 mcast:8023 errors:0
collisions:0 unsupported proto:0
RX bytes:1858k TX bytes:714k

Total entries: 1 interface.
SwitchA Ping SwitchB local address and global unicast address. If the configuration is right, then the two types of IPv6 addresses can be pinged successfully.
SwitchA(config)#ping6 FE80:12::2%sw0

Pinging FE80:12::2%sw0 (FE80:12::2%sw0) with 56 bytes of data:
Reply from FE80:12::2%sw0 bytes=56 time=10ms hlim=64
Reply from FE80:12::2%sw0 bytes=56 time=10ms hlim=64
Reply from FE80:12::2%sw0 bytes=56 time=10ms hlim=64
Reply from FE80:12::2%sw0 bytes=56 time=10ms hlim=64

--- FE80:12::2%sw0 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 4840 ms
rtt min/avg/max = 10/10/10 ms

SwitchA(config)#ping6 2001::2

Pinging 2001::2 (2001::2) with 56 bytes of data:
Reply from 2001::2 bytes=56 time=10ms hlim=64
Reply from 2001::2 bytes=56 time=10ms hlim=64
Reply from 2001::2 bytes=56 time=10ms hlim=64
Reply from 2001::2 bytes=56 time=10ms hlim=64

--- 2001::2 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 4840 ms
rtt min/avg/max = 10/10/10 ms

show SwitchA neighbor list
SwitchA(config)#show ipv6 neighbors all

Information of neighbor cache

Neighbor	Mac_Address	Vlan	Port	Type	Expire	Status
2001::2	00:01:7a:e9:68:58	1	e0/0/1	Dynamic	941 s	stale
FE80:12::2%sw0	00:01:7a:e9:68:58	1	e0/0/1	Dynamic	946 s	stale

Total entries:2

9. ARP Configuration

9.1 ARP Overview

9.1.1 ARP Function

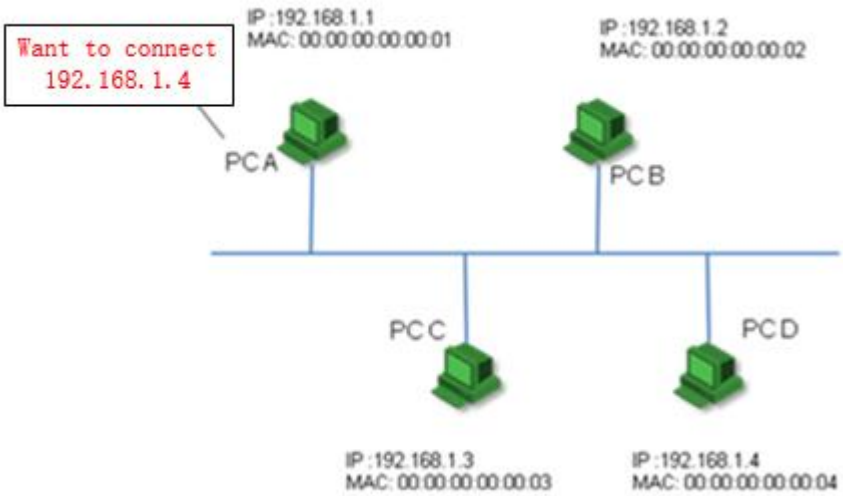
ARP, Address Resolution Protocol, is one of the most important protocols in TCP/IP family. An IP address is the address of a host at the network layer. To send a network layer packet to a destination host, the device must know the data link layer address (such as the MAC address) of the destination host. To this end, the IP address must be resolved into the corresponding data link layer address.

Unless otherwise stated, the data link layer addresses that appear in this chapter refer to the 48-bit Ethernet MAC addresses.

9.1.2 Operating Process of ARP

Take FTP communication for example to describe the operating process of ARP.

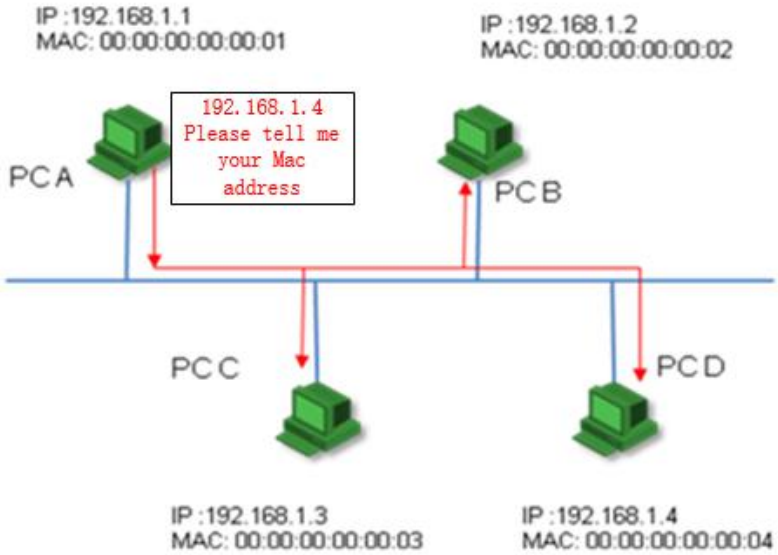
As shown below, host A expects to access the host with IP address of 192.168.1.4.



Network diagram

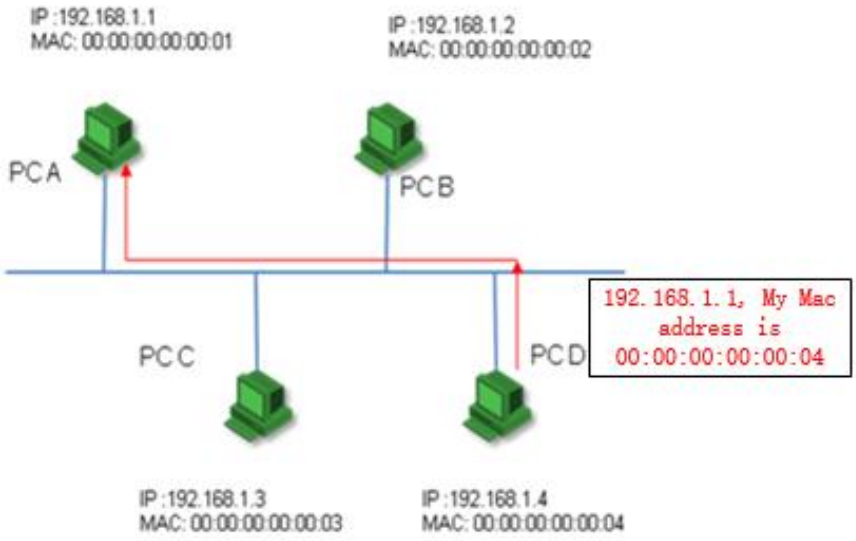
Assume that this is an Ethernet, and each host all doesn't know other host in local area network (LAN). On this occasion, host should know the host MAC address of 192.168.1.4 before establishing communication.

According to ARP protocol, host A will send an ARP request to ask for the MAC address of 192.168.1.4. This request is a broadcast message, so all the hosts in the local area network will receive this request. As shown below.



Forward ARP request

According to protocol, only host D can be able to reply the ARP request of host A, and this ARP reply is a unicast message.



Forward ARP reply

Host A will record the IP address and MAC address of host D in ARP cache after receiving the reply of host D. In that case, host doesn't need to send ARP request to ask for the MAC address of destination host unless this table is aged.

9.1.3 ARP Table

After the equipment analyzed the destination MAC address via ARP, it will add the mapping table to its ARP table, such as IP address, MAC address, port and so forth.

ARP table is divided into dynamic ARP table and static ARP table. Layer-2 equipment only supports dynamic ARP table, and layer-3 equipment supports both dynamic ARP table and static ARP table.

（1）dynamic ARP table is generated and maintained via ARP packet, it can be aged, updated by the new ARP packet and covered by static ARP table. If it has reached the aged time, it will delete corresponding ARP table when the port is down.

（2）static ARP table is mainly configured and maintained by manual. It cannot be aged or covered by dynamic ARP table.

Static ARP table is divided into short static ARP table and long static ARP table.

When configuring long static ARP table, you should configure the IP address and MAC address as well as the VLAN and egress port of this ARP. Long static ARP packet can be used to transmit the packet directly.

When configuring the short static ARP table, you just need to configure the IP address and MAC address. Short static ARP table cannot be used to transmit the packet directly. When you need to use the short static ARP table, you should send ARP request packet firstly, if the source IP address and source MAC address of the reply packet are the same as the configured IP address and MAC address, just complete this ARP table, and then you can used it to transmit the IP data packet.

Note:

When configuring the long static ARP table manually, the ip address of the ARP table should be in the same network segment with ip address of the egress port. Or the adding operation will not succeed.

9.2 ARP Configuration

9.2.1 ARP Table Configuration

Procedure	Command	Operation
Enter global configuration mode	**configure terminal**	

Configure short static arp table	**arp {ipaddress mac** *mac* **}**	
Configure long static arp table	**arp {ipaddress mac** *mac* **vid** *vid* **port** *port* **}**	
Configure the aging time	**arp aging-time** *aging-time*	20min by default
Configure the dynamic arp to be static arp	**arp bind dynamic {***ipaddress* **| all}**	
Delete arp table	**no arp { dynamic | static | all |** *ipaddress* **}**	
Display arp table	**Show arp { dynamic | static | all}**	

9.2.2 ARP peer

ARP Peer means that two Switches learn each other's ARP only by the specified port.

Procedure	Command	Operation
Enter global configuration mode	**configure terminal**	
Configure arp peer table	**arp peer {ipaddress** *mac port* **}**	
Delete arp peer table	**no arp peer**	

Instance Example:

Switch(config)#arp peer 192.168.1.1 00:56:3A:40:5A:01 0/0/1

The MAC address of above Peer is 00:56:3A:40:5A:01. In addition, the ARP message corresponds to this MAC address takes the effect only when it comes from Ethernet0/0/1. IP 192.168.1.1 only acts as a label.

9.2.3 ARP overwrite

The Switch deals with the ARP conflict via this command. If the port enables this function, ARP conflict table will be updated to this port. Or the ARP conflict will not be dealt.

Operation	Command	Remarks
Enter port configuration mode	**interface ethernet** *device/slot/port*	
Configure the ARP overwrite function	**arp overwrite**	Disabled by default
Forbid the ARP overwrite function	**no arp overwrite**	

9.2.4 Linkup gratuitous-arp

By default, the Switch will not take the initiative to forward gratuitous ARP message from the Switch port when this port in linkup state. You can configure the Switch forwards gratuitous ARP message from this port so as to detect the ip conflict.

Operation	Command	Remarks
Enter port configuration mode	**interface ethernet** *device/slot/port*	
Forward gratuitous arp message when configuring port linkup	**linkup gratuitous-arp**	Disabled by default
Limit the port forwarding gratuitous arp message when the port is in the state of linkup	**no linkup gratuitous-arp**	

9.2.5 Arp-reply-repeat

The Switch responds to each ARP request message with only one reply message by default (The premise is that the request message asks for the ARP of the Switch). You can configure the port responds to each ARP request message with multiple ARP reply message so as to support a certain kind of protocol.

Operation	Command	Remarks
Enter port configuration mode	**configure terminal**	
Configure the interval and frequency of arp-reply-repeat	**arp-reply-repeat interval** *interval* **times** *times*	The unit is millisecond and default parameter refers to repeat reply every 20 milliseconds.
Enter interface mode	**interface ethernet** *device/slot/port*	
Configure arp-reply-repeat function	**arp-reply-repeat**	disabled by default
Disabled arp-reply-repeat function	**no arp-reply-repeat**	

9.2.6 ARP Detection

The principle of ARP detection is to configure the remote IP address, that is, set the corresponding ARP table state as PROBE state and configure the aging time of this ARP table as retransmission interval. If it receives the remote ARP reply, it will update the aging time of this ARP table as normal value (20 minutes). Or it will retransmit. In addition, if it reaches the retransmission times while it still hasn't received a reply, the ARP table will be deleted.

Procedure	Command	Operation
Enter global configuration mode	**configure terminal**	
Configure the remote device ip	**arp probe ip { ip }**	You can configure 4 ip for the maximum.
Delete the remote device ip	**no arp probe ip { all | ip }**	
Configure the parameter of arp probe ip	**arp probe [poll-timer** *value* **| retransmit { count value | interval** *value* **}]**	Poll-timer: value range is 60-300 seconds , and the default value is 180 seconds. Count: retransmission times, value range is 2-5 , and the default value is 3 times Interval:retransmission interval , value range is 1-3 seconds , and the default value is 3 seconds.
Show arp probe parameter.	**show arp probe**	

Note: It is not allow configuring poll-timer during arp-probe running procedure.

9.2.7 ARP - Proxy

arp-proxy: ARP request message is broadcast message, so it cannot pass through VLAN. If ARP is enabled, hosts of sub-VLAN in the same superVLAN interface can be able to perform ARP interaction, that is, hosts can be able to communicate with each other.

arp-proxy broadcast: sub-VLAN can be able to perform arp-proxy broadcast to other sub-VLAN by default if sub-VLAN enable arp-proxy function. The command of “no arp-proxy broadcast” can be used to limit the ARP request message performing broadcast to other sub-VLAN.

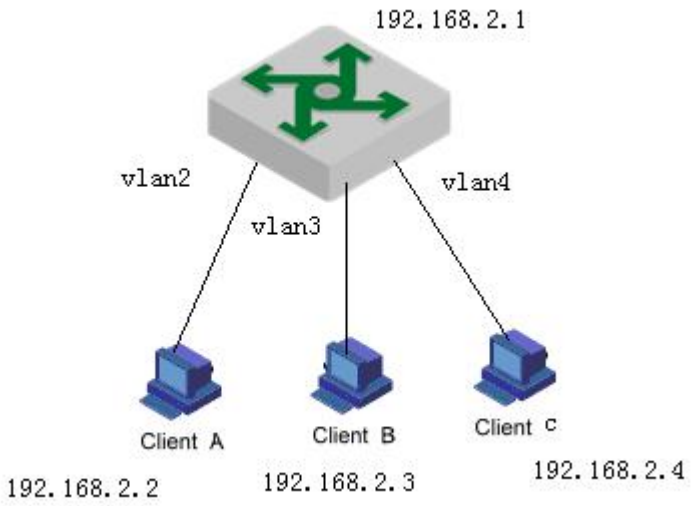
application environment: arp-proxy is adopted when the host ip of sub-VLAN and the superVLAN interface ip are in the same network segment. They can perform 3-layer forward via gateway instead of adopting arp-proxy if they are in different segment.

Operation	Command	Remarks
Enter vlan configuration mode	**vlan <vlanid>**	-

Enable arp-proxy	**arp-proxy**	
Disable arp-proxy	**no arp-proxy**	optional
Enable arp-proxy broadcast	**arp-proxy broadcast**	optional
Disable arp-proxy broadcast	**no arp-proxy broadcast**	optional
Display the ARP proxy information configured on the current system	**show arp-proxy**	It can be executed in all modes.

Configuration instance:

As shown in the following figure: VLAN 2, 3, 4 are the sub-VLAN of supervlan-interface1, and they are connected to computerA, computerB, computerC respectively with the arp-proxy enabled. In addition, VLAN 4 is in the state with arp-proxy broadcast disabled.




```
DUT2(config)#interface supervlan-interface 1
DUT2(config-if-superVLANInterface-1)#subvlan 2-4
DUT2(config-if-superVLANInterface-1)#ip address 192.168.2.1 255.255.255.0
DUT2(config-if-superVLANInterface-1)#exit
DUT2(config)#vlan 2-4
DUT2(config-if-vlan)#arp-proxy
Config arp-proxy enable successfully.
DUT2(config-if-vlan)#exit
DUT2(config)#vlan 4
DUT2(config-if-vlan)#no arp-proxy broadcast
Config arp-proxy broadcast disable successfully.
```

1、A forwards arp request packet， B\C can be able to respond， A PING B ，communication-capable

2、C forwards arp request packet，A\C cannot receive，C ping A，communication-disable

10.Mirroring

Mirroring is to copy packets matching the specified rule to the mirroring destination port. Generally, the destination port is connected to the data detection device. Users can analyze the mirrored packets, monitor the network, and troubleshoot faults. Mirroring is divided into port mirroring, remote port mirroring, and flow mirroring.

10.1 Port Mirroring

Port mirroring, which is used to copy the packets received or sent on the specified port to the mirroring destination port. Switch supports one-to-one and many-to-one mirroring, which can support multiple mirroring sources.

- mirrored: it can be a port or a packet that the CPU receives or sends.
- mirror: For the Switch, the destination port of the mirror can only be one. If the mirroring destination port is configured, only the mirroring destination port of the last configuration takes effect.

10.1.1 Configure Port Mirroring

Configuring Port Mirroring

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure mirrored	**mirror source-interface** {**ethernet** *device/slot/port* | **cpu** **}** {**ingress** | **egress** | **both**}	Required; You can configure multiple mirroring source ports
Configure mirror	**mirror destination-interface ethernet** *device/slot/port*	Required; You can specify only one mirroring destination port
Delete a mirroring group	**no mirror {soure-interface {cpu | ethernet** *device/slot/port* **}** **| destination-interface ethernet** *device/slot/port* | **all }**	optional
Display mirroring groups	**show mirror**	optional

10.1.2 Configuration Example for Port Mirror

1.Network requirements
Mirror the packet of CPU, e 0/0/1, e 0/0/2 to e 0/0/4.

2.Configuration steps
Switch(config)#mirror source-interface cpu both
Switch(config)#mirror source-interface ethernet 0/0/1 both
Switch(config)#mirror source-interface ethernet 0/0/2 both
Switch(config)#mirror destination-interface ethernet 0/0/4

3.Result validation
Switch(config)#show mirror
Information about mirror port(s)
The monitor port : e0/0/4
The mirrored egress ports : cpu,e0/0/1-e0/0/2.
The mirrored ingress ports : cpu,e0/0/1-e0/0/2.

The packet of CPU, e 0/0/1, e 0/0/2 can be mirrored to port e 0/0/4.

10.2 RSPAN

RSPAN, that is, Remote Switched Port Analyzer, breaks the restriction that mirrored ports and mirror ports must be on the same Switch. RSPAN allows mirrored and mirrored ports to span multiple devices in the network, facilitating the management of remote Switch devices.

There are three types of Switches that can implement RSPAN functions:

- Source Switch: The Switch where the monitored port resides is responsible for forwarding the traffic to the intermediate Switch or the destination Switch via rspan vlan.
- Intermediate Switch: Switches between the source Switch and the destination Switch transmits the mirrored traffic to the next intermediate Switch or destination Switch through the rspan vlan. If the source Switch is directly connected to the destination Switch, there is no intermediate Switch.
- Destination Switch: The Switch where the remote mirroring destination port located forwards the mirrored flow received from the rspan vlan to the monitoring device through the mirroring destination port

The ports that participate in mirroring on each Switch are shown in the following table:

The ports that participate in mirroring on each Switch

Switch	The ports that participate in mirroring	Function
Source switch	Source port	The monitored user port copies the user data packets to the specified local destination port through local port mirroring. There can be multiple source ports.
	Destination port of local mirror	to receive the user data packet of local port mirror
Intermediate switch	Trunk port	forward the mirrored packets to the destination switch On the intermediate switch, it is recommended to configure two trunk ports, which are connected to the devices on both sides
Destination switch	Trunk port	to receive remote mirror packets
	Destination port	Monitor port for remote mirror packets

In order to implement remote port mirroring, you need to define a special VLAN, called rspan vlan. All the mirrored packets are transmitted from the source switch of this VLAN to the mirroring port of the destination switch to monitor the source packets of the remote switch port based on the destination switch.

Rspan vlan has the following characteristics:

- It is recommended that you configure the device interconnection ports in the VLAN as trunk ports.
- You cannot configure the default VLAN and management VLAN as rspan vlan.
- You need to configure the rspan vlan to ensure Layer2 interoperability from source switch to destination switch.

10.2.1 Configure Remote Port Mirror

Source device configurations

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the local mirror source	**mirror source-interface** {**ethernet** *device/slot/port* | **cpu}** {**ingress** | **egress** | **both**}	required
Configure the destination port for local mirror	**mirror destination-interface ethernet** *device/slot/port*	required
Enter interface configuration mode	**interface ethernet** *device/slot/port*	
Enable remote mirror	**remote_mirror rspan local vlan** *vlan-id* **tpid [tpid** *tpid***]**	required
Delete remote mirror	**no remote_mirror rspan local**	required
Verify the operation	**show remote_mirror**	optional

Note:

remote_mirror rspan enable vlan is for the source mirroring device. Only one remote source vlan can be configured on a device.

---

intermediate device configurations

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the remote mirror VLAN	**[no]remote_mirror rspan middle vlan** *vlan-id*	required
Verify the	**show remote_mirror**	optional

operation		

Note:

remote_mirror rspan disable vlan is for the intermediate mirror device and it can configure multiple remote mirroring vlans. A device can either be a mirrored source device or an intermediate device.

destination device configurations

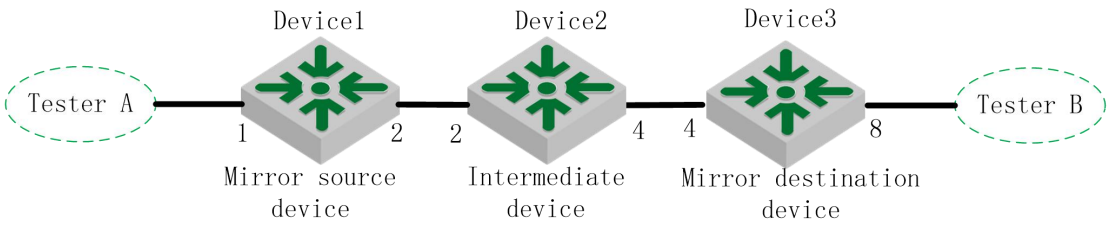
Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter interface configuration mode	**interface ethernet** *device/slot/port*	
Configure the remote mirror VLAN	**remote_mirror rspan target vlan** *vlan-id* **[tpid** *tpid***]**	required
Delete remote mirror vlan	**no remote_mirror rspan target**	
Verify the operation	**show remote_mirror**	optional

10.2.2 Configuration Example for Remote Port Mirroring

1.Network requirements

The packets from Device 1 on port 1 can be mirrored to port 8 on Device 3.

Network diagram is as follows:



2.Configuration steps

#Device1 Configuration:
Switch(config)#mirror source-interface ethernet 0/0/1 both
Switch(config)#mirror destination-interface ethernet 0/0/2
Switch(config)#vlan 100
Switch(config-if-vlan)#switchport ethernet 0/0/2
Switch(config)#interface ethernet 0/0/2
Switch(config-if-ethernet-0/0/2)#switchport mode trunk
Switch(config-if-ethernet-0/0/2)#remote_mirror rspan local vlan 100 tpid 9100

#Device2 Configuration:
Switch(config)#vlan 100
Switch(config-if-vlan)#switchport ethernet 0/0/2 ethernet 0/0/4
Switch(config-if-vlan)#exit
Switch(config)#interface ethernet 0/0/4
Switch(config-if-ethernet-0/0/4)#switchport mode trunk
Switch(config-if-ethernet-0/0/4)#exit
Switch(config)#remote_mirror rspan middle vlan 100

#Device3 Configuration:
Switch(config)#vlan 100
Switch(config-if-vlan)#switchport ethernet 0/0/4 ethernet 0/0/8
Switch(config-if-vlan)#exit
Switch(config)#interface ethernet 0/0/8
Switch(config-if-ethernet-0/0/4)#interface ethernet 0/0/8
Switch(config-if-ethernet-0/0/8)#remote_mirror rspan target vlan 100 tpid 9100
3.Result validation
The packets from Device 1 on port 1 can be mirrored to port 8 on Device 3.

10.3 Flow Mirror

Flow mirror is to copy the service flow matching ACL rules to the specified destination port for packet analysis and monitoring. Before configuring flow mirror, you need to define the ACL rules that meet the requirements. The device references these ACL rules for flow identification.

10.3.1 Configure Flow Mirror

Configure Flow Mirror

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure flow mirror	**mirrored-to** {ip-group<1-199>|link-group<200-299> [**subitem** <0-127>]	required
Remove flow mirror	**no mirrored-to** {ip-group<1-199>|link-group<200-299> } [**subitem** <0-127>]	optional
Verify the operation	**show mirror**	optional

10.3.2 Configuration Example for Flow Mirror

1.Network requirements
Mirror the packets whose source IP address is 10.1.1.1 to e 0/0/7.

2.Configuration steps

Switch(config)#access-list 100 permit 10.1.1.1 0 any
Switch(config)#mirror destination-interface ethernet 0/0/7
Switch(config)#mirrored-to ip-group 100

3.Result validation
The e 0/0/7 port can catch packets with source IP 10.1.1.1

11.SNMP Login Management

11.1 SNMP Overview

SNMP (Simple Network Management Protocol) is an important network management protocol on TCP / IP networks, implementing network management by exchanging packets on the network. The SNMP protocol provides the possibility of centralized management of large networks. Its goal is to ensure the management information is transmitted between any two points. SNMP is convenient for the network administrator to retrieve information from any node on the network, make modifications, find faults, and complete fault diagnosis, capacity planning and report generation.

SNMP structure is divided into two parts: **NMS** and **Agent**. **NMS** (Network Management Station) is a workstation that runs client programs while **Agent** is a server-side software running on a network device. The NMS can forward GetRequest, GetNextRequest, and SetRequest packets to the Agent. Upon receiving the NMS request message, the agent performs Read or Write operations according to the packet type and generates a Response packet to return to the NMS. On the other hand, when the device encounters an abnormal event such as hot / cold start, the agent will forward a trap packet to NMS to report the events.

The system supports SNMP v1, SNMP v2c and SNMP v3. SNMP V1 provides a simple authentication mechanism, does not support the administrator-to-manager communications, and v1 Trap has no confirmation mechanism. V2c enhanced v1 management model (on security), management information structure, protocol operation, manager and communication ability between managers to increase the creation and deletion of the table, the communication ability between managers, reducing the storage side of the agent. V3 implements the user authentication mechanism and packet encryption mechanism, which greatly improves the security of the SNMP protocol.

This function cooperates with the network management software to log on to the switch and manage the switch.

11.2 Configuring the Basic Parameters

Configuring the Basic Parameters

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enable/disable SNMP	snmp-server [enable | disable]	optional. Enabled by default.
Configure sysContact	[no] snmp-server contact ***syscontact***	optional, with default parameters
Display sysContact configuration	show snmp contact	optional
Configure sysLocation	[no] snmp-server location ***syslocation***	optional, with default parameters
Display sysLocation configuration	show snmp location	optional
Configure sysName	[no] snmp-server name ***sysname***	optional, with

		default parameters
Display sysName configuration	show snmp name	optional
Configure maximum length of snmp protocol packets	[no] snmp-server max-packet-length ***length***	optional
Display the mib node information	show snmp mib [modeule ***module-name***]	optional

Note:

1. The device that does not use the configuration command of *snmp-server [enable | disable]* does not need to be configured. It is enabled by default and can not be disabled.

11.3 Configure the Community Name

SNMP adopts the community name authentication scheme. SNMP packets that do not match the community name will be discarded. SNMP community is named by a string, known as the community name. Different communities can have read-only or read-write access permission. A community with read-only access can only query system information. However, in addition to query the system information, the community with read-write access permission can perform the system configuration. It defaults to no community name.

Configure the Community Name

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure whether to display the community name encrypted or not	snmp-server community encrypt { enable | disable }	Optional; The default is not encrypted.
Configure the community name	snmp-server community name { ro | rw } { permit | deny } [view ***view-name***]	Optional; The iso view is used by default.
Display the community name	show snmp community	optional, with default parameter
Remove the community name	no snmp-server community ***community- index***	optional

Note:

1. The community name encryption function is irreversible. That is, after the encryption is configured, if the encryption function is disabled, the previously encrypted community will not become a plain text, and only the newly configured community will be encrypted.

11.4 Configure the Group

This configuration task can be used to configure an access control group. By default, there

are two snmpv3 groups: (1) The initial group with the security level of auth； (2) The initial group with the security level of noauthpriv(No authentication is required and no encryption is required)

Configure the Group

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure the group	snmp-server group ***group-name*** 3 [auth | noauth | priv] read ***read-view*** write ***write-view*** notify ***notify-view***	required
Configure the context of the group	[no]snmp-server group group-name 3 context ***context-name***	optional
Display the group configuration	show snmp group [***group-name***]	optional

📖 Note:

1. In the configuration control group, the write-view and the notify-view have no default values, so you must enter the view name in the configuration. *readview* defaults to *iso* and the security level defaults to *auth*.

11.5 Configure the User

It is used to configure the user for the local engine or for the remote engine that can be identified. By default, the following users exist: (1)initialmd5, (2) initialsha, (3) initialnone.

The above three users are reserved for the system and cannot be used by the user. When configuring a user, you need to ensure that the engine to which this user belongs is identifiable. When an identifiable engine is deleted, the users it contains are also deleted.

Configure the User

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure whether the display of password is encrypted or not	snmp-server encrypt { enable | disable }	Optional; The default is encryption
Configure the user	snmp-server user ***username groupname*** [remote ***ipaddress*** [udp-port ***port-number***]] [auth { md5 | sha } { auth-password ***authpassword*** | auth-key ***authkey*** } [priv des priv-key { auth-key ***privkey*** | auth-password ***privpassword*** }]]	required
Remove the user	no snmp-server user ***username*** [remote ***ipaddress*** [udp-port ***port***-number]]	optional
Display the user configuration	show snmp user [***username***]	optional

📖 Note:

1. remote ipaddress [udp-port port-number] : It means to configure the remote engine user. If you do not enter, it means to configure the local engine user. It defaults to local engine user. Port is the remote engine port number; if you do not enter, the default port number is 162.

3. There are three levels of user privilege levels: a.:noauthpriv (No authentication is required and no encryption is required), It is the default configuration； b. auth(Authentication is required but not encrypted)； c. authpriv(It requires authentication and encryption). The user's security level must be the same as the corresponding group security group.

11.6 Configure the Views

It is used to configure the views available to access control and the subtrees that they contain. The *iso*, *internet*, and *sysview* exist by default. Delete and modify the *internet* is not supported.

Configure the Views

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure the views	snmp-server view ***view-name oid-tree*** { included | excluded }	required
Remove the views	no snmp-server view ***view-name*** [***oid-tree***]	optional
Display the views configuration	show snmp view ***view-name***	optional

11.7 Configure SNMP Notification

Configure SNMP Notification

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure the source IP of the notification packets	[no]snmp-server trap-source { loopback-interface |vlan-interface | supervlan-interface } ***if-id***	optional
Enable notification function	[no]snmp-server enable [[informs | traps] [bridge | gbn | gbnsavecfg | interfaces | rmon | snmp]]	required
Display notification configuration	show snmp notify	optional
Configure to notify the destination host	[no]snmp-server host ***ipaddress*** [version {1 | 2c | 3 [auth | noauthpriv | priv] } ***security-name*** [udp-port ***port-number***] [notify-type [bridge | gbn | gbnsavecfg | interfaces | rmon | snmp]]	required
Display the configuration of notification host	show snmp host	optional

11.8 Configure Engine ID

It is used to configure the engine id of the local snmp and the engine id of the remote snmp. The local engine id defaults to 134640000000000000000000, and it can be modified but cannot be deleted. The remote engine id can be added and removed. By default, the remote engine id is not identified. Once an identifiable remote engine is deleted, its corresponding users will also be deleted. The maximum number of configurable remote engines is 32. This command uses the command of ***no*** to restore the default local engine id or delete the remote engine id.

Configure Engine ID

Operation	Command	Remarks

Enter global configuration mode	configure terminal	required
Configure engine id	snmp-server engineid { local ***engine-id*** | remote ***ipaddress*** [udp-port **port-number**] ***engine-id*** }	optional
Display engine id configuration	show snmp engineid { local | remote } [***id***]	optional
Remove engine id	no snmp-server engineid { local | remote ***ipaddress port-number*** }	optional

11.9 Configuration Example for Snmp

1.Network requirements
Before accessing the switch with the mib-browser, make sure that the mib-browser terminal is able to communicate with the switch properly.
a. Configure the community test2, and then make the mib-browser accesses the Switch through snmp v1 / v2;
b. Configure group name g3, user name u3, security levels are auth, make mib-browser access Switch through the snmp v3;
c. Configure snmp notice. Notify v2 and v3 respectively;

2.Configuration steps
Enable snmp（There is no need to configure a device without this command）
Switch(config)#snmp-server enable

Configure the community test2, mib-browser accesses the Switch through snmp v1 / v2
Switch(config)#snmp-server community test2 rw permit view iso

Configure the group name g3, the user name u3, the security levels are auth, mib-browser can access the Switch through snmp v3.
Switch(config)#snmp-server group g3 3 auth notify iso read iso write iso
Switch(config)#snmp-server user u3 g3 auth md5 auth-password password
Switch(config)#show snmp group g3
groupname: g3
securitymodel: 3 auth
readview: iso
writeview: iso
notifyview: iso
context: default value(NULL)

Switch(config)#show snmp user u3
User name: u3
Engine ID: 134640000000000000000000
Authentication Protocol: HMACMD5AuthProtocol
Group-name: g3
Validation: valid

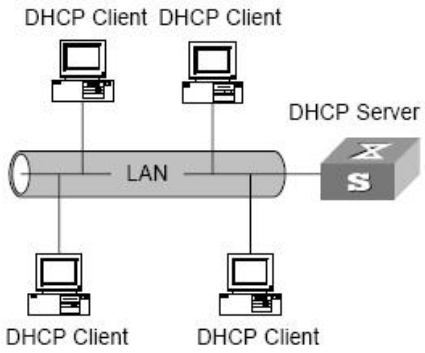
Configure the notification function
#Enable the notification function
Switch(config)#snmp-server enable traps
configure to notify the host
Switch(config)#snmp-server host 192.168.1.10 version 2 test2
Switch(config)#snmp-server host 192.168.1.10 version 3 auth u3

12.DHCP Configuration

12.1 DHCP Overview

As the increasing of network size and complexity of network, network configuration becomes more and more complicated. Computer position changes (laptop or wireless) and the number of computers over the distribution of IP addresses often occur. DHCP(Dynamic Host Configuration Protocol） develops to fulfill these needs. DHCP adopts Client/Server mode. DHCP client requests dynamically configuration from DHCP Server, DHCP Server sends back the related configuration based on the strategy.

In DHCP typical case, there is a DHCP server and multiple clients (PC or laptop), shown as follow.



DHCP typical case

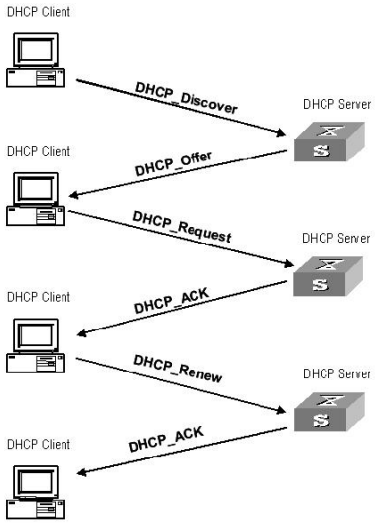
12.1.1 IP Address Allocation Strategy

DHCP provides 3 IP address strategies according to different requirements:

- IP address manual assigning: administrator binds fixed IP address statically for some particular clients(WWW server). Send configured fixed IP address to client by DHCP.
- IP address automatic assigning: DHCP assigns an IP address with an infinite lease to the client.
- IP address dynamic assigning: DHCP assigns a time-limited IP address to the client. After it is invalid, client re-applies the address. Most of clients get the IP address from this method.

12.1.2 IP Address Dynamic Acquisition Process

The packet interaction between DHCP Client and DHCP Server shown as follow.



message interaction between DHCP Client and DHCP Server

In order to acquire legal dynamic IP address, DHCP client exchanges different messages at different stages with the server. Usually there is 3 modes:

(1) Client logs in internet for the first time

DHCP client access the internet for the first time, there is 4 stages to connect the DHCP server.

- Discover stage. DHCP client discovers DHCP server. The client sends DHCP-discover packet in broadcast mode, only the DHCP server will response.
- Offer stage. DHCP server provides IP address. DHCP server receives DHCP-discover packet, and chooses an unallocated IP address to the client from the IP address pool and sends a DHCP-Offer packet containing the lease IP address and other settings to the client.
- Selection stage, that is to say the DHCP client chooses IP address. There are several DHCP servers send DHCP-Offer packet, the client only receives the first received DHCP-Offer then sends back DHCP-Request packet in broadcast mode to all DHCP servers. The packet includes the contents of the IP address requested from the selected DHCP server.
- Acknowledgement stage, that is to say DHCP server confirms IP address provided from DHCP server. When the DHCP server receives a DHCP-Request packet from the DHCP client, it sends a DHCP-ACK packet containing the IP address and other settings provided by the DHCP server to the client. Otherwise, the DHCP-NAK packet is returned, indicating that the address cannot be assigned to the client. When client receives the DHCP-ACK from the server, it will send

ARP (the destination address is the assigned address)in broadcast mode to do address detection. If there is no response in a certain time, the client uses this address.

Except the server that DHCP client chooses, other DHCP server's unallocated IP address could be used in other client's IP address apply.

(2) DHCP client logs in the internet more than one time

When DHCP client logs in the internet again, there are the main steps to connect the DHCP server.

- After the DHCP client logs in to the network for the first time, when it logs in network again later, it only needs to broadcast the DHCP-Request packet containing the IP address last assigned, and does not need to send the DHCP-Discover packet again.
- When DHCP server receives DHCP-Request packet, if the address applied by the client is not assigned, a DHCP-ACK packet is returned to inform the DHCP client that the previous IP address is to be used.
- If the IP address cannot be assigned to the DHCP client (already assigned to other client), DHCP server will return DHCP-NAK packet. When the client receives it, it will send again DHCP-Discover packet to request the new IP address.

(3) DHCP client extends the validity of lease for IP address.

The dynamic IP address assigned by the DHCP server usually has a certain lease duration. and the server reclaims the IP address after expiration. If the DHCP client wants to continue to use the IP address, the IP lease needs to be updated.

In actual use, DHCP client sends a DHCP-Request packet to the DHCP server to complete the update of the IP lease when the lease time of the IP address reaches half. If the IP address is valid, the DHCP server responds with a DHCP-ACK packet to inform the DHCP client that it has obtained a new lease.

12.1.3 DHCP Packet Structure

There are 8 types of packets and every packet carries the same form, but some fields value is different.

DHCP packet format is based on BOOTP packet format, the format is as shown as picture 1-3 (the number represents the bytes):

op(1)	htype(1)	hlen(1)	hops(1)
Xid(4)			
secs(2)		flags(2)	
ciaddr(4)			
yiaddr(4)			
siaddr(4)			
chaddr(16)			
sname(64)			
file(128)			
option(variable)			

DHCP packet structure

Explanation:

- op: DHCP packet operating type, is divided into request packet and response packet. 1 represents request packet, and 2 is response packet.
- htype, hlen: The hardware address type and length of the DHCP client.
- hops: The number of DHCP relays that DHCP packet passes through. Every time request packet from DHCP server passes through a DHCP relay, the field will increase 1.
- xid: The random number selected by the client when initiating a request, used to represent the process of one-time address request.
- secs: The time after the DHCP client starts a DHCP request.
- flags: The first byte is the broadcast response identifier. It is used to identify whether the DHCP server responds with unicast or broadcast. The remaining bits are reserved.
- ciaddr: IP address of DHCP client.
- yiaddr: IP address of client assigned by the DHCP server
- siaddr: server IP address of DHCP client for getting IP address
- giaddr: IP address of first DHCP relay that DHCP client request packet passed through
- chaddr: hardware address of DHCP client
- sname: server name of DHCP client getting IP address
- file: The name of the startup configuration file that the DHCP server assigns to the DHCP client.
- Option: Optional variable-length option field, including the type of the packet, valid tenancy term, DNS (Domain Name System) server IP address, WINS server IP address and so on.

12.2 DHCP Server Configuration

12.2.1 DHCP Server Application Environment

In the following cases, the DHCP server is usually used to complete the IP address assignment：

- Large-scale network, manual configuration requires a lot of work, and difficult to centrally manage the entire network
- The number of hosts in the network is greater than the number of IP addresses supported by the network. Not every host can be assigned a fixed IP address and there are restrictions on the number of users accessing the network at the same time (for example, the Internet access service provider). A large number of users obtain their dynamic IP addresses through the DHCP service.
- Only a few hosts in the network need a fixed IP address. Most hosts do not demand a fixed IP address.

12.2.2 DHCP Address Pool

The DHCP server selects and assigns IP addresses and other related parameters from the address pool to the clients. When an equipment as a DHCP server receives a DHCP request from a DHCP client, it selects an appropriate address pool according to the configuration and selects an idle IP address from it, and then send to the client together with other parameters such as the DNS server address and the lease duration of the address, etc.

The Prioritization of the DHCP server assigning IP addresses to clients from the address pool as follows:

- The address used by the clients
- Static binding IP address of the client’s MAC address in the DHCP server
- The address requested by the customers
- Addresses that are available in the address pool

12.2.3 Configure the Address Pool

According to the actual needs of the network, static address binding or dynamic address allocation can be chosen. Dynamic address allocation needs to specify the address range for

allocation, and static address binding needs to configure some MAC and IP corresponding binding table.

1. Configure the static binding for address allocation

Some clients (FTP servers, Web servers, etc.) need fixed IP addresses, which can be implemented by binding the MAC address of the client to the IP address. When a client with this MAC address requests an IP address, the DHCP server searches for the corresponding IP address based on the MAC address of the client and assigns the IP address to the client.

Static Binding for Address Allocation

Operation	Command	Explanation
Enter the global configuration mode	**configure terminal**	-
Enable static binding for address allocation	**dhcp-client bind**	optional
Enable to assign addresses for non binding users	**dhcp-client unknown-client assign**	Optioanal This function enables the non binding users to get dynamic IP address from IP address pool; otherwise, the non binding users couldn't get any IP address
Configure static binding list	**dhcp-client** *HH:HH:HH:HH:HH:HH A.B.C.D vid user_name*	optional

Note:
The "vid" and "user_name" parameters configured in the static binding table are used only to identify and record the client and will not be assigned to the client.

2、Configure the Dynamic Address Assignment

Configuration of the address pool range is demanded for addresses that are dynamically assigned to the client, including both permanent and lease-limited dynamic addresses. Up to eight IP address segments can be configured in the same address pool, and each address segment can hold up to 1024 IP addresses. Only one gateway is allowed in the same address pool (this gateway is used to determine the IP address range allocated, not the gateway address assigned to the DHCP client). The address in the address pool must be in the same network segment with the gateway.

The DHCP server needs to exclude the occupied IP addresses (such as gateways, FTP servers, and so on) when assigning addresses. Otherwise, assigning the same IP address to two clients will

cause IP address conflicts.

For different DHCP address pools, the DHCP server can specify different address lease duration, but the addresses in the same address pool have the same lease duration.

Dynamic Address Assignment

Operation	Command	Explanation
Enter the global configuration mode	**configure terminal**	-
Create address pool and enter address pool configuration mode	**ip pool** *ip-pool-name*	required
Configure the corresponding gateway in the address pool	**gateway** *ip(A.B.C.D) mask(A.B.C.D)*	required
Configure assignable address segment in the address pool	**section** <0-7> *start-ip end-ip*	required
Configure the lease of assignable address in the address pool	**lease** *ddd:hh:mm*,	optional
Configure an IP address in the DHCP address pool that does not participate in auto-allocation	**ip { enable | disable }** *A.B.C.D*	optional
Configure gateway address assigning for DHCP client	**router** *A.B.C.D*	optional

Note:
The gateway in command line of gateway *ip mask* refers to the gateway used for IP address allocation, which is the same as the value of the gateway field in DHCP request packets. It is used to determine the range of IP addresses to be allocated, not the gateway address assigned to the DHCP client. Assigning the gateway address to the DHCP client must be configured using the router *A.B.C.D* command.

12.2.4 Configure the DHCP Server to Assign the DNS Server Address

When a host accesses the Internet through a domain name, it needs to resolve the domain name to an IP address. This is achieved through the DNS (Domain Name System). In order for a DHCP client to successfully access the Internet through a domain name, the DHCP server ought to allocate an IP address to the client and in the meantime assign the DNS server address. At present, each address pool can be configured with up to four DNS server addresses.

On the DHCP server, you can specify a domain name for each address pool. When assigning an IP address to a client, the domain name would be sent to the client, too.

DHCP Server Supporting DNS Service

Operation	Command	Explanation
Enter the global configuration mode	**configure terminal**	-

Enter the address pool configuration mode	**ip pool** *ip-pool-name*	-
Configure a domain name for DHCP client	**dns suffix** *name*	optional
Configure DNS server address for DHCP client	**dns { primary-ip | second-ip | third-ip | fourth-ip }** *A.B.C.D*	optional

12.2.5 Configure the DHCP Server to Assign WINS server Addresses

For clients that use the Windows Microsoft operating system, WINS (Windows Internet Naming Service) servers provide host name-to-IP address resolution for hosts that communicate through the NetBIOS protocol. Therefore, most Windows network clients need to set WINS. Currently, each DHCP address pool can be configured with up to 2 WINS server addresses.

Configure the DHCP Server to Assign WINS server Addresses

Operation	Command	Explanation
Enter the global configuration mode	**configure terminal**	-
Enter the address pool configuration mode	**ip pool** *ip-pool-name*	-
Configure the WINS server address to be assigned to DHCP clients	**wins { primary-ip | second-ip }** *A.B.C.D*	optional

12.2.6 Configure the DHCP Customization Option

The contents of some options are not stipulated in RFC 2132. Manufacturers can define the contents of the options, such as Option 43. By configuring the DHCP customization option, you can provide the DHCP client with vendor-specified information. The IP address is assigned to the client and the contents of the custom option are also sent to the client.

Configure DHCP Self-defined Option 43

Operation	Command	Explanation
Enter the global configuration mode	**configure terminal**	-
Enter the address pool configuration mode	**ip pool** *ip-pool-name*	-
Configure DHCP self-defined option 43	**option 43 { ascii** *string* **| hex** *hexvalue* **}**	optional

12.2.7 Configure the DHCP Server to Support Option 60

The DHCP server supports the processing of DHCP packets with option 60 fields. After the option 60 is configured on a VLAN interface or super VLAN interface, when the interface receives a DHCP packet from the client, the option 60 field is matched with the value configured on the interface.

- If a match is found, the response information in the match is used as the content of the option 60 to the client.
- If no match is found or no response is configured in the match, the response packet to the client does not contain the option 60 field.

Configure DHCP Server Supporting Option 60

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enter VLAN interface configuration mode or Enter the super VLAN interface configuration mode	**interface vlan-interface** *vid* 或者 **interface supervlan-interface** *super-vid*	-
Configure interface option 60	**dhcp option60** { **equals** | **starts-with** } { **ascii** *string* | **hexadecimal** *hexdata* } **gateway** *A.B.C.D* [**dhcp-server** *group-id*] [**server-reply** { **ascii** *string* | **hexadecimal** *hexdata* }]	optional

12.2.8 Enable the DHCP Server Function

When the DHCP relay function is enabled, the device supports built-in DHCP server function. For the device to successfully assign IP addresses, the following requirements must be met:

（1）Enable the DHCP relay function;

（2）Configure the IP address of the device as the DHCP server IP;

（3）Configure the DHCP address pool correctly;

Refer to DHCP Relay Configuration for enabling DHCP relay function and configuring DHCP server IP.

12.2.9 DHCP Server Display and Maintenance

After completing the operations above, you can use the following command to view the current configuration and DHCP server status.

DHCP Server Display and Maintenance

Operation	Command	Explanation
show ip address pool, address section, gateway, assigned IP, ip lease, DNS server and WINS server information	**show ip pool** [*name*] **show ip pool** *name section-id* **show ip pool-brief**	Configure in any modes
show status of static binding allocation	**show dhcp-client bind**	
show static binding table	**show dhcp-client [ip** *A.B.C.D* **|MAC** *HH:HH:HH:HH:HH:HH* **]**	
show address information assigned to client	**show dhcp-server clients [ip [** *mask* **]]** **show dhcp-server clients [** *HH:HH:HH:HH:HH:HH* | *poolname*]	
show option 60 configured on interface	**show dhcp option60** [**interface** { **vlan-interface** *vid* | **supervlan-interface** *super-vid* }]	

12.3 DHCP Relay Configuration

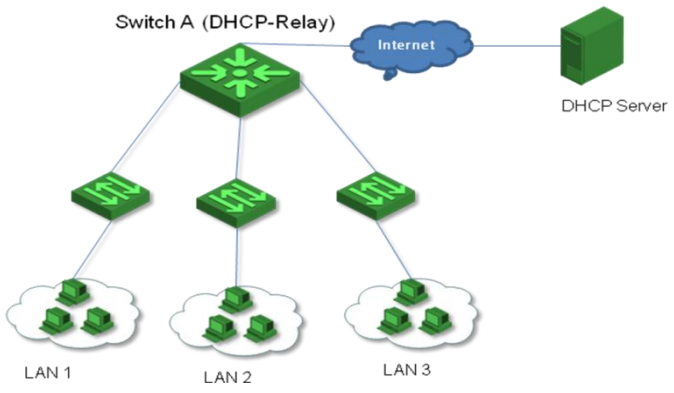
12.3.1 DHCP Relay Application Environment

DHCP is used only when the DHCP client and the server are on the same subnet, because IP packets are broadcasted during dynamic IP address acquisition. For dynamic host configuration, you need to set up DHCP servers on all network segments, which is obviously uneconomical.

DHCP relay function can solve this problem: the clients on the subnet can communicate with DHCP servers on other subnets through by DHCP relay, and finally get IP addresses. In this way, DHCP clients on multiple networks can use the same DHCP server, which saves costs and facilitates centralized management.

12.3.2 Basic Principles of DHCP Relay

Typical case of DHCP relay is shown as follow:



DHCP typical case of DHCP relay

The DHCP relay agent provides transparent transmission of DHCP broadcast packets and transparently transmits broadcast packets of DHCP clients (or servers) to DHCP servers (or clients) on other network segments.

During the process of dynamic configuration through DHCP relay, the DHCP client and the DHCP server are processed in the same way as when the DHCP relay agent is not used. The following describes only the forwarding process of the DHCP relay agent. For details about the packet exchange process, refer to section "1.2.2 IP Address Dynamic Acquisition Process".

The DHCP relay works as follows:

1) After receiving a DHCP-DISCOVER or DHCP-REQUEST packet, the DHCP relay agent fills the giaddr field with the IP address of the DHCP relay agent, and then forward unicast packets to the specified DHCP server according to the configuration.

2) The DHCP server assigns parameters such as the IP address to the client according to the giaddr field and forwards the configuration information to the client through the DHCP relay to complete the dynamic configuration of the client.

12.3.3 How DHCP Relay Agent handle DHCP packets

When the DHCP relay function is enabled on a device, the device will process the DHCP packets according to the following configuration when receiving DHCP packets from the DHCP client:

- Server mode: When the configured IP address of the DHCP server is the IP address of the device, the DHCP server function is enabled. Upon receiving the DHCP packet from the DHCP client, the address of the same address range is allocated in the address pool of the local DHCP server according to the "giaddr" field.
- Relay mode: When the configured IP address of the DHCP server is not the IP address of the device, it indicates that an external DHCP server is used. When receiving a DHCP packet from a DHCP client, it forwards the packet to an external DHCP server. The DHCP server assigns an address.

12.3.4 Configure DHCP Server Group

To improve reliability, multiple DHCP servers can be set in a network. Each DHCP server corresponds to a DHCP server group. When a DHCP server group is referenced by a VLAN interface or a super VLAN interface, DHCP packets from the client will be forwarded to all the servers in the server group.

Configure DHCP server IP

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Configure the DHCP server group	**dhcp-server** *group-id* **ip** *server-ip*	required
Enter VLAN interface configuration mode or Enter the super VLAN interface configuration mode	**interface vlan-interface** *vid* 或者 **interface supervlan-interface** *super-vid*	-
Configure the DHCP server group referenced by the interface	**dhcp-server** *group-id*	required

12.3.5 Configure DHCP Relay Agent to Support Option 60 Function

The DHCP relay agent supports DHCP packets with the option 60 option field. After the option 60 is configured on a VLAN interface or super VLAN interface, when the interface receives the DHCP packet from the client, if the packet contains the option 60 field, it will be matched with the value configured on the interface.

- If a match is found, the packet is relayed using the gateway address in the matched entry, and the DHCP packet is forwarded to the server address in the match.
- If no match is found, it will be relayed according to the requested IP address or the IP address of the client.

Configure DHCP Relay Agent to Support Option 60 Function

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enter VLAN interface configuration mode or Enter the super VLAN interface configuration mode	**interface vlan-interface** *vid* 或者 **interface supervlan-interface** *super-vid*	-
Configure interface option 60	**dhcp option60** { **equals** | **starts-with** } { **ascii** *string* | **hexadecimal** *hexdata* } **gateway** *A.B.C.D* [**dhcp-server** *group-id*] [**server-reply** { **ascii** *string* | **hexadecimal** *hexdata* }]	optional

12.3.6 Enable the DHCP Relay Function

If the DHCP server is not on the same subnet as the DHCP client, or the device is configured as a DHCP server, the DHCP relay function should be enabled.

Sometimes, for network security reasons, the network administrator does not want the DHCP client to know the address of the DHCP server. To meet this requirement, a DHCP relay

agent can be configured to hide the real DHCP server address. In this way, the DHCP client considers the DHCP relay agent to be the DHCP server, so as to hide the real DHCP server. Of course, if the DHCP relay device is the DHCP server, then this function is no longer applicable.

Enable DHCP Relay Function

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enable global DHCP relay function	**dhcp-relay**	required
Hide the IP address of the real DHCP server	**dhcp-relay hide server-ip**	optional
Set the maximum number of hops of DHCP packets	**dhcp max-hops** *hops*	optional
Enter the port configuration mode	**interface ethernet** *device/slot/port*	-
Enable interface DHCP relay	**dhcp-relay**	optional

12.3.7 DHCP Relay Display and Maintenance

After completed the preceding operations, you can use the following commands to check the configuration and the current status of the DHCP relay agent.

DHCP Relay Display and Maintenance

Operation	Command	Explanation
Show DHCP server group	**show dhcp-server** [*group-id*]	Configure in any mode
Show DHCP server group referenced by the interface	**show dhcp-server interface** [**vlan-interface** *vid* | **supervlan-interface** *super-vid*]	
Show DHCP relay status	**show dhcp-relay**	
Show the enable status of the hidden server function	**show dhcp-relay hide server-ip**	
Show option 60 configured on interface	**show dhcp option60** [**interface** { **vlan-interface** *vid* | **supervlan-interface** *super-vid* }]	

12.4 DHCP Snooping Configuration

12.4.1 DHCP-Snooping Overview

For the sake of security, the IP addresses used by online DHCP clients need to be tracked for the administrator to verify the corresponding relationship between the IP addresses the DHCP clients obtained from DHCP servers and the MAC addresses of the DHCP clients. Switches can track DHCP client IP addresses through the DHCP snooping function, which monitors DHCP

broadcast packets.

DHCP snooping monitors the following two types of packets to retrieve the IP addresses the DHCP clients obtain from DHCP servers and the MAC addresses of the DHCP clients:

- DHCP-ACK packet
- DHCP-REQUEST packet

When an unauthorized DHCP server exists in the network, a DHCP client may obtains an illegal IP address. To ensure that the DHCP clients obtain IP addresses from valid DHCP servers, you can specify a port to be a trust port or an untrusted port by the DHCP snooping function:

- Trusted ports can be used to connect DHCP servers or ports of other Switches. Untrusted ports can be used to connect DHCP clients or networks.
- Untrusted ports drop the DHCP-ACK and DHCP-OFFER packets received from DHCP servers. Trusted ports forward any received DHCP packets to ensure that DHCP clients can obtain IP addresses from valid DHCP servers.
- Trusted vlan: untrusted port will not drop the DHCP-ACK and DHCP-Offer.

12.4.2 DHCP Snooping Configuration

Configure DHCP Snooping

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enable DHCP Snooping	**dhcp-snooping**	required
Enter vlan configuration mode	**vlan** *vid*	-
Enable DHCP Snooping of VLAN	**dhcp-snooping**	optional
Configure the corresponding vlan to be trust vlan	dhcp-snooping trust	optional
Enter port configuration mode	interface ethernet port_id (device/slot/port)	-
Specify the port connected to the DHCP server as a trust port	**dhcp-snooping trust**	required

Note:

If an aggregation group is used in an application, you must ensure that the configuration of each member port in the aggregation group is consistent. Otherwise, the packet processing exception will be caused. For example, the uplink port requires to use the aggregation group, aggregation group has a total of four member ports, you must manually configure the four member ports to be trust ports. All other port configurations are in accordance with this requirement.

12.4.3 Configure Link-Down Operation

When the connection link is down, you can perform the following operation on the dynamic table which Dhcp-snooping has learned:

- enable fast-remove, delete Dhcp-snooping dynamic table immediately when the port is down.
- disable fast-remove， normally aging the dynamic table according to the tenancy term instead of deleting the Dhcp-snooping dynamic table immediately when the port is down.

Configure Link-Down Operation

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Configure link-down operation of the port	**dhcp-snooping port-down-action fast-remove**	optional

12.4.4 Configure Max Clients Number

If the attacker exists, it will disguise as multiple users to ask DHCP Server for address to use up the Server allocable address. As a consequence, Server has no address to allocate to the user who needs the IP address. For this problem, network administrator can take the following measures:

1) Restrict the DHCP-Client number connected to Switch port. In this case, only the clients connected to the same port with the attacker will suffer the attack.

2) Restrict the DHCP-Client number in specified VLAN. In this case, only the clients in the same VLAN with the attacker will suffer the attack.

Configure max clients number

Operation	Command	Explanation
Enter port configuration mode	**interface ethernet** *port_id*	-
Configure max DHCP-Client number connected to Switch port	**dhcp-snooping max-clients** *<0-2048>*	optional
Enter vlan configuration mode	**vlan** *vlan_list*	-
Configure max DHCP-Client number in specified VLAN.	**dhcp-snooping max-clients** *<0-2048>*	optional

12.4.5 IP-Source-Guard Overview

IP-Source-Guard is in the layer-2 used to limit IP traffic based on DHCP Snooping-monitoring table and manual configuration IP-source-binding. IP-Source-Guard can be able to prevent the traffic attack caused by neighbor IP embezzled.

After enabling IP-Source-Guard, Switch will stop the untrust port receiving the IP flow which is different from DHCP Snooping-monitoring table or IP-source-binding information. The IP flow includes source MAC address, source IP address, source port number, VLAN.

Configure IP-Source-Guard

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Configure the ip-source-bind ip	**ip-source-guard bind ip** *A.B.C.D* [**mac** *HH:HH:HH:HH:HH:HH* **interface ethernet** *device/slot/port* **vlan** *vid*]	optional
Configure the IP-Source-Guard VLAN list	**ip-source-guard vlan** *vlan-list*	optional
Configure IP-Source-Guard to allow IGMP packets to pass through	**ip-source-guard permit igmp**	optional
Enter interface configuration mode	**interface ethernet** *device/slot/port*	-
Enable the IP-Source-Guard function on the untrusted port and specify the port filtering mode	**ip-source-guard [ip | ip-mac | ip-mac-vlan]**	optional

Note:

The network administrator can flexibly configure the IP source binding table according to the actual situation. The IP source binding table supports two binding modes:

- Source IP
- Source IP+ Source MAC+ Source port +VLAN

When you enable the IP-Source-Guard function, you can also specify the port filtering mode. There are three types of port filtering mode:

- ip: The port only filters packets based on the source IP address of the ip packet, regardless of the source MAC address and the VLAN ID;
- ip-mac: The port filters packets based on the source IP address of the ip packet and the mac address, regardless of the vlan;
- ip-mac-vlan: The port filters packets based on source IP, MAC and vlan;

If you do not specify the port filtering mode when the IP-Source-Guard function is enabled, the effective port filtering mode is ip-mac-vlan, that is, the effect of the ip-source-guard command is the same as that of the ip-source-guard ip-mac-vlan command.

12.4.6 DHCP Snooping Display and Maintenance

After completing the above configurations, you can use the following commands to view the related information.

- DHCP Snooping Display and Maintenance

Operation	Command	Explanation
Display the mapping relationship between the MAC address and the user IP address recorded by the DHCP Snooping	**show dhcp-snooping clients**	It can be viewed in any mode
Display DHCP Snooping state (enable/disable) , trusted port information, number of DHCP clients allowed on the physical port and number of DHCP clients that are currently connected	**show dhcp-snooping interface** [*port-list*]	
Display the DHCP snooping state (enable/disable) and the number of DHCP clients that belong to the specified VLAN	**show dhcp-snooping vlan** [*vid*]	
Display the status of the IP-Source-Guard function on a port (enable/disable)	**show ip-source-guard**	
Display the source IP binding table of IP-Source-Guard	**show ip-source-guard bind** [**ip** *A.B.C.D*]	
Display the IP-Source-Guard VLAN information	**show ip-source-guard vlan**	

Configure IP-Source-Guard to allow IGMP packets to pass through	**show ip-source-guard permit igmp**	
Remove the dynamic entries recorded by DHCP snooping	**clear dhcp-snooping** [**ip** *A.B.C.D* **|** **mac** *H:H:H:H:H:H* **|** **vlan** *vid* **|** **interface ethernet** *port-id*]	It can only be removed in global configuration mode

12.4.7 Configuration Example for DHCP Snooping

- Network requirements

As shown below:

Ethernet0/ 0/1 is connected to the DHCP server and Ethernet0 /0/ 2 is connected to the DHCP client network. Enable DHCP snooping on the Switch. Set the port Ethernet0 / 0/1 of the Switch to be a trust port.

- network diagram

Configuration Example for DHCP Snooping

- Configuration steps

The following configuration is performed on the switch that serves as the DHCP snooping device.

(1) Enter global configuration mode

Switch# configure terminal

Switch(config)#

(2) Enable DHCP Snooping

Switch(config)#dhcp-snooping

Config DHCP Snooping successfully.

(3) Enter the interface configuration mode of Ethernet0/0/1

Switch(config)#interface ethernet 0/0/1

(4) Configure Ethernet0 / 0/1 as the trusted port

Switch(config-if-ethernet-0/0/1)#dhcp-snooping trust

Config DHCP Snooping mode of port successfully.

12.5 DHCP Option 82 Overview

Option: A length-variable field in DHCP packets, carrying information such as part of the lease information and packet type. It includes at least one option and at most 255 options.

Option 82: Also known as relay agent information option. This option is a part of the Option field in DHCP packet. According to RFC3046, option 82 lies before option 255 and after the other options. Option 82 includes at least one sub-option and at most 255 sub-options. Currently, the commonly used sub-options in option 82 are sub-option 1 and sub-option 2.

Sub-option 1: A sub-option of option 82. Sub-option 1 represents the agent circuit ID, namely Circuit ID. It holds the port number and VLAN-ID of the Switch port connected to the DHCP client, and is usually configured on the DHCP relay. Generally, sub-option 1 and sub-option 2 must be used together to identify the information of DHCP source.

Sub-option 2: A sub-option of option 82. Sub-option 2 represents the remote agent ID, namely Remote ID. It holds the MAC address of the DHCP relay, and is usually configured on the DHCP relay. Generally, sub-option 1 and sub-option 2 must be used together to identify information about a DHCP source.

12.5.1 Configure DHCP Option82

The DHCP Option 82 function must be used together with DHCP relay or DHCP snooping.

When the Switch receives a DHCP packet with the Option 82 field, the following three strategies are supported:

- drop: Drop all DHCP messages carrying the Option 82 field.
- keep: Keep the Option 82 field in the packet and forward it.
- replace: According to the local situation, replace the original Option 82 with the new option 82 and forward it.

Enable DHCP Option82

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enable DHCP Option82	**dhcp option82**	required
Configure DHCP option82 format	**dhcp option82 format** { **normal | verbose | henan** }	optional
Configure DHCP option82 format to be verbose node-identifier	**dhcp option82 format verbose node-identifier { mac | hostname | user-defined** *node-id* }	optional

Enter interface configuration mode	**interface ethernet** *port-id*	-
# Configure the Switch to process DHCP messages carrying the Option 82 field	**dhcp option82 strategy { drop | keep | replace | append { hostname | hostname-ip } }**	optional
Configure the circuit-id of DHCP option82	**dhcp option82 circuit-id string** *id*	optional
Configure the remote-id of DHCP option82	**dhcp option82 remote-id string {** *string* **| hostname }**	optional

12.5.2 DHCP Option82 Display and Maintenance

DHCP Option82 Display and Maintenance

Operation	Command	Explanation
Display DHCP Option 82 state (enabled/disable) and the processing mode on DHCP messages carrying the Option 82 field	**show dhcp option82** [**interface** *port-list*]	Run in any mode

12.6 DHCPv6 Snooping Overview

For security reasons, network administrator might need to record the IPv6 address what user uses when surfing the internet so as to confirm the corresponding relationship between user MAC address and IPv6 address obtaining from DHCPv6 Server. Switch supports to monitor DHCP message via DHCPv6 Snooping function, recording users IP address information.

Security mechanism of DHCPv6 Snooping allows configuring the port as trust port or untrust port:

- Trust port connects with DHCPv6 server or the ports of other Switches; untrust port connects with user or network.
- Untrust port discards the DHCPv6-reply message and DHCPv6-advertise message which from DHCPv6server; trust port will forward the DHCP message so as to ensure the user to obtain the correct IPv6 address.

12.6.1 Configure DHCPv6 Snooping

Configure DHCPv6 Snooping

Operation	Command	Explanation

Enter global configuration mode	**configure terminal**	-
Enable global DHCPv6 Snooping	**dhcpv6-snooping**	required
Enter interface configuration mode	**interface ethernet** *port_id (device/slot/port)*	-
Specify the port which connects with DHCPv6 Server to be trust port	**dhcpv6-snooping trust**	required

12.6.2 Configure the Port Operation When Link Down

When port link is down, the dynamic items learned by DHCPv6 Snooping could be dealt with as follows:

- Enable fast-remove—Delete dynamic items learned by DHCPv6 Snooping immediately.
- Disable fast-remove—Do not delete dynamic items learned by DHCPv6 Snooping immediately but aging table according to rental agreement.

Configure the port operation when link down

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Configure the port operation when link is down	**dhcpv6-snooping port-down-action fast-remove**	optional

12.6.3 Limit DHCPv6-Client Number Accessed to the Port

If there is such a network attacker, who disguised as multiple user request ip addresses from the DHCPv6 Server, the addresses that can be assigned are exhausted, so that the server has no more address to assign to the users. For such attacks, the network administrator can do the follows:

- Limit the number of DHCPv6 Client collected to Switch physical port. This will limit the impact of the attack and the attacker range of users connected to the same physical port, rather than the user of the whole network.
- Limit the number of DHCPv6 Client accessed to a specified VLAN. Limit the attack to the same range of users as the attacker belongs to, rather than the entire network of users. This feature requires taking effect together with DHCPv6 Snooping.

This function should be used together with DHCPv6 Snooping

Configure physical port and vlan allow DHCPv6 Client number of connections

Operation	Command	Explanation
Enter port configuration mode	**interface ethernet** *port_id*	-

Configure the number DHCPv6 Client accessed to physical port according to network status	**dhcpv6-snooping max-clients** *<0-2048>*	optional
Enter VLAN configuration mode	**vlan** *vlan_list*	-
Configure the number DHCPv6 Client belongs to specific vlan according to network status	**dhcpv6-snooping max-clients** *<0-2048>*	optional

12.6.4 Configure IPv6-Source-Guard

IPv6-Source-Guard is a kind of security feature which is used to limit IP traffic. You can use IPv6-Source-Guard to prevent traffic theft caused by the theft of neighbor IP.

After enabling IPv6-Source-Guard, Switch will prevent the untrust port from receiving the IPV6 traffic which differs from DHCPv6 Snooping monitor table or IPv6 source binding table. This information includes: source MAC address, source IPV6 address, source port number, VLAN.

Configure IPV6-Source-Guard

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Configure IPV6 source binding table	**ipv6-source-guard bind ip** *A.B.C.D* [**mac** *HH:HH:HH:HH:HH:HH* **interface ethernet** *device/slot/port* **vlan** *vid*]	optional
Configure IPV6-Source-Guard VLAN list	**ipv6-source-guard vlan** *vlan-list*	optional
enter interface configuration mode	**interface ethernet** *device/slot/port*	-
Enable the IPV6-Source-Guard function of untrust port and then specify filter approach of the port	**ipv6-source-guard**	optional

Note:

Network administrators can flexibly configure IPv6 source binding table according to the actual situation. IPv6 source binding table supports four bound manners:

- Source IPv6 address

- Source IPv6 address+source MAC
- Source IPv6 address+source MAC+source port
- Source IPv6 address+source MAC+source port+VLAN

If you want to match the table with mac, just configure the acl matching approach as ipv6-acl-key-mode mac

12.6.5 DHCPv6 Snooping Display and Maintenance

After finishing the above configurations, the commands below could be used to display relevant information.

DHCPv6 Snooping display and maintenance

Operation	Command	Explanation
Display the mapping between IPV6 address and MAC address recorded by DHCPv6 snooping	**show dhcpv6-snooping clients**	It supports to display in any mode.
Display the DHCPv6 snooping enable status, trusted port information, the number of DHCPv6 clients allowed to access, and the number of DHCPv6 clients that are currently connected	**show dhcpv6-snooping interface** [*port-list*]	
Display the DHCPv6-snooping enable/disable state and the number of DHCPv6 clients allowed to belong to the specified VLAN	**show dhcpv6-snooping vlan** [*vid*]	
Display the IPv6-Source-Guard status	**show ipv6-source-guard**	
Show the IPv6-Source-Guard's sourceIPv6 binding table	**show ipv6-source-guard bind** [**ipv6** *A.B.C.D*]	
Display the IPV6-Source-Guard VLAN information	**show ipv6-source-guard vlan**	
Delete the dynamic entries recorded by DHCPv6 snooping	**clear dhcpv6-snooping** [**ip** *A.B.C.D* **|** **mac** *H:H:H:H:H:H* **|** **vlan** *vid* **|** **interface ethernet** *port-id*]	

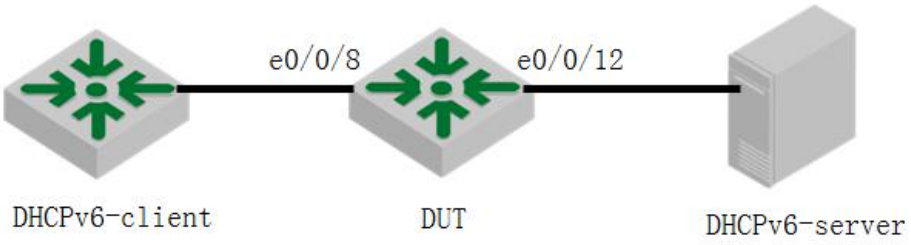
12.6.6 Configuration Example for DHCPv6 Snooping

- Network requirements

As shown below:

Etherent0/0/12 of DUT connects with DHCPv6 Server，and Ethernet0/0/8 of DUT connects with the network of DHCPv6 Client; DUT enables DHCPv6 Snooping function. Set Ethernet0/0/12 of DUT to be trust port.

- network diagram



Configuration Example for DHCPv6 Snooping

- Configuration steps

Perform the following configuration on the DUT which acts as the DHCPv6 snooping device.

(1) Enter global configuration mode

Switch # configure terminal

Switch (config)#

(2) Enable DHCPv6 Snooping

Switch (config)#dhcpv6-snooping

Config DHCPv6 Snooping successfully.

(3) Enter Ethernet0/0/12 port configuration mode

Switch (config)#interface ethernet 0/12

(4) Set Ethernet0/0/12 to be trusted port

Switch (config-if-ethernet-0/1)#dhcpv6-snooping trust

Config DHCPv6 Snooping mode of port successfully.

- Result Validation

Switch(config)#sh dhcpv6-snooping clients

DHCPv6 client information:

d - days, h - hours, m - minutes, s - seconds

IPv6Address	mac	vlan	port	Lease
2001::B4F8:72C6:C276:A4C8	00:00:00:0f:11:bd	1	e0/8	1h52m43s

12.7 DHCPv6 Option 18 and DHCPv6 Option 37

DHCPv6 Option 18 is an option of interface-id. Relay agent can be able to send interface id to identify the port which has received the information from client side. If the relay agent receives the relay-relay information with interface-id option, relay agent will relay this information to client side via the interface with the interface-id option.

DHCPv6 Option 37 is Remote ID. Generally, this option should be configured on DHCP relay device, asking for carrying the hostname of relay device or IPv4/IPv6 address information in the process of transmission.

12.7.1 DHCPv6 Option18 Configuration and Display

DHCPv6 Option18 should be used together with DHCPV6 Relay or DHCPV6 Snooping.

Enable DHCPV6 Option18

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enable dhcpv6-snooping information	**dhcpv6-snooping information option** [18|37]	required
Enable DHCPV6 Option18	**dhcpv6-snooping information option 18**	required
Display DHCPv6 Option18	**show dhcpv6-snooping information**	

12.7.2 DHCPv6 Option37 Configuration and Display

Enable DHCPV6 Option37

Operation	Command	Explanation
Enter global configuration mode	**configure terminal**	-
Enable dhcpv6-snooping information	**dhcpv6-snooping information option** [18|37]	required
Enable DHCPV6 Option37	**dhcpv6-snooping information option 37**	required

Configure remote id	**dhcpv6-snooping information remote-id [hostname | ipv4-address | ipv6-address | string]**	required

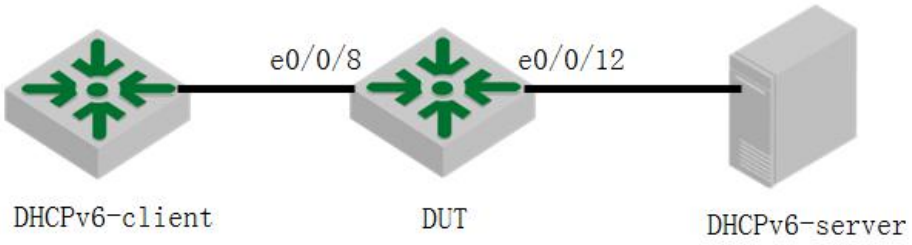
12.7.3 Configuration Examples of DHCPv6 Option18 and DHCPv6 Option37

- network requirement

As shown below:

Etherent0/0/12 of DUT connects with DHCPv6 Server, Ethernet0/0/8 connects with network of DHCPv6 Client side; enable DHCPv6 Snooping on DUT → configure Ethernet0/0/12 to be trusted port→ DUT enables option 18.

- network diagram



DHCPv6 Option18 configuration example

Perform the following configurations on the DUT which acts as the DHCPv6 snooping device.

Perform the following configuration based on 2.6 configurations:

Enter global configuration mode

Switch #configure terminal

Switch (config)#

Enable dhcpv6 option18

Switch(config)#dhcpv6-snooping information option 18

- Result Validation

Capture packet on DUT Ethernet0/0/12, you can find that both dhcpv6 solicit packet and dhcpv6 request packet carry Option 18 field.

DHCPv6 Option37 configuration and Option18 configuration are basically identical. However, it asks to configure remote-id field in DUT.

Please refer to the configuration table of DHCPv6 Option37.

13.ACL Configuration

In order to filter data packet, it needs configuring a series of matching rules to identify the object which needs filtration. After recognizing special object, it can configure to permit or deny corresponding data packet passing according to the scheduled strategy. Access Control List (ACL) is used to realize this function.

ACL classifies packets according to a series of matching conditions, which can be the source address, destination address, and port number of a packet and so forth. A Switch or an xPON device detects packets based on the conditions specified in the ACL to determine whether to forward or drop the packet.

Data packet matching rules defined by ACL can be introduced to other situation which needs distinguish flow, such as the flow classification in QoS.

According to the purpose of application, ACL can be divided into the following categories:

Standard ACL: Defines rules based on source IP addresses only.

Extended ACL: Defines rules based on the source IP address, destination IP address, protocol type, and protocol attributes of packets.

Layer 2 ACL: Layer 2 information such as source MAC address, destination MAC address, VLAN priority, and Layer 2 protocol type.

Both switches and ePON products support standard ACLs, extended ACLs, and Layer 2 ACLs.

13.1 ACL Matching Order

Since the same ACL can configure multiple sub-items, so there is a problem of matching order. ACLs support two match orders:

config: Match ACL rules according to the configuration order

auto: Match ACL rules according to the depth-first rule.

Depth-first means that the longest sub-item matches first.

ACL Matching Order

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the ACL matching order	access-list *access-list* match-order {auto| config}	Optional; Config by default.

13.1.1 Configuration Example for ACL Matching Order

Device(config)#access-list 1 deny any // Configure two sub-items of the same ACL
Config ACL subitem successfully.

Device(config)#access-list 1 permit 1.1.1.1 0
Config ACL subitem successfully.

a. If the match order is the configuration order, configure sub-item to be 0 first. The configurations are as follow:

Device(config)#show access-list config 1 // Configuration order by default
Standard IP Access List 1, match-order is config, 2 rule:
0 deny any
1 permit 1.1.1.1 0.0.0.0

b . If the match order you choose is automatic, the longest ACL matching rule will be sub-item 0. The configurations are as follow:

Device(config)#access-list 1 match-order auto // Set the matching order to automatic order
Config ACL match order successfully.

Device(config)#access-list 1 deny any
Config ACL subitem successfully.

Device(config)#access-list 1 permit 1.1.1.1 0
Config ACL subitem successfully.

Device(config)#show access-list config 1
Standard IP Access List 1, match-order is auto, 2 rule:
0 permit 1.1.1.1 0.0.0.0
1 deny any
The device complies with the “first activation takes precedence” matching rule.

13.2 Standard ACLs

The standard ACL only performs the corresponding analysis and processing on the packets according to the rules specified in the Layer 3 source IP address.

If the standard ACL is identified by numbers, the sequence number ranges from 1 to 99. Up to 99 standard ACLs can be created; If the basic ACL is identified by names, up to 1000 entries can be defined. At the same time, Switch or Xpon can define up to 128 sub-rules for each ACL.

If you want to configure a rule with time range parameters, you need to define the corresponding time range first. For the configurations of time-range, refer to 2.6.

ACL identified by numbers

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Define an ACL	**access-list** *num* { **permit** | **deny** } { *source-IPv4/v6 source-wildcard* | **any** | **ipv6any** } [**time-range** *name*]	required
Remove the ACL based on number	**no access-list** [*num* *subitem* |all]	*All* refers to all ACLs

ACL identified by names

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure a standard ACL based on the name and enter ACL configuration mode	**access-list standard** *name*	required
Configure an ACL rule	**{ permit | deny } {** *source-IPv4/v6 source- wildcard* **| any | ipv6any } [time-range** *name* **]**	required

Remove the ACL based on nam	**no access-list** [*name* | subitem] all	

description of standard ACL rules

permit | deny	ACL scope rules	***permit*** means to allow access, ***deny*** means to deny access
{ *source-IPv4/v6 source-wildcard* | **any** | **ipv6any** }	Specify the source address of the ACL rule	*source-IPv4/v6sour-wildcard* is used to determine the source IP address (IPv4 / v6) range of the packet. IPv4 addresses are represented in dotted decimal notation ； IPv6 addresses are represented in hexadecimal； When *source-wildcard* is 0, it indicates the host address； *any*| *ipv6any* refers to any source address.
{ dest-IPv4/v6 dest-wildcard| **any** | **ipv6any** }	Specify the destination address of the ACL rule	*dest-IPv4/v6 dest-wildcard* is used to determine the destination IP address (IPv4 / v6) range. IPv4 addresses are represented in dotted decimal notation ； IPv6 addresses are represented in hexadecimal； When *source-wildcard* is 0, it indicates the host address； *any*| *ipv6any* refers to any source address.
time-range *name*	Specifies the time range in which the rule takes effect	Please refer to 2.6

13.2.1 Configuration Example

// Define a basic ACL identified with numbers to forbid the packets whose source IP address is 10.0.0.1

Device# configure terminal
Device(config)#access-list 1 deny 10.0.0.1 0
Config ACL subitem successfully.

// Define a basic ACL identified with names to forbid the packets whose source IP address is 10.0.0.2

Device(config)#access-list standard stdacl
Create ACL item successfully.

Device(config-std-nacl-stdacl)#deny 10.0.0.2 0
Config ACL subitem successfully.

13.3 Extended ACL

An extended ACL can make rules based on information such as the source IP address, destination IP address, protocol type, protocol characteristics, and so on.

If the extended ACL is identified by numbers, the sequence number ranges from 100 to 199. Up to 100 extended ACLs can be created. If the extended ACL is identified by names, up to 1000 entries can be defined. At the same time, SWITCH or Xpon can define up to 128 sub-rules for each ACL.

If you want to configure a rule with time range parameters, you need to define the corresponding time range first. For the configuration of time range, refer to 2.6.

ACL identified by numbers

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Define an extended ACL	**access-list** *num* { **permit** | **deny** } [*protocol*] [**established**] { *source-IPv4/v6 source-wildcard* | **any** | **ipv6any** } [*source-port wildcard*] { *dest-IPv4/v6 dest-wildcard* | **any** | **ipv6any** } [*dest-port wildcard*] [*icmp-type icmp-code*] [*igmp-type*] [**traffic-class** *traffic-class*][**precedence** *precedence*] [**tos** *tos*] | [**dscp** *dscp*] [**fragments**][**time-range** *name*]	required
Remove the ACL based on number	**no access-list** [*num* | all }	*All* refers to all ACLs

description of extended ACL rules

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure an extended ACL based on the name and enter ACL configuration mode	**access-list extended** *name*	required
Configure an ACL rule	{ **permit** | **deny** } [*protocol*] [**established**] { *source-IPv4/v6 source-wildcard* | **any** | **ipv6any** } [*source-port wildcard*] { *dest-IPv4/v6 dest-wildcard* | **any** | **ipv6any** } [*dest-port wildcard*] [*icmp-type icmp-code*] [*igmp-type*] [**traffic-class** *traffic-class*][**precedence** *precedence*] [**tos** *tos*] | [**dscp** *dscp*] [**fragments**][**time-range** *name*]	required
Remove the ACL based on name	**no access-list** [*name* | all }	*All* refers to all ACLs

The detailed parameters in the extended ACL are described in the following table.

description of extended ACL rules

Parameter	Function	Explanation
protocol	IP protocol type	It is in the range of 1 to 255 by number. By name, you can select GRE, ICMP, IGMP, IPinIP, OSPF, TCP, UDP, ICMPv6.
established	SYN flag in TCP	SYN = 1 is active
{ *source-IPv4/v6*	Specify the source	*source-IPv4/v6sour-wildcard* is used to determine

source-wildcard | **any** | **ipv6any** }	address of the ACL rule	the source IP address (IPv4 / v6) range of the packet. IPv4 addresses are represented in dotted decimal notation ； IPv6 addresses are represented in hexadecimal； When *source-wildcard* is 0, it indicates the host address； *any*| *ipv6any* refers to any source address.。
{ dest-IPv4/v6 dest-wildcard| **any** | **ipv6any** }	Specify the destination address of the ACL rule	*dest-IPv4/v6 dest-wildcard* is used to determine the destination IP address (IPv4 / v6) range. IPv4 addresses are represented in dotted decimal notation ； IPv6 addresses are represented in hexadecimal； When *source-wildcard* is 0, it indicates the host address； *any*| *ipv6any* refers to any source address.
source-port/ dest-port wildcard	TCP/UDP source and destination port numbers	*Wildcard*—The inverse number determines the range of port numbers
icmp-type icmp-code	ICMP protocol packet type	It is valid only when the protocol is configured as icmp / icmpv6
igmp-type	IGMP protocol packet type	It is valid only when the protocol is configured as igmp
traffic-class	**traffic-class** in Ipv6	Available for ipv6 messages only
precedence	precedence priority	The IP precedence ranges from 0 to 7
tos	tos priority	The value ranges from 0 to 15
dscp	DSCP priority	The value ranges from 0 to 63
fragments	fragment packets	The rule is valid only for non-first fragmented packets

13.3.1 Configuration Example for Extended ACL

// Define a number-based extended ACL to deny FTP packets whose source IP address is 10.0.0.1.

Device(config)#access-list 100 deny tcp 10.0.0.1 0 ftp any
Config ACL subitem successfully.

// Define a name-based extended ACL to deny FTP packets whose source IP address is 10.0.0.2.

Device(config)#access-list extended extacl
Create ACL item successfully.

Device(config-ext-nacl-extacl)#deny tcp 10.0.0.2 0 ftp any
Config ACL subitem successfully.

13.4 Layer2 ACL

Layer 2 ACLs can be configured based on Layer 2 information such as source MAC address, destination MAC address, VLAN priority, and Layer 2 protocol type.

If Layer 2 ACL is identified by numbers, the number ranges from 200 to 299. You can create up to 100 Layer 2 ACLs identified by numbers； If the Layer 2 ACL is identified by names, up to 1000 entries can be defined. At the same time, Switch/Xpon can define up to 128 sub-rules for each ACL.

If you want to configure a rule with time range parameters, you need to define the corresponding time range first. For the configuration of time range, refer to 2.6.

ACL identified by numbers

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Define Layer 2 ACL	**access-list** *num* { **permit** | **deny** } [*protocol*] [**cos** *vlan-pri*] **ingress** { { [**inner-vid** *vid*] [*start-vlan-id end-vlan-id*] [*source-mac-addr source-mac-wildcard*] [**interface** *interface-num*] } | **any** } **egress** { { [*dest-mac-addr dest-mac-wildcard*] [**interface** *interface-num* | **cpu**] } | **any** } [**time-range** *name*]	required
Remove ACL	**no access-list [num | all }**	*All* refers to all ACLs

ACL identified by names

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Define Layer 2 ACL and enter ACL configuration mode	**access-list link** *name*	required
Configure ACL rule	{ **permit** | **deny** } [*protocol*] [**cos** *vlan-pri*] **ingress** { { [**inner-vid** *vid*] [*start-vlan-id end-vlan-id*] [*source-mac-addr source-mac-wildcard*] [**interface** *interface-num*] } | **any** } **egress** { { [*dest-mac-addr dest-mac-wildcard*] [**interface** *interface-num* | **cpu**] } | **any** } [**time-range** *name*]	required
Remove ACL	no access-list [name | all }	*All* refers to all ACLs

description of layer2 ACL rules

Parameter	Function	Explanation
protocol	Protocol type carried by the Ethernet frame	In hexadecimal notation, range 0 to FFFF. Optional for arp, ip, rarp
Cos	The priority of the Vlan tag	
Ingress	Ingress direction	
inner-vid	The inner vid value of a double-tagged packet	
start-vlan-id *end-vlan-id*	It is used to indicate the range of VLANs	As for double tag packet, it is the vid range of the outer tag ; as for single tag packet, it is the vid

		range of the tag itself.
source-mac-addr *source-mac-wildcard*	Source mac address options	The source-mac-wildcard can be used to indicate the source mac range.
interface *interface-num*	The physical port number	Divided into ingress port and egress port
CPU		Indicates the data will be forwarded to the CPU
any	Any address	Divided into ingress direction and egress direction

13.4.1 Configuration Example for Layer2 ACL

// Define a Layer 2 ACL that is identified by number, and disable the ARP packets whose source MAC address is 00: 00: 00: 00: 00: 01.

Device(config)#access-list 200 deny arp ingress 00:00:00:00:00:01 0 egress any
Config ACL subitem successfully.

// Define a Layer 2 ACL that is identified by name, and disable the ARP packets whose source MAC address is 00: 00: 00: 00: 00: 02.

Device(config)#access-list link lnkacl
Create ACL item successfully.

Device(config-link-nacl-lnkacl)#deny arp ingress 00:00:00:00:00:02 0 egress any
Config ACL subitem successfully.

13.5 Time Range

The configuration of the time range includes the periodic time range and absolute time range. Configuring a periodic time period takes the form of the day of the week. The absolute time range is configured from the start time to the end time.

Configure the time range

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Create a time range and enter the time range configuration mode	**time-range** *name*	-
Configure an absolute time range	**absolute start** *HH:MM:SS YYYY/MM/DD* [**end** *HH:MM:SS YYYY/MM/DD*]	required
Configure a periodic time range	**periodic** *days-of-the-week hh:mm:ss* **to** [*day-of-the-week*] *hh:mm:ss*	
Remove time ranges	**no time-range [all | name** *name***]**	All means to remove all time ranges

Note:

If a time range only defines a periodic time range, only when the system clock in the

periodic time range can the time range enter the active state. If multiple periodic time ranges are defined in a time range, the relationship between the periodic time ranges is “OR”.

If a time range only defines an absolute time range, only when the system clock in the absolute time range can the time range enter the active state. If multiple absolute time ranges are defined in a time range, the relationship between the absolute time ranges is “OR”.

If a time range defines periodic time range and absolute time range simultaneously, only when the definition of the absolute time period and the period time period is satisfied at the same time can the time range be able to enter the active state.

For example:

A time range defines an absolute time range: From 0:00 on January 1, 2009 to 23:59 on December 31, 2009. At the same time, the time range also defines a periodic time range: Every Wednesday from 12:00 to 14:00. In this case, the time range can be able to enter the active state only when the time in every Wednesday from 12:00 to 14:00 in 2009.

When configuring an absolute time range, if you do not configure the start date, the time range ranges from the earliest time supported by the system to the configured end date. If the end date is not configured, the time range is from the configured start date to 2100/12/31 23:59.

13.5.1 Configuration Examples

Configure the absolute time range, which ranges from 16:00 on March 30, 2015 to 16:00 on March 31, 2015.

```
Device(config)#time-range b
Config time range successfully.
Device(config-timerange-b)#absolute start 16:00:00 2015/03/30 end 16:00:00 2015/03/31
Config absolute range successfully .

Device(config-timerange-b)#show    time-range name b
Current time is: 10:19:16      2015/03/30    Monday
time-range: b ( Inactive )
absolute:    start    16:00:00    2015/03/30    end    16:00:00    2015/03/31
```

Configure the periodic time range, which ranges from 8:00 to 18:00 and Monday to Friday

```
Device(config)#time-range d
Config time range successfully.

Device(config-timerange-d)#periodic weekdays 8:00:00 to 18:00:00
Config periodic range successfully .

Device(config-timerange-d)#show time-range name d
Current time is: 10:23:33      2015/03/30    Monday

time-range:d ( Inactive )
periodic:    weekdays 08:00:00    to                18:00:00
```

2.7 Remove the ACL rule

13.6 Activate ACL

ACLs need to be activated before they take effect, and follow the rules of “first activation takes precedence”.

Activate ACL

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
The rule for ACL activating	**access-group** [**ip-group** *name* | *num*] [**link-group** *name* | *num*] [**subitem** *num*]	required
Deactivate the specified ACLs	**no access-group** [**ip-group** *name* | *num*][**link-group** *name* | *num*] [**subitem** *num*]	
Deactivate all ACLs	**no access-group all**	

13.6.1 Configuration Examples for ACL activating

Example 1: Configure one ACL and then activate it.

Case 1:

```
//Before ACL activating
Device(config)#access-list 1 deny any
Config ACL subitem successfully.

Device(config)#access-list 1 permit 1.1.1.1 0
Config ACL subitem successfully.

Device(config)#show access-list config 1
Standard IP Access List 1, match-order is config, 2 rule:
  0   deny      any
  1   permit    1.1.1.1    0.0.0.0


// Configure to activate ACL
Device(config)#access-group ip-group 1 subitem 1
Activate ACL successfully .

Device(config)#access-group ip-group 1 subitem 0
Activate ACL successfully .
```

According to the principle that the first activation takes precedence, the device only allows packets with source IP address 1.1.1.1 to pass through.

Case 2:

```
Before ACL activating
Device(config)#access-list 1 match-order auto
Config ACL match order successfully.

Device(config)#access-list 1 deny any
Config ACL subitem successfully.

Device(config)#access-list 1 permit 1.1.1.1 0
```

Config ACL subitem successfully.

Device(config)#show access-list config 1
Standard IP Access List 1, match-order is auto, 2 rule:
  0 permit 1.1.1.1 0.0.0.0
  1 deny any

// Configure to activate ACL
Device(config)#access-group ip-group 1 subitem 0
Activate ACL successfully .

Device(config)#access-group ip-group 1 subitem 1
Activate ACL successfully .

According to the principle that the first activation takes precedence, the device only allows packets with source IP address 1.1.1.1 to pass through.

Example 2: Configure multiple ACLs and then activate them to achieve IP + MAC + port binding.

Device(config)#access-list 1 permit 1.1.1.1 0
Config ACL subitem successfully.

Device(config)#access-list 200 permit ingress 00:00:00:00:00:01 0 interface ethernet 0/1 egress any
Config ACL subitem successfully.

Device(config)#access-group ip-group 1 link-group 200
Activate ACL successfully .

13.7 ACL Display and Debug

After you completed the above configurations, you can use the following commands to view the configurations.

ACL Display and Debug

Operation	Command	Remarks
count up the number of the ACLs	**show access-list config statistic**	All modes are executable
Display the ACLs	**show access-list config** { **all** | *num* | **name** *name* }	
count up the number of the activated ACLs	**show access-list runtime statistic**	
Displayed the activated ACLs	**show access-list runtime** { **all** | *num* | **name** *name* }	

14. QACL Configuration

ACL (Access Control List), mainly refers to the network equipment in order to filter data packets, user needs to configure a series of matching rules to identify the need to filter the object. After a specific object is identified, the corresponding packet can be allowed or disabled according to a pre-defined policy.

QOS (Quality of Service) refers to a network to use a variety of basic technologies, to provide better service capabilities for specified network communication, which is a network security mechanism, and a technology used to solve network latency and congestion and other issues. Ethernet technology is today the most widely used network technology. At present, Ethernet becomes not only the dominant technology in various independent LANs, many LANs in the form of Ethernet have also become an integral part of the Internet. With the continuous development of Ethernet technology, Ethernet access method will become one of the main access methods of ordinary Internet users. Therefore, to achieve the end-to-end QoS solution for the whole network, it is inevitable to consider the QoS guarantee problem about the Ethernet. This requires Ethernet switch equipment to apply Ethernet QoS technology to provide different levels of QoS guarantee for different types of business process, and especially support business process that requiring higher for delay and dithering.

QACL (QOS & ACL), refers to the function of associating traffic rules with traffic operations by using ACL. That is, QoS functions are performed by referencing access control list, including packet filtering, commit access rate, traffic mirroring, traffic statistic, redirection, VLAN rewrite or insert, Precedence re-marking and traffic copying to the CPU, two rate three color and other functions.

14.1 QACL Related Concepts

1. Traffic

Traffic refers to message passing through the switch.

2. Traffic classification

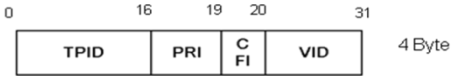
Traffic classification refers to use certain rules to identify messages that meet certain characteristics. The classification rule refers to the filtering rule configured by the administrator according to the management requirement, which may be simple, for example, the traffic of different Precedence can be identified according to the ToS field in the IP message header, or it can be complicated. Such as Layer 2, Layer 3, and Layer 4 information including MAC address, IP address, source IP address, destination IP address, port number of the application and other related information to classify messages . That is complex traffic classification rules. The general classification is limited to the header information of the package message, and the content of the message is the standard of the classification, which is relatively rare.

3. Precedence

Figure 2-1 802.1p Precedence

802.1p Precedence is in the Layer 2 message header, and applicable to occasions where Layer 3 messages don't need to be analyzed and QoS needs to be guaranteed in the Layer 2.

As described in the VLAN Configuration section, each host that supports the 802.1Q protocol, adds a 4-byte 802.1Q tag header behind the source address in the original Ethernet frame header when sending a data packet. As shown in Figure 1-1.



- 802.1Q tag

In the above figure, the PRI field is the 802.1p Precedence, which consists of 3 bits. The value ranges from 0 to 7. These three bits indicate the Precedence of the frame, which consists eight priorities, mainly is used to determine Precedence to send data packets when the switch is blocked.

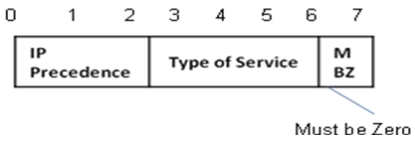
802.1Q values description

cos（Decimal）	cos（Binary）	implication
0	000	spare
1	001	background
2	010	best-effort
3	011	excellent-effort
4	100	controlled-load
5	101	video
6	110	voice
7	111	network-management

IP Precedence, ToS, and DSCP

In the IP packet header of an IPv4 message, the service type (TOS) field has 8 bits.

The Type of Service (TOS) field includes a 3-bit IP Precedence field, a 4-bit TOS field, and a 1-bit unused bit, which must be zero. Four bits TOS respectively represent: minimum latency, maximum throughput, maximum reliability and minimal cost. Among four bits at most one bit can be set at the same time. If 4 bits are 0, it means the general service.



- IP precedence and TOS precedence

There are eight priorities for IP precedence

Description of IP Precedence values

IP Precedence（Decimal）	IP Precedence（Binary）	implication
0	000	routine
1	001	priority
2	010	immediate
3	011	flash
4	100	flash-override
5	101	critical
6	110	internet
7	111	network

TOS precedence has 5 priority levels.

Description of the TOS values

TOS（Decimal）	TOS（Binary）	implication
0	0000	normal
1	0001	min-monetary-cost
2	0010	max-reliability
4	0100	max-throughput
8	1000	min-delay

Soon after, RFC2474 redefines the TOS domain of the IP message header, called the DS domain, where the DSCP precedence is represented by the first 6 bits (0-5 bits) of the domain, and the range is 0 to 63. The first 3 bits of the DSCP are used as class selectors, the 4 to 5 bits indicate the drop precedence, and the 6th bit is set to 0 to indicate that the device is a service class set as the DS model. The last two bits are reserved bits.

- DSCP precedence

The Diff-Serv network defines four types of traffic:

Expedited Forwarding (EF) class, which is applicable to low-delay, low-loss, low-jitter, and bandwidth-priority services (such as virtual leased lines), regardless of whether other traffic share its link.

Assured Forwarding (AF) class is divided into four sub-categories (AF1/ 2/3/4). Each AF class is divided into three drop precedence, which can be used to classify the AF business. AF Classes have lower QoS level than EF classes.

Class selector (CS) evolves from the IP TOS field, a total of eight categories.

Best Effort (BE) is a special category of CS, there is no guarantee. The AF class can be downgraded to BE class after overrun, the existing IP network traffic is also defaulted to this category.

DSCP value description

DSCP（Decimal）	DSCP（Binary）	KEY Word
0	000000	be
46	101110	ef
10	001010	af1
18	010010	af2
26	011010	af3
34	100010	af4
8	001000	cs1
16	010000	cs2
24	011000	cs3
32	100000	cs4
40	101000	cs5
48	110000	cs6
56	111000	cs7

Note: Before configuring these ACL tasks, configure permit ACL according to your requirements. For more information about ACL configuration, see the ACL Configuration Guide.

14.2 Configure Traffic Speed Limit

Traffic-based speed limit can monitor the rate of a traffic entering the switch. If the traffic exceeds the configured threshold, corresponding measures are taken, such as dropping those messages that exceed the threshold or resetting their priorities.

Configure traffic speed limit

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure traffic speed limit	**rate-limit** { **input** | **output** } { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] } *target-rate*	Optional, some devices only support inbound direction，some devices support inbound and outbound direction.

14.3 Configure Two Rate Three Color Marker

Two rate three color marking function is defined in RFC 2698, and mainly based on four kinds of flow parameters to evaluate: CIR、CBS、PIR、PBS.

Configuration of two rate three color

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure color-aware mode	**two-rate-policer mode {color-aware | color-blind}**	optional
Different DSCP messages are color-coded	**two-rate-policer set-pre-color** {*dscp-value* **{green | red | yellow}**}	optinal
Configure a two rate three color application policy (processing action for three different color messages)	**rate-limit input**{ [**ip-group** { *num* | *name* } [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] } **two-rate-policer cir** *cir-value* **cbs** *cbs-value* **pir** *pir-value* **pbs** *pbs-value* **conform-action {copy-to-cpu | drop | set_dscp_value** *dscp_value* **|transmit } exceed-action {copy-to-cpu | drop | set_dscp_value** *dscp_value* **|transmit } violate-action{copy-to-cpu | drop | set_dscp_value** *dscp_value* **|transmit }**	optional

Note: The color-aware mode corresponds to the color-blind mode; the system default is the color-blind mode. The difference between the two is as follows:

Figure 56: TrTCM in Color-Blind Mode

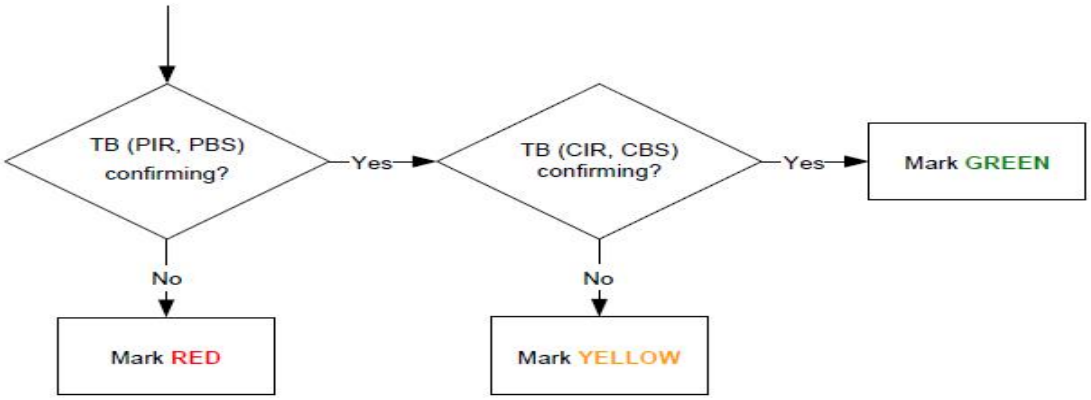
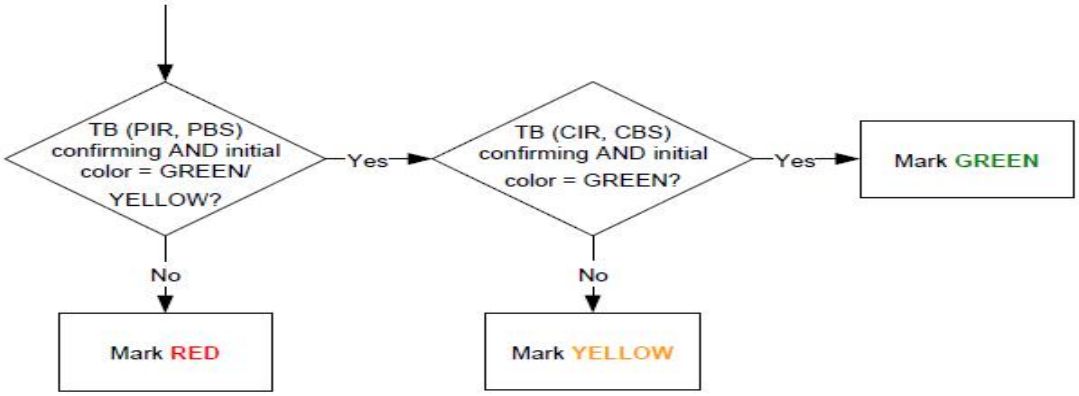


Figure 57: TrTCM in Color-Aware Mode



14.4 Configure Message Redirection

The message redirection is that the message forwarded is redirected to an outgoing port.

Configure message redirection

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure message redirection	**traffic-redirect** { [**ip-group** { *num* | *name* } [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] } { [**interface***interface-num* | **cpu**] }	optional

14.5 Configure Message Copy to CPU

After the configuration message is copied to the CPU function, the switch automatically copies a specified message to the CPU.

Configure messages copy to CPU

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
configuration messages copy to CPU	**traffic-copy-to-cpu** { [**ip-group** { *num* | *name* } [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] }	optional

14.6 Configure Precedence Marker

The precedence marker function is a strategy of remarking the priority to match ACL message. The precedence marker function can remark IP precedence, ToS, DSCP and 802.1p precedence for message. User can also specify the local precedence for these messages matching ACL.

Configure precedence marker

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure precedence marker	**traffic-priority** { [**ip-group** { *num* | *name*} [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] } { [**dscp***dscp-value*] [**cos** { *pre-value* | **from-ipprec** }] [**local-precedence***pre-value*] }	optional

NOTE: If both the 802.1p precedence and the local precedence are specified, the switch will use 802.1p precedence to put the message into the corresponding port output queue.

14.7 Configure Traffic Statistic

The traffic statistic function can be used to collect matching ACL rule messages. The statistic is an accumulated value that can be cleared by a command. If user re-configures the traffic statistic, the traffic statistic will be cleared.

Configure traffic statistic

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure traffic statistic	**traffic-statistic** { [**ip-group** { *num* | *name* } [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] }	optional
Clear statistic information	**clear traffic-statistic** { [**all** | [**ip-group** { *num* | *name* } [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]]] }	optional

14.8 Configure VLAN Rewrite

The VLAN ID of the business process matching ACL rule is rewritten the configured VLAN ID.

Configure message VLAN rewrite

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure message VLAN rewrite	**traffic-rewrite-vlan** { [**ip-group** { *num* | *name*} [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] } *rewrite-VLAN -ID*	optional

14.9 Configure VLAN Insert

Message matching ACL rule is inserted into an outer VLAN. The VLAN ID is the inserted VID. The VLAN precedence is the port precedence.

Configure message insert VLAN function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure message VLAN insert	**traffic-insert-vlan** { [**ip-group** { *num* | *name*} [**subitem***subitem*]] [**link-group** { *num* | *name* } [**subitem***subitem*]] } *insert-VLAN -ID*	optional

14.10 QACL Display and Maintenance

QACL display and maintenance

Operation	Command	Remarks
Display all QoS parameter setting	**show qos-info all**	-
Display the statistics of all QoS	**show qos-info statistic**	
Display the parameter setting for copying the message to the CPU	**show qos-info traffic-copy-to-cpu**	
Display the parameter setting of traffic mirroring	**show qos-info mirrored-to**	
Display the parameter setting of precedence marker	**show qos-info traffic-priority**	
Display the parameter setting of redirect	**show qos-info traffic-redirect**	
Display the traffic statistic	**show qos-info traffic-statistic**	
Display the parameter setting of VLAN insert	**show qos-info traffic-insert-vlan**	
Display the parameter setting of two rate three color	**show two-rate-policer**	
Display all the configuration information of rate-limit	show qos-interface all	
Display the rate-limit configuration information of all ports	show qos-interface rate-limit	
Statistics rate-limit Number of rules	show qos-interface statistic	

show qos-infodisplays the configuration information related to the traffic configuration
show qos-interfacedisplays the configuration information related to rate-limit

14.11 QACL Configuration example

一、Requirements and networking diagram

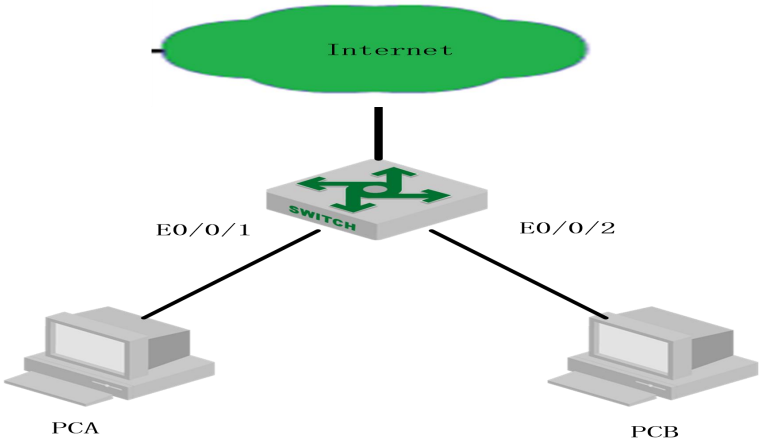
The interconnection between A and B is achieved by the Ethernet switch. A and B do not belong to the same network segment. A connects to the switch through Eth0/0/1, and B connects to the switch through Eth0/0/2.

PCB equips with data detection equipment. The specific needs are as follows:

Statistic non-workday traffic on the Internet through HTTP on port E0/0/1;

Redirect traffic through the port E0/0/1 by HTTP to access internet to E0/0/2

The networking diagram is as follows:



二, configuration steps

1、 Configure the time period

Switch(config)#time-range a

Switch(config-timerange-a)#periodic weekdays daily 08:30:00 to 18:00:00

Switch(config)#time-range b

Switch(config-timerange-b)#periodic weekdays 00:00:00 to 08:30:00

Switch(config-timerange-b)#periodic weekend 00:00:00 to 23:59:00

2、 Configure ACL, according to the different time period to access the Internet by HTTP message classification

Switch(config)#access-list 100 permit tcp any 192.168.0.1 0 80 time-range a

Switch(config)#access-list 100 permit tcp any 192.168.0.1 0 80 time-range b

15.Cos Control

15.1 Overview for Cos Control Function

When the network is congested, it is necessary to solve the problem that multiple messages are competing for resources at the same time, which is usually solved by queue scheduling. Common queue scheduling algorithms include FIFS, Strict-Priority Queue scheduling, Weighted Round Robin (WRR) scheduling, and SP + WRR scheduling.

FCFS (First Come First Serve): First in first out is usually called FIFO. The first in first out queue does not classify message. When the message arrives at an interface faster than the interface can send, the FCFS forwards the message to the queue according to the order that the message arrives at the interface. At the same time, the FCFS sends the message at the queue exit according to the order into the team out of the team. First in message will be the first out from the team. Later in message will come out after the team.

SP (Strict-Priority Queuing), is designed for critical business applications. One important feature of critical business is the requirement to prioritize services in order to reduce the latency of the response when congestion occurs. The priority queue classifies all messages into eight classes (follow by 7, 6, 5, 4, 3, 2, 1, 0 queue), and their priorities are reduced in turn.

In the queue scheduling, SP strictly sends a higher priority queue in accordance with the order of priority from high to low. When the higher priority queue is empty, then the group of the lower priority queue in the queue is sent. In this way, the group of critical services is put into the queue of higher priority, the non-critical business is put into lower priority queue, which can ensure that the key business of the message is prioritized. The group of non-critical business is transmitted in idle gap of handling critical business data.

The disadvantage of the SP is that if groups exist for a long time in a higher priority queue, the messages in the low priority queue are "starved", because they cannot get services.

WRR queue scheduling divides each port into 8 output queues (followed by 7, 6, 5, 4, 3, 2, 1, 0 queue, and their priorities are reduced in turn). The queues are scheduled in turn, and ensure that each queue has a certain service time. WRR can be configured with a weighted value (w7, w6, w5, w4, w3, w2, w1, w0) for each queue. The weighted value represents the weight of the resource. Such as a 100M port, configure the weighted value of WRR queue scheduling algorithm for 80, 70, 60, 50, 50, 40, 30, 20(correspond w7, w6, w5, w4, w3, w2, w1, w0 in turn). This ensures that the lowest priority queue can obtain at least 5 Mbit / s bandwidth. This avoids the disadvantage that message in low priority queues may not receive services for a long time when SP scheduling is adopted.

An advantage of WRR queuing is that although multiple queues are scheduled by polling, each queue is not assigned a fixed time slot - if a queue is empty, it immediately switches to the next queue schedule, So the bandwidth resources can be fully utilized.

In the Sp + WRR queue scheduling, if the weight value of a queue is set to 0, the queue will perform the strict priority algorithm. Otherwise, it is the weight value of the WRR queue.

WFQ is the same as the WRR principle. The difference is that WRR uses pps to calculate the weight of the queue, and WFQ uses bps to calculate the weight of the queue. Such as a 100M port, configure the weighted value of WFQ queue scheduling algorithm for 80, 70, 60, 50, 50, 40, 30, 20(correspond w7, w6, w5, w4, w3, w2, w1, w0 in turn), So the lowest priority queue is guaranteed to have at least a bandwidth of 20 / (80 + 70 + 60 + 50 + 50 + 40 + 30 + 20) *% 100, where bandwidth is calculated using bits / Bytes, For example, in the above weights, the export message statistics bits / bytes ratio of each queue approximate 8: 7: 6: 5: 5: 4: 3: 2, that is, the export bandwidth of the w0 queue is 5 M.

Sp + WFQ is also calculated in accordance with bps queue weight.

15.2 CoS Control Configuration

15.2.1 Configure CoS Control

Queue scheduling doesn't have function switch (enable/disable), which is enabled by default, and uses SP (strict priority) scheduling.

In SP + WRR, a queue with a weight of 0 uses SP, and the other queues are forwarded by WRR weight.

Configure CoS

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	-
Use the SP scheduling mode	queue-scheduler strict-priority	optional
Use the WRR scheduling mode	queue-scheduler wrr *w1 w2 w3 w4 w5 w6 w7 w8*	optional
Use the SP+WRR scheduling mode	queue-scheduler sp-wrr *w1 w2 w3 w4 w5 w6 w7 w8*	optional
Restore the default schedule	no queue-scheduler	optional
Information View	show queue-scheduler	optional

Some devices support 4 queues, some support 8 queues;

The default scheduling method, some devices are SP, some devices are FIFO.

15.2.2 Configure 802.1p and Hardware Queue Mapping

The system maps the 802.1p priority and hardware queue priority of the message. For each message entering the switch, the system maps the specific hardware queue priority according to the 802.1p priority of the message. By default, the mapping relation between 802.1p and hardware priority is as follows:

802.1p	Hardware priority queue
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7

By changing the mapping relation between 802.1p priority and hardware queues, we can change the mapping relation between 802.1p priorities and output queues.

As the chip queue scheduling use a random algorithm, if the two 802.1p priorities are

mapped to the same hardware priority queue, messages of two 802.1p priorities cannot be forwarded with 1: 1 forwarding.

Configure 802.1p and hardware queue mapping

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure 802..1p and hardware queue mapping	**queue-scheduler cos-map** queue-v priority-v	optional
Map Information View	**show queue-scheduler cos-map**	optional

15.2.3 Configure DSCP and 802.1P Mapping

By default, the mapping relation between 802.1p and hardware priority is as follows:

DSCP	Hardware priority queue	DSCP	Hardware priority queue	DSCP	Hardware priority queue	DSCP	Hardware priority queue
0	0	16	2	32	4	48	6
1	0	17	2	33	4	49	6
2	0	18	2	34	4	50	6
3	0	19	2	35	4	51	6
4	0	20	2	36	4	52	6
5	0	21	2	37	4	53	6
6	0	22	2	38	4	54	6
7	0	23	2	39	4	55	6
8	1	24	3	40	5	56	7
9	1	25	3	41	5	57	7
10	1	26	3	42	5	58	7
11	1	27	3	43	5	59	7
12	1	28	3	44	5	60	7
13	1	29	3	45	5	61	7
14	1	30	3	46	5	62	7
15	1	31	3	47	5	63	7

User can change the mapping relation between DSCP precedence and output queues by changing the mapping between DSCP priorities and 802.1p priorities according to the actual network requirements.

Dscp mapping is disabled by default. User needs to enable it before you can perform related configurations. If both 802.1p and dscp are used, the dscp-map is preferred.

Configure DSCP and 802.1p mapping

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the dscp mapping function	[no] queue-scheduler dscp-map	optional
Modify the dscp and 802.1p mapping	queue-scheduler dscp-map ***dscp-v priority-v***	optional
Dscp mapping information view	show queue-scheduler dscp-map	optional

15.3 COS Control Configuration Example

1. Configuration step

View the default queue scheduling mode

SW1(config)#show queue-scheduler

Queue scheduler status : enable

Queue scheduler mode : SP (Strict Priority)

View the priority mapping relationship between 802.1p and hardware queues
SW1(config)#show queue-scheduler cos-map
Information about map of cos:
802.1P Priority Queue of class
---
0 0
1 1
2 2
3 3
4 4
5 5
6 6
7 7
Modify the priority mapping relationship between 802.1p and hardware queues: map packets with priority = 0 to queue 1, only for demonstration, the actual use is the default value;
SW1(config)#queue-scheduler cos-map 1 0
Config successfully.
SW1(config)#show queue-scheduler cos-map
Information about map of cos:
802.1P Priority Queue of class
---
0 1
1 1
2 2
3 3
4 4
5 5
6 6
7 7

#Use wrr queue scheduling
SW1(config)#queue-scheduler wrr 1 2 3 4 5 6 7 8
Config queue scheduler successfully.

SW1(config)#show queue-scheduler
Queue scheduler status : enable
Queue scheduler mode : WRR (Weighted Round Robin)
Queue0 weight is 1
Queue1 weight is 2
Queue2 weight is 3
Queue3 weight is 4
Queue4 weight is 5

Queue5 weight is 6
Queue6 weight is 7
Queue7 weight is 8

Restore the default queue schedule
SW1(config)#no queue-scheduler
Recover queue scheduler to default value(strict-priority) successfully.

SW1(config)#show queue-scheduler
Queue scheduler status : enable
Queue scheduler mode : SP (Strict Priority)

16.Forward Control

16.1 bandwidth-control

Bandwidth-control is mainly to achieve the bandwidth of the export or import restriction; and limit the total rate of incoming or outgoing packets from the port.

16.1.1 Configure Bandwidth Limit for Port.

In the port, configure the bandwidth limit for the inbound / outbound direction of the port.

Configure bandwidth limit for port

operation	command	remark
Enter port mode	interface ethernet ***port-number***	-
Configure the bandwidth limit for port outbound	[no]bandwidth egress ***rate***	optional
Configure the bandwidth limit for port inbound	[no]bandwidth ingress ***rate***	optional
Configure the bandwidth limit of the port based on the queue	bandwidth queue ***queue-id*** { maximum | minimum} ***rate***	optional
Cancel the bandwidth limit of the port based on the queue	no bandwidth queue ***queue-id*** { maximum | minimum }	optional

16.1.2 Bandwidth Limit Display and Maintenance

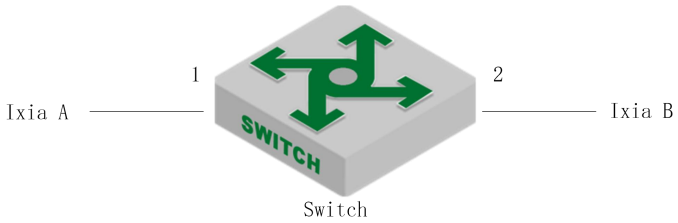
Bandwidth Limit Display and Maintenance

operation	command	remark
View the bandwidth limit information of the port	show bandwidth-control interface [ethernet ***port-number***]	optional
View the queue-based bandwidth limit of port	show bandwidth queue interface [ethernet ***port-number***]	optional

16.1.3 bandwidth-control configuration example

I. Network requirement

Set the ingress speed of port 1 to 1024 (1M).



Bandwidth control Schematic Diagram

2. Configuration steps

Configure bandwidth control

Switch(config)#interface ethernet 0/0/1

Switch(config-if-ethernet-0/0/1)#bandwidth ingress 1024

Switch(config-if-ethernet-0/0/1)#exit

View the configuration information

Switch(config)#show bandwidth-control interface ethernet 0/0/1

port Ingress bandwidth control Egress bandwidth control

e0/0/1 1024 kbps disable

Total entries: 1.

3. Validation results

16.2 Storm-control Function

When the network appears loop or malicious attacker, there will be a lot of messages, these messages wasted bandwidth and even make the network equipment in the edge of the collapse. The Storm-control function is used to avoid excessive messages in the network. It monitors the number of message on each port of the switch to understand the bandwidth usage, and controls the bandwidth of the message within the configured threshold. Packets exceeding the threshold will be discarded.

16.2.1 Storm-control Configuration

The Storm-Control function is implemented in the port configuration mode. That is, the administrator can set different storm suppression policies for different ports.

Configure Storm-control

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *device/slot/port*	-
Configure storm suppression message type and suppression threshold	**Storm-control {broadcast |multicast |unicast}** *target-rate*	optional

16.2.2 Storm-control Display and Maintenance

After you complete the above configuration, you can use the following command to view the configuration.

Storm-control display and maintenance

operation	command	remark
Display the storm suppression message type and suppression threshold for the port	**show interface ethernet** *device/slot/port*	All modes are executable

16.2.3 Storm-control Configuration Example

Enable storm suppression on port 1 and set the broadcast storm suppression value for 128 pps, the unknown multicast storm suppression value is 256 pps, and the unknown unicast storm suppression value is 512 pps.

Switch(config)# interface ethernet 0/0/1

Switch(config-if-ethernet-0/0/1)#storm-control broadcast 128

Switch(config-if-ethernet-0/0/1)#storm-control multicast 256

Switch(config-if-ethernet-0/0/1)#storm-control unicast 512

View the port storm control settings status

Switch(config-if-ethernet-0/0/1)#show interface ethernet 0/0/1

Ethernet e0/0/5 is enabled, port link is down

Hardware address is 00:00:53:28:00:0a

SetSpeed is auto, ActualSpeed is unknown, Duplex mode is unknown

Current port type: 1000BASE-T

Priority is 0

Flow control is disabled

Broadcast storm control target rate is 128Kbps

Multicast storm control target rate is 256Kbps

Unicast storm control target rate is 512Kbps

PVID is 1

Port mode: hybrid

Untagged VLAN ID : 1

Input : 0 packets, 0 bytes

0 broadcasts, 0 multicasts, 0 unicasts

Output : 0 packets, 0 bytes

0 broadcasts, 0 multicasts, 0 unicasts

17.Attack Protection

17.1 Anti-DDOS Attack Function

Dos is short of Denial of Service. DoS attack caused by the attack is known as DdoS. Its purpose is that let computer or network not provide normal services.

Dos attack is a simple and effective attack method which is very harmful to many network attack technologies. It is through various means to consume network bandwidth and system resources, or attack system defects, so that the normal service of normal system is paralyzed state, and cannot service normal user. It achieves to deny normal user accessing services, so in the internet anti-DOS attack is more important. Configure the anti-TTL attack.

According to the relevant standard, the TTL field in the IP header must be greater than 0. By default, if the message of TTL = 0 is received, the switch discards the message as an attack, but allows the message of ttl = 0 to be discarded.

17.1.1 Anti-TTL Attack

Configure anti-TTL attack

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enable anti-TTL attack	anti-dos ip ttl	Optional. By default, messages with ttl = 0 are discarded
Disable the anti-TTL attack	no anti-dos ip ttl	Optional, After configuration, normal messages are processed
View the configuration information	show anti-dos	optional

17.1.2 Configure Anti-IP Fragment Attack

If the number of an IP message fragment is many, the switch will take up too many system resources and may affect other messages. Therefore, a reasonable limit for the length of the IP message does not allow too many fragments. If the number of fragment exceeds the specified value, the message is discarded as an attack message. By default, an IP message has 800 fragments.

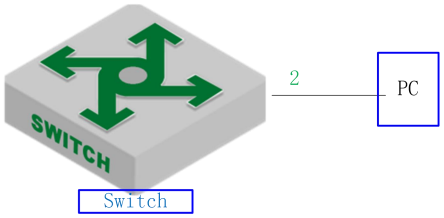
Configure Anti-IP Fragment Attack

operation	command	remark
Enter the global configuration mode	configure terminal	-
Set the maximum number of IP messages allowed	[no]anti-dos ip fragment ***max-numbers***	Optional, no command restores the default value of 800

17.1.3 Configuration Example

1、Network Requirement

The PC directly connects to the switch and communicates. Verify how the DUT handles more than the permitted fragment and the normal fragment, respectively. The switch: ip=10.5.2.134； PC IP=10.5.2.91



Anti-dos attack

2、Configuration steps

Configure an IP message to have up to two fragments

DUT needs two fragments of the IP message, you can communicate properly

Switch(config)#ping -l 2800 10.5.2.91

PING 10.5.2.91: with 2800 bytes of data:

reply from 10.5.2.91: bytes=2800 time<10ms TTL=64

reply from 10.5.2.91: bytes=2800 time<10ms TTL=64

reply from 10.5.2.91: bytes=2800 time<10ms TTL=64

reply from 10.5.2.91: bytes=2800 time<10ms TTL=64

reply from 10.5.2.91: bytes=2800 time<10ms TTL=64

----10.5.2.91 PING Statistics----

5 packets transmitted, 5 packets received, 0% packet loss

round-trip (ms) min/avg/max = 0/0/0

DUT needs three fragments of the ip message, you cannot communicate

Switch(config)#ping -l 3000 10.5.2.91

PING 10.5.2.91: with 3000 bytes of data:

Request timed out.

Request timed out.

Request timed out.

Request timed out.

Request timed out.

no answer from 10.5.2.91

Delete ip fragmentation configuration (restore the default value of 800), and then need to send three pieces of ip messages, communication is normal

Switch(config)#no anti-dos ip fragment

Switch(config)#ping -l 3000 10.5.2.91

PING 10.5.2.91: with 3000 bytes of data:

reply from 10.5.2.91: bytes=3000 time=10ms TTL=64

reply from 10.5.2.91: bytes=3000 time<10ms TTL=64

reply from 10.5.2.91: bytes=3000 time=10ms TTL=64

reply from 10.5.2.91: bytes=3000 time<10ms TTL=64
reply from 10.5.2.91: bytes=3000 time<10ms TTL=64

----10.5.2.91 PING Statistics----
5 packets transmitted, 5 packets received, 0% packet loss
round-trip (ms) min/avg/max = 0/4/10

17.2 CPU-car Function

A large number of messages on the CPU will cause the CPU busy. This function is used to limit the rate of receiving messages by the CPU.

17.2.1 Configure Cpu-car

Cpu-car is enabled by default and does not support the shutdown function

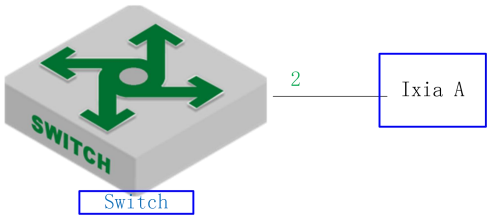
Configure cpu-car

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the cpu-car rate	[no]cpu-car ***value***	Optional, The no command restores the default value 400pps
View the configuration information	show cpu-car	optional
View cpu receiving packet port statistics	show cpu-statistics [ethernet ***port-number***]	optional
Clear cpu receiving packet port statistics	clear cpu-statistics	optional
View cpu receiving packet classification statistics	show cpu-classification [interface ethernet ***port-number***]	optional
Clear cpu receiving packet classification statistics	clear cpu-classification [interface ethernet ***port-number***]	optional
View cpu utilization	show cpu-utilization	optional

17.2.2 Configuration Example

1、Network Requirement

Limit the rate of message less than 50 pps on the switch.



cpu-car schematic diagram

2. Configuration steps

Configure the cpu-car speed to 50 pps

Switch(config-if-ethernet-0/0/2)#port-car-rate 50

View Configuration Information

Switch(config)#show cpu-car

Send packet to cpu rate = 50 pps.

3. Validation results

Ixia A sends icmp request messages to the DUT: at a rate of 100 pps for 10 seconds, the total number of messages on the dut is 600, indicating that the cpu-car function takes effect.

Switch(config)#clear cpu-statistics

Switch(config)#clear cpu-classification

Switch(config)#clear interface

Switch(config)#show cpu-statistics ethernet 0/0/2

Show packets sent to cpu statistic information

port	64Byte	128Byte	256Byte	512Byte	1024Byte	2048Byte
e0/0/2	600	0	0	0	0	0

Switch(config)#show cpu-classification

Type	Count	Percent(%)
Total	600	100
BPDU	0	0
ERRP	0	0
ARP	0	0
MLD	0	0
IGMP	0	0
ICMP	600	100
OSPF	0	0
RIP	0	0
DHCP	0	0
SNMP	0	0
Telnet	0	0
PIM	0	0

BGP 0 0
SSH 0 0
Other 0 0
Switch(config)#show statistics interface ethernet 0/0/2

Port number : e0/0/2
last 5 minutes input rate 5248 bits/sec, 10 packets/sec
last 5 minutes output rate 433832 bits/sec, 771 packets/sec
64 byte packets:1048
65-127 byte packets:0
128-255 byte packets:0
256-511 byte packets:0
512-1023 byte packets:0
1024-1518 byte packets:0
1048 packets input, 67072 bytes , 0 discarded packets
1048 unicasts, 0 multicasts, 0 broadcasts
0 input errors, 0 FCS error, 0 symbol error, 0 false carrier
0 runts, 0 giants
19 packets output, 1216 bytes, 0 discarded packets
0 unicasts, 9 multicasts, 10 broadcasts
0 output errors, 0 deferred, 0 collisions
0 late collisions
Total entries: 1.

17.3 Shutdown-Control Overview

When the network appears loop or malicious attack, there will be a lot of messages, these messages waste bandwidth or even make the network equipment in the collapse of the edge, and affect the normal use of other users. The shutdown-control function is used to avoid excessive messages in the network. It monitors the bandwidth of each port on the switch. When the number of unknown messages received by the port exceeds the security set by the administrator threshold, the

shutdown-control function automatically shuts down the port to ensure that the other links and devices are protected from the impact in the network.

17.3.1 Enable/disable Shutdown-Control

Enable/disable Shutdown-Control

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enter the port configuration mode	interface ethernet *port-num*	-
Enable and configure the shutdown rate	shutdown-control **{broadcast |multicast |unicast}*rate***	required
Shutdown function	no shutdown-control **{broadcast |multicast |unicast}**	optional，default close
View the configuration information	show shutdown-control interface [ethernet ***port-numer***]	optional

17.3.2 Configure Recovery Mode

Port is shutdown, need to manually restore by default. Administrators can configure automatic recovery, and set the recovery cycle, the default is 480s.

Configure recovery mode

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the recovery mode	[no] shutdown-control-recover mode {**automatic** | **manual**}	Optional, manual and no commands are used to restore the default configuration
Configure the automatic recovery period	[no] shutdown-control-recover automatic-open-time ***value***	optional，Default 480s, only valid for automatic recovery
View configuration information	show shutdown-control interface [ethernet ***port-number***]	optional

17.3.3 Manually Restore Shutdown Port

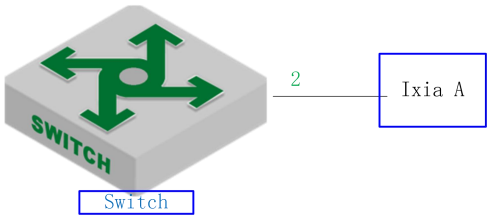
Manually restore shutdown port

operation	command	remark
Enter the shutdown port configuration mode	interface ethernet ***port-number***	required
Command the shutdown port	shutdown	required
Restore the port	no shutdown	required

17.3.4 Configuration Example

1、Network Requirement

Port 2 receiving unknown unicast rate is limited for 1000pps, if it is shutdown, automatic recovery, the default value 480s is used for recovery cycle.



Shutdown-control sketch map

2.configuration steps

Enable the unknow unicast shutdown-control function and set the rate to 1000 pps

Switch(config)#interface ethernet 0/0/2

Switch(config-if-ethernet-0/0/2)#shutdown-control unicast 1000

Switch(config-if-ethernet-0/0/2)#ex

View Configuration Information

Switch(config)#show shutdown-control interface ethernet 0/0/2

port shutdown control recover mode : automatic

Port recover time(second) : 480

port shutdown control information :

PortID	Broadcast status	Broadcast value	Multicast status	Multicast value	Unicast status	Unicast value	RemainTime
e0/0/2	disable	-	disable	-	enable	1000	-

Total entries: 1 .

3. Validation results

Switch(config)#logging monitor 0

The tester sends an unknown message to the DUT 0/0/2 at a rate of 1100 pps.

Switch(config)#05:12:04: Switch: %DEVICE-3-LINKUPDOWN: e0/0/2 LinkDown.

05:12:04: Switch: %OAM-5-SHUTDOWN-CTRL: port e0/0/2 **was shutdown.**

Switch(config)#show shutdown-control interface ethernet 0/0/2

port shutdown control recover mode : automatic

Port recover time(second) : 480

port shutdown control information :

PortID	Broadcast status	Broadcast value	Multicast status	Multicast value	Unicast status	Unicast value	**RemainTime**
e0/0/2	disable	-	disable	-	enable	1000	**07min48sec**

Total entries: 1 .

Switch(config)#show interface brief ethernet 0/0/2

Port	Desc	Link	shutdn	Speed	Pri	PVID	Mode	TagVlan	UtVlan
e0/0/2		down	**ERROR**	auto	0	1	hyb		1

Total entries: 1 .

Switch(config)#05:20:06: Switch: %DEVICE-3-LINKUPDOWN: e0/0/2 LinkUp.
05:20:08: Switch: %OAM-5-PORTRECOVER: port e0/0/2 **recover.**

17.4 Anti-DHCP Attack

Normally, when the dhcp client obtains ip from the dhcp server, the rate of dhcp message sent by the dhcp client is very small. Generally, it doesn't cause the dhcp server disabled. However, a malicious attacker may send large rate dhcp message to the dhcp server, which will cause the dhcp server busy, affect the allocation of ip for other clients, and even cause panic.

The anti-dhcp attack function restricts the dhcp message rate of the dhcp client. Over-rate client will be considered as malicious attackers, according to a good strategy to deal, so as to protect the dhcp server to work normally.

17.4.1 Enable/disable Anti-DHCP

Enable/disable Anti-DHCP

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enable/disable Anti-DHCP	[no] dhcp anti-attack	required，off by default
View configuration information	show dhcp anti-attack [interface ethernet ***port-numbwer***]	optional

17.4.2 Configure Processing Policy

After the switch detects an attack, it can take two actions: 1) Discard all the messages of the client (based on the source MAC address of messages to distinguish) 2) Discard only the dhcp message of the client (according to the source MAC address of the message to distinguish), that is, the client is not assigned ip.

When the switch detects an attack, it sends the source MAC address of the attack message to the attack entry. If the policy drops all packets, user can manually bind the attack entry to a black hole MAC address.

Configure processing policy

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure processing policy	dhcp anti-attack action [deny-all | deny-dhcp]	Optional, deny-dhcp by default
Bind black hole mac table	dhcp anti-attack bind blackhole [all | mac-address]	Optional, It can be configured only when deny-all is specified
View configuration information	show dhcp anti-attack [interface ethernet ***port-number***]	optional

17.4.3 Configure Rate Threshold

In the anti-dhcp attack, the rate of dhcp message sent by the same user is determined whether there is attack. If the rate is equal to or higher than 16 pps, the message is considered as an attack. The administrator is allowed to modify the rate threshold.

Configure rate threshold

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure rate threshold	[no] dhcp anti-attack threshold ***value***	Optional, 16pps by default

View configuration information	show dhcp anti-attack [interface ethernet ***port-number***]	optional
Enter the port configuration mode	interface ethernet ***port-number***	-
Configure rate threshold	[no]dhcp anti-attack threshold ***value***	optional

17.4.4 Configure Recovery Function

When the switch detects an attack, it sends the source MAC address of the attack message to the attack table item. The attack table item maintains an aging time. When the aging time expires, the table item is deleted. The default aging time is 10 minutes. If you do not want to delete a table item, you can configure 0.

Configure recovery function

operation	command	remark
Enter the global configuration mode	configure terminal	-
View Configuration Information	show dhcp anti-attack [interface ethernet ***port-number***]	Optional, and display attack table item
Configure recovery time	dhcp anti-attack recover-time ***value***	optional，10m by default，0 means no aging
Configure manual recovery	dhcp anti-attack recover [all | ***mac-address***]	The table items are restored immediately, and do not need to wait for the aging time to expire

17.4.5 Configure Trusted Ports

By default, all ports are considered to be untrustworthy after the global anti-dhcp attack is enabled, and you need to monitor whether the dhcp attack exists or not. If the port does not appear dhcp attack, it can be modified into a trust port, so you do not need to monitor whether there dhcp attack or not.

Configure trusted port

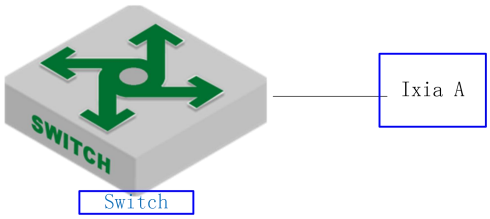
operation	command	remark
Enter the port configuration mode	interface ethernet ***port-number***	-
Configure whether the port is a trusted port	[no]dhcp anti-attack trust	optional ， un-trusted port by default
View configuration information	show dhcp anti-attack interface ethernet ***port-number***	optional

17.4.6 Configuration Example

I. Network requirement

The switch is regarded as dhcp server, and the anti-dhcp attack function is enabled. Ixia A

emulates the PC to send dhcp message. To demonstrate the effect, configure the anti-dhcp attack rate threshold to 1pps, and enable the auto-recovery function.



Anti-DHCP attack

2、configuration example

DHCP server configuration omitted, please refer to the relevant manual

#configure anti-DHCP attack

Switch(config)#dhcp anti-attack

Switch(config)#dhcp anti-attack action deny-dhcp

Switch(config)#dhcp anti-attack threshold 1

Switch(config)#dhcp anti-attack recover-time 3

Display the log information (No need to configure the actual use)

Switch(config)#logging monitor 0

Switch(config)#debug dhcp

3. Validation results

Ixia A sends dhcp request message to dut at the rate of 2 pps. The log information is as follows:

Switch(config)#

05:26:56: Switch: %DHCP-4-DHCP: 19616:33: Deny user 00:00:00:01:11:23,dhcpRate 2pps

05:26:58: Switch: %DHCP-4-DHCP: 19618:33: Deny user 00:00:00:01:11:23,dhcpRate 2pps

Send the attack entry

Switch(config)#show dhcp anti-attack

Dhcp anti-attack: enabled

Dhcp rate limit:1pps

User recovery time:3 minutes

Reject type:DenyDHCP

DeniedSrcMAC	Port	Vlan	DenyType	RemainAgingTime(m)
00:00:00:01:11:23	e0/0/1	2	DenyDHCP	3

Total entry: 1.

#After 3 minutes, the attack entry is aged out

Switch(config)#show dhcp anti-attack

Dhcp anti-attack: enabled

Dhcp rate limit:1pps

User recovery time:3 minutes
Reject type:DenyDHCP
DeniedSrcMAC Port Vlan DenyType RemainAgingTime(m)

Total entry: 0.

17.5 ARP Spoofing and Flood Attack

17.5.1 Overview for ARP Spoofing

If two hosts need to communicate, they should know each other's MAC address. ARP protocol makes this procedure transparent to users. However, there is no certification instructions in ARP protocol, it left the door open for attacker as a consequence.

All devices in LAN can receive the ARP request of host A, so if host C is an attacker, he pretends to be host B to send ARP reply to host A "my address is 00:00:00:00:00:03" , host A will unconditionally believe in this reply and then add or cover the intrinsical APR table. However, the IP of this table is 192.168.1.4 while its corresponding MAC is 00:00:00:00:00:03. So the host C can be able to intercept and capture the message which should be sent to host B. Due to host A is treat by false ARP, this is also called the ARP spoofing attacks.

After enabling this function, all ARP which will go through Switch will be redirected to CPU for a check. The ARP packets will be checked one by one whether they are complete matched with static arp table, ip-source-guard static binding table and dhcp-snooping table. If there exists one cannot be complete matched, it will stop the follow-up inspection and this arp packet can be transmitted. If there exists one incomplete matching (partial matching) table, the arp packet will be discarded. If there is no corresponding static ARP table, static ip-source-guard table and

dhcpsnooping will be handled according to configured strategy: discard or flood (send to all ports), the function of anti-ARP spoofing attack will be disabled by default.

17.5.2 Overview for ARP Flooding Attack

Arp flood attack is generally attack the network device (for example: router, Switch, server and so on) with large number of message traffic, exhausting the CPU resource of network device and then leading to the network paralysis.

When facing to such kind of flood attack, the most important thing is to ensure the normal operation of the network device, preventing widespread network paralysis. There are various flood attacks, and the most damage to device is ARP attack. According to the above mentioned ARP mechanism, all network devices will send the ARP request packet to CPU to handle after they receives the ARP request packet. Only in this way can they judge if they are the other equipment who request its MAC address. ARP flood attack takes advantages of this ARP mechanism flaw, randomly sending a lot of ARP request packet to attack the network equipment in the local area network (LAN) .

Main purpose of ARP flood attacker is to impact the network equipment's CPU, and then run out the CPU resources of the core equipment. Switch should judge it ahead of time and forbid the transmission of flood packet so as to defense the attack of this type.

arp anti-flood function can be able to identify each ARP flow and then judge whether it's ARP flood attack according to configured safe ARP rate-value. The Switch will take it as flood attack if the ARP traffic of a certain host exceeds the configured ARP rate-value, and it will put this virus-host into blacklist to forbid the packet transmission from this host.

In order to facilitate the management and maintenance of network administrators, it can be able to perform auto-protect and save relevant warning message. As to those users who have been forbidden, administrator can configure it as manual recovery or automatic recovery.

The whole process on the Switches are as follows:

Enable arp anti-flood function, report the ARP packet to CPU, identify different flow according to the source MAC address of ARP packet.

Configure the safe ARP rate. Switch will take it as ARP attack if the rate exceeds the configured threshold value.

If you select the above deny-all command, when one ARP traffic exceeds configured threshold value, the Switch will put this MAC address into blackhole address list and forbid the this address's packet transmission according to source MAC address.

If you select the above deny-arp command, the Switch will not deal with the subsequent ARP message based on source mac address when ARP traffic is larger than the configured threshold.

As to the recover for those messages which are forbidden to forward, administrator could configure the recovery time as automatic recovery or handwork recovery.

17.5.3 Anti-Spoofing Configuration

Operation	**Command**	Remarks
Enter global configuration mode	**configure terminal**	
Enable arp anti-spoofing	**arp anti-spoofing**	
Disable arp anti-spoofing	**no arp anti-spoofing**	
Configure the approach for unknown message: discard or flooding	**arp anti-spoofing unknown {diacard | flood}**	unknown arp packet refers to the ip of those arp packets which cannot match with any item of the ip options of arp static table, ip-soure-guard binding table, dhcp-snooping table. In other word, this ip does not exist in the table.

17.5.4 Host Protection Configuration

Configure ip+port binding when configuring to discard the unknown arp packet, and then the arp packet of this ip can flood to other ports only via this valid port. If the arp packet of this ip enters from other ports, it will be discarded.

Operation	**Command**	Remarks
Enter global configuration mode	**configure terminal**	
Enable arp anti-spoofing	**arp anti-spoofing**	
Configure the process mode of unknown ARP message to be discard	**arp anti-spoofing unknown flood**	
Configure host protection	**host-guard bind ip** *ipaddress* **interface ethernet** *device/slot/port*	
Delete host protection	**no host-guard bind { ip** *ipaddress* **| interface ethernet** *device/slot/port* **}**	

17.5.5 Configure Source-MAC Consistency Inspection

As to a certain ARP attack packet, their source-MAC in the head of Ethernet data is different from the source-MAC in ARP protocol. After enabling source-MAC consistency inspection, Switch will inspect whether the Ethernet source MAC address in ARP packet is the same as the source MAC in ARP protocol packet. If they are not the same, Switch will discard the packet.

This function is disabled by default.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enable arp anti-spoofing	**arp anti-spoofing**	
Enable source-mac consistency inspection	**arp anti-spoofing valid-check**	
Disable source-mac consistency inspection	**no arp anti-spoofing valid-check**	

17.5.6 Configure Anti-Gateway-Spoofing for Layer-3 Equipment

When the layer-3 Switch acts as the gateway for some certain LAN equipment, this Switch will list the attacker who wants to simulate the switch into blacklist, and it will send gratuitous ARP to notice the LAN equipment that “It is I who is the correct gateway”.

This function is disabled by default.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enable arp anti-spoofing	**arp anti-spoofing**	
Enable anti-gateway-spoofing	**arp anti-spoofing deny-disguiser**	
Disable anti-gateway-spoofing	**no arp anti-spoofing deny-disguiser**	

17.5.7 Configure the Trust Port

The trust port will not perform attack and spoof check when it receives arp message.

Operation	Command	Remarks
Enter port configuration mode	**interface ethernet** *device/slot/port*	
Configure the port to be trust port	**arp anti trust**	Untrust by default
Recover the port to be untrust port	**no arp anti trust**	

17.5.8 Anti-Flood Attack Configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable anti-ARP flooding attack	**arp anti-flood**	required
Disable anti-ARP flooding attack	**no arp anti-flood**	
Configure safety trigger threshold	**arp anti-flood threshold** *threshold*	optional By default, the safety trigger threshold is 16PPS.
Configure approach for the attacker	**arp anti-flood action {deny-arp|deny-all}**	optional By default, approach for the attacker is deny arp.
Configure automatically banned user recovery time	**arp anti-flood recover-time** *time*	optional Configurable time range to be<0-1440> minutes; if you set the value to be 0, it means that you should manually restored. By default, the user automatically banned recovery time is 10 minutes.
Banned user manual resume forwarding	**arp anti-flood recover {***H:H:H:H:H:H* | **all}**	optional
To bind the dynamic blackhole MAC to be static blackhole MAC	**arp anti-flood bind blackhole {***H:H:H:H:H:H* **| all}**	Only when the process mode is deny-all can it generate as the mac of non-stationary black hole
Enter interface configuration mode	**interface ethernet** *device/slot/port*	
Configure threshold limit values for the port	**arp anti-flood threshold** *threshold*	It takes effect only when threshold limit values of the port is smaller than global threshold limit values.

17.5.9 Display and Maintain

Operation	**Command** 行	Remarks
Display arp anti-spoofing configuration	**show arp anti-spoofing**	
Display ARP anti-flood configuration and attackers list	**show arp anti-flood**	It can be executed under any mode.
Display the state of interface	**show arp anti interface**	

17.5.10 Example for Anti- ARP Spoofing Configuration

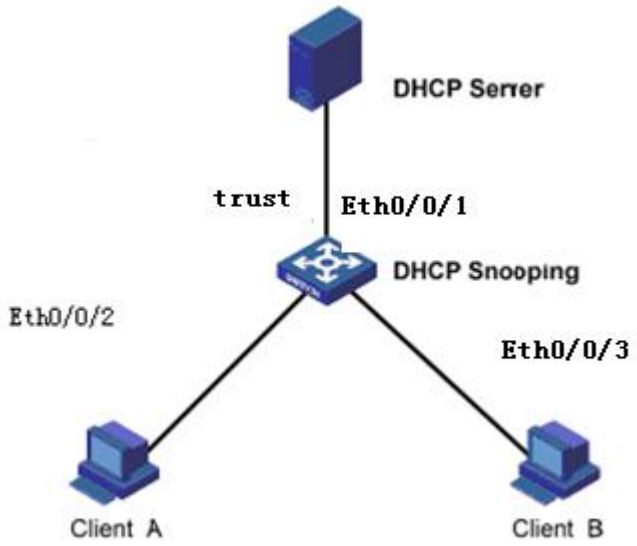
Network requirements

As shown in the figure, Eth0/0/1 port of SwitchA connects to DHCP server, Eth0/0/2 port and Eth0/0/3 port connect to Client A and Client B respectively. Moreover, these three ports are geared to VLAN 1.

Combine dhcp-snooping to use

Enable DHCP Snooping, set port Eth0/0/1 as the trust port of DHCP Snooping to enable anti-arp spoofing.

Network diagram



Configuration steps

Enable DHCP Snooping

Switch(config)#dhcp-snooping

Set port Ethernet 0/0/1 as the trust port of DHCP Snooping

Switch(config-if-ethernet-0/0/1)#dhcp-snooping trust

Config DHCP Snooping mode of port successfully.

Ip-soure-guard binding table

Switch(config)#ip-source-guard bind ip 192.168.5.10 mac 40:16:9f:f2:75:a8 in

terface ethernet 0/0/3 vlan 1

Add ip-source-guard bind entry successfully.

Enable anti-arp spoofing function

Switch(config)#arp anti-spoofing

Switch(config)#arp anti-spoofing unknown discard

Switch(config)#interface ethernet 0/0/1

Switch(config-if-ethernet-0/0/1)#arp anti trust

Client A DHCP obtains ip to form the dhcp-snooping clients table.

```
Switch(config)#show dhcp-snooping clients
DHCP client information:
d - days, h - hours, m - minutes, s - seconds
IPAddress        mac                vlan  port     LeaseTime      ExceedTime
192.168.5.13     00:00:00:00:32:33  1      e0/0/2  10m0s          7m53s

Total entries: 1. Printed entries: 1.
```

Client A forwards arp quest message to dhcpserver, dhcpserver can be able to receive this arp quest message

Client B configure static ip=192.168.5.10 mac=40:16:9f:f2:75:a8，Client B forwards arp quest message to dhcpserver, dhcpserver can be able to receive this arp quest message

If client B enable anti-arp spoofing, source ip of arp message=Client A， the equipment will discard the message if it found this arp message is spoof message.

This instance estimates whether this arp message is spoof message or not according to dhcp-snooping clients table or ip-soure-guard bind table. In addition, ayer-3 equipment can be able to realize this function via static arp table. All of this shares the same principle, no more tautology here.

18.Single Spanning Tree

18.1 STP Overview

Single spanning tree includes spanning tree (STP) and rapid spanning tree (RSTP).

18.1.1 STP Practical Application

STP is a part of the IEEE802.1D bridge protocol, its primary function is to clear the layer 2 loop from the topology.

18.1.2 Bridge Protocol Data Unit

In order to run STP, user needs to share information between switches. They shared information is the bridge protocol data unit, which is sent in the form of multicast information, and only other Layer 2 devices to listen to the bridge data unit. The switch learns the network topology using BPDU: what devices are connected to other devices, and whether there are certain Layer 2 loops in the network based on this topology.

If some loops are found, the switch disables one or some of the ports in this topology to ensure that there are not loops in the network. That is, in a switch network, only one path is available from one device to any other device. If there is any change in the layer 2 network, such as a link breaks down, new links add, new switches add or a switch fails, switches in the network share this information, which causes the STP algorithm to re-create implementation, and produces a new acyclic topology.

18.1.3 Basic Concepts of STP

Root Bridge

After the STP algorithm runs, the first step is to elect the root switch. The root switch is at the top of the spanning tree topology. The switch with the lowest switch ID is selected as the root. The switch ID consists of two parts:

- The priority of the switch. By default, the priority of all switches is 32,768.
- The MAC address of the switch.

Administrator can specify a switch as the root by changing the switch ID. When the network topology changes, such as a root switch fails or a new switch is added to the network, the root switch election process is re-triggered.

Root Port

After selecting the root switch, you also need to select a port closest to the root switch on all the non-root switches in the network to communicate with the root switch.

Designated Bridge

In each individual LAN there is a switch called the designated bridge, which belongs to the bridge of the least expense in the LAN root path. A root switch is an election bridge for all the LANs to which it is connected.

Designated Port

After electing the root switch and the root port, you need to elect a port for reaching the root switch in each link, which is the designated port. Specifying a port requires the following conditions:

- In two switches of a link, the ports on the switch that have the lowest accumulated path cost to the root switch will be selected. If the cumulative cost of the two switches is the same, select the switch with the lowest switch ID.
- If multiple links on the same switch are connected to the root switch, the switch port with the lowest priority is selected as the designated port. If these ports have the same priority, select the port with the lowest physical port number as the designated port.

18.2 RSTP Introduction

Rapid Spanning Tree Protocol is an optimized version of the STP protocol. “Rapid” is shown that when a port is selected as a root port and a designated port, the delay of entering the forwarding state is greatly shortened under certain condition, thus which will shorten the time required for the network to reach the topology stability finally.

In RSTP, the condition for rapid transition of the port state of the root port is: the old root port on the device has stopped to forward data and the upstream port has started to forward data.

In RSTP, the condition for rapid transition of the port state of a specified port is: the specified port is an edge port or the specified port is connected to a point-to-point link. If the specified port is an edge port, the specified port can directly enter the forwarding state. If the specified port connects with a point-to-point link, the device can enter the forwarding state immediately after receiving the handshake from the downstream device.

RSTP can converge quickly. However, there are the following drawbacks as STP: all the bridges in a LAN share a spanning tree, and cannot block redundant links according to VLAN, all VLAN messages are forwarded along a spanning tree.

18.3 Configure STP/RSTP

The STP / MSTP configuration command is the same

18.3.1 Enable the Spanning Tree

After global startup spanning tree, all the ports will participate in the calculation of the spanning tree topology by default. If the administrator wishes to exclude some ports from spanning tree calculation, user can use the no spanning-tree command in the configuration mode to disable spanning tree function.

Enable STP

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enable the spanning tree globally	**spanning-tree**	This command is valid for STP / RSTP / MSTP
Shut down the spanning tree globally	**no spanning-tree**	
Select the spanning tree mode	**spanning-tree mode** { stp | rstp | mstp }	optional
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Port Open the spanning tree	**spanning-tree**	optional
Port Close the spanning tree	**no spanning-tree**	

Note: after the spanning tree is enabled globally, the system works in RSTP mode by default.

18.3.2 Set the Bridge Priority of the Switch

The bridge priority size of the switch determines whether the switch can be selected as the root of the spanning tree. By configuring a smaller bridge priority, you can specify that a switch becomes the root of the spanning tree.

By default, the bridge priority of the switch is 32768.

Set STP/RSTP priority

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Set STP priority	**spanning-tree priority** *bridge priority*	optional

18.3.3 Configure Time Parameter

The switch has three time parameters: Forward Delay、 Hello Time and Max Age. User can configure these three parameters on the switch for STP / RSTP calculation.

Configure time parameter

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure the hello message interval	**spanning-tree hello-time** *seconds*	optional
Configure the forwarding delay of the system	**spanning-tree forward-time** *seconds*	optional
Configure the aging time of the system	**spanning-tree max-age** *seconds*	optional

Note:

An excessively long Hello Time value causes the bridge to consider a link failure and start recalculate the spanning tree due to link message loss. An excessively short Hello time value causes the bridge to send configuration information frequently, and increase internet and CPU load. Hello Time is in the range of 1 to 10 seconds. It is recommended to use the default value 2 seconds. The hello time must be less than or equal to forward delay-2.

If the forward delay is too small, a temporary redundant path may be introduced; if the forward delay is too large, the network may not resume communication for a long time. The value of forward delay is in the range of 4 to 30 seconds. It is recommended to use the default value 15 seconds. The time of forward delay must be greater than or equal to hello time + 2.

Max Age sets the maximum time interval for STP message. If it times out, the message is discarded. If this value is too small, spanning tree calculation may be more frequent, network congestion may be mistaken for network link failure; if this value is too large, it is not conducive to detect link failure timely. The Max Age value ranges from 6 to 40 seconds. The Max Age time value depends on the network diameter of the switching network. The default value 20 seconds is recommended. Max Age must be greater than or equal to 2 * (Hello Time + 1), and less than or equal to 2 * (Forward Delay-1).

18.3.4 Configure Path Cost of Port

By configuring the path cost of a port, the port can become a root port or a designated port easily.

The path cost of a port depends on the link rate of the port. The larger the link rate is, the smaller the parameter is. STP can automatically detect the link rate of the current port and translate it into the corresponding path cost.

Configuring the path cost of an Ethernet port will cause the spanning tree to recalculate. The path cost of a port ranges from 1 to 65,535. It is recommended to use the default value. The STP protocol calculates the path cost of the current port. By default, the path cost is determined based on the speed of a specific port.

The default value of port path cost is determined by the port speed. When the port speed is 10M, the default value is 20,00,000. The default value is 200,000 when 100 M, and 20,000 is the default value at 1000M. When the port rate is not available, the path cost is 200,000 by default.

Configure the path cost of a port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	optional
Modify the path cost of the port	**spanning-tree cost** *path-cost*	optional

Several ports are aggregated into one aggregation group. The default path cost of the aggregation group is p [1 - (n-1/10)], p is the path cost of the port, and n is the number of the aggregation group ports.

Of course, you can manually set the path cost of the aggregation group.

Configure aggregation group port cost

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Set the path cost of the aggregation group	**channel-group** *id* **spanning-tree cost** *path-cost*	optional

18.3.5 Configure the Priority of Port

By configuring the priority of a port, you can make a port more easily become a root port.

The lower the priority value is, the higher the priority is. Changing the priority of an Ethernet port will cause the spanning tree to recalculate. The spanning tree priority of a port ranges from 0 to 240. It must be an integer multiple of 16. By default, the port spanning tree priority is 128.

Configure the priority of port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure the STP priority of the port	**spanning-tree port-priority** *priority*	optional

18.3.6 Configure Mcheck Function

The switch operating in RSTP mode can connect with an STP switch to ensure compatibility. However, after the neighbor changes the work mode for RSTP, the two ports connected to each other still work in STP mode by default. The Mcheck function is used to force the port to send RSTP message and confirm whether the adjacent port can work in RSTP mode. If yes, the switch automatically works in RSTP mode.

Note: the mcheck function requires that the port must send BPDU. So it is only useful on the specified port.

Configure mcheck function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Perform the mcheck function	**spanning-tree mcheck**	optional

18.3.7 Configure Point-to-Point Link

In RSTP, a port enters a forwarding state quickly. It is required that the port must be a point-to-point link, and not a shared media link. User can manually specify the link type of a port, or determine the link type automatically based on the port duplex mode.

The port is in the configured mode, if the port is in full duplex mode, it is judged as a

point-to-point link. If it is half duplex, it is a non-point-to-point link.

When the switch is in the force true mode, the port is a point-to-point link.

When the switch is in the force false mode, the port is non-point-to-point link.

Configure point-to-point link

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure the port link type	**spanning-tree point-to-point {auto | forcetrue| forcefalse}**	optional

18.3.8 Configure Port to Edge Port

An edge port refers to a port connected with a terminal device such as a host, and these ports can enter the forwarding state in a short time after the linkup. The edge port is valid only for RSTP.

Configure port to edge port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure the port as a edge port	**spanning-tree portfast**	
Configure the port as a non-edge port	**no spanning-tree portfast**	

18.3.9 Set Port to Send the Maximum Rate of BPDU

The maximum rate of BPDU message sent by port is the maximum number of BPDU message sent in each Hello time.

By default, the rate of BPDU sent by port sends 3 per Hello time.

Set port to send the Maximum Rate of BPDU

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Set port to send the Maximum Rate of BPDU	**spanning-tree transit-limit***transit-limit*	optional

18.3.10 Configure Root Protection of a Port

The root bridge may receive the higher priority configuration message due to misconfiguration of the maintenance personnel or malicious attacks in the network. Thus, the current root bridge can lose the status of the root bridge and cause incorrect network topology changes. Assume that the original traffic is forwarded over a high-speed link, this illegal change will cause the traffic passing through the high-speed link to be traced to the low-speed link, which results in network congestion. Root protection can prevent from happening.

For a port with root protection enabled, the port role can only be the designated port. Once a high-priority configuration has been received on the port, there are two options for configuring the status of these ports:

Block-port: The port state is set for discarding, discarding BPDU configuration messages and not forwarding data packets.

Drop-packets: Port state is forwarding, only the BPDU configuration is discarded, and ordinary packets can be forwarded.

Configure Root Protection of a Port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure the action of the root protection port to process message	**spanning-tree root-guard action {block-port | drop-packets}**	

Enter the port configuration mode	**interface ethernet** *interface-num*	-
Enable the root protection function of the port	**spanning-tree root-guard**	optional
Disable the root protection function of the port	**no spanning-tree root-guard**	

18.3.11 Configure Loop-guard Function

Loop-guard function: To prevent a blocked port because of un-normal link (not two-way communication) not receive the BPDU configuration information, which becomes forwarding state. When the port is configured with this option, the port remains blocked even if BPDU configuration is not received.

Configure port loop-guard function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface * thernet** *interface-num*	
Enable the loop-guard function	**spanning-tree loop-guard**	not share with root-guard
Disable the loop-guard function	**no spanning-tree loop-guard**	

18.3.12 Configure Bpdu-guard Function

For an access layer device, an access port is usually directly connected with a user terminal (such as a PC) or a file server. In this case, the access port is configured as an edge port to implement rapid transition. When these ports receive BPDU messages, the system will automatically set these ports as non-edge ports and recalculate spanning trees to cause network topology changes. These ports should normally not receive BPDU message. If someone forges BPDU to attack the device maliciously, the network will become unstable.

The device provides the BPDU guard function to prevent such attacks: after the BPDU guard function is enabled on a device, if a port configured with an edge port attribute receives BPDU message, the device will SHUTDOWN the port and prompt the user with the syslog information. The port is restored by manual NO SHUTDOWN.

Configure the bpdu-guard function on a port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enable bpdu-guard globally	**spanning-tree bpdu-guard**	In global mode, this function is enabled on all ports
Disable bpdu-guard globally	**no spanning-tree bpdu-guard**	
Enter the port configuration mode	**interface ethernet** *interface-num*	
Enable the bpdu-guard function	**spanning-tree bpdu-guard**	This function takes effect on a port
Disable the bpdu-guard function	**no spanning-tree bpdu-guard**	

Note:

The port BPDU guard function takes effect only on the port configured with the edge port attribute. For a port that is configured with edge port attribute, receives BPDU message from another port and becomes a non-edge port again, if the BPDU guard function is enabled, the port can take effect only when it is restarted as an edge port.

18.3.13 Configure Bpdu-filter Function

After the bpdu-filter is set on the edge port, the device will discard the received BPDU message, and the port will not send BPDU message.

Configure bpdu-filter function for port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enable bpdu-filter globally	**spanning-tree bpdu-filter**	In global mode, this function is enabled on all ports
Disable bpdu-filter globally	**no spanning-tree bpdu-filter**	
Enter the port configuration mode	**interface ethernet** *interface-num*	
Enable the bpdu-filter function	**spanning-tree bpdu-filter**	This function takes effect on a port
Disable the bpdu-filter function	**no spanning-tree bpdu-filter**	

18.3.14 Bpdu-car Function

When a large number of bpdu messages are on the CPU, it causes likely CPU busy, so the bpdu-car function limits the rate of bppu message on the cpu.

Configure BPDU-Car

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enable/disable Port-Car	[no]port-car	Optional, enable by default
Configure the bpdu rate of cpu	port-car-rae ***value***	Optional, (number of ports * 30) pps by default
Enter the port configuration mode	interface ethernet ***port-number***	-
Enable/disable Port-Car	[no]port-car	Optional, open by default
Configure the bpdu rate of cpu	port-car-rae ***value***	Optional, 30pps by default
View the configuration information	show port-car	optional

18.3.15 Discard-BPDU Function

The Discard-bpdu function is used to drop spanning tree message. If the device does not want to receive BPDU message from other networks and cause the switch spanning tree to vibrate. This function can be opened.

This function is usually enabled on the edge port.

The Discard-BPDU function is disabled by default. Global configuration and port configuration are mutually exclusive: globally, all ports are enabled. If you only need to enable certain designated ports and other ports are not enabled, you need not configure them globally to directly enter the specified port enabling function.

Configure the global Discard-BPDU

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enable/disable BPDU	[no] discard-bpdu	Required, off by default
View the configuration information	show discard-bpdu	optional

Configure port Discard-BPDU

operation	command	remark
Enter the port configuration mode	interface ethernet ***port-num***	-
Enable/disable BPDU	[no]discard-bpdu	Required, off by default
View configuration information	show discard-bpdu	optional

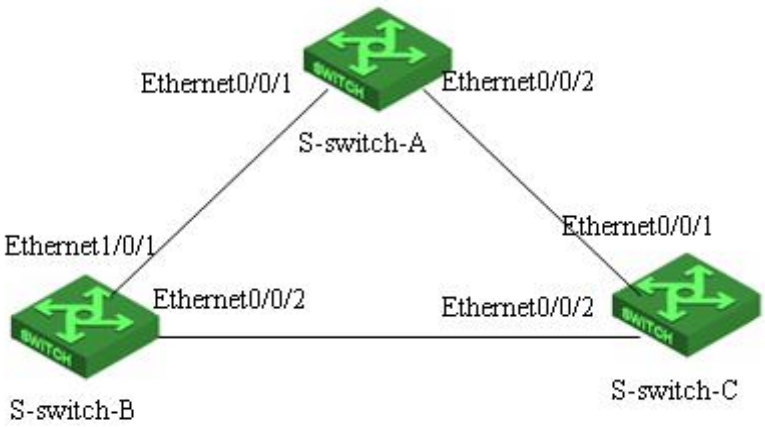
18.3.16 Display and Maintenance

After you complete the above configuration, you can use the following command to view the configuration.

STP/RSTP display and maintenance

operation	command	remark
Display the spanning tree status of port, that is, the spanning tree configuration parameters. Root switch MAC address and so on.	**show spanning-tree interface [brief [ethernet** *interface-list* **]**	All modes are executable

18.3.17 RSTP Configuration Example



I. Network requirement

As shown in above figure, S-switch-A is the core switch and acts as the root bridge. S-switch-B acts as the designated bridge. The links connecting S-switch-B and S-switch-C are backup links. S-switch-B, S-switch-A, or S-switch-C fails, the backup link works.

2、Configuration procedure

The spanning tree default mode is RSTP. Therefore, you need to enable the global spanning tree switch when you enable RSTP. The RSTP takes the default time.

The configuration of Switch A

#configure port Ethernet0/0/1 and port Ethernet0/0/2 for trunk port

S-switch-A(config)#interface range ethernet 0/0/1 ethernet 0/0/2

S-switch-A(config-if-range)#switchport mode trunk

Configure the priority of S-switch-A bridge to 0. Ensure that S-switch-A is the root bridge.

S-switch-A(config)#spanning-tree priority 0

#Start RSTP globally.
S-switch-A(config)#spanning-tree
S-switch-A(config)#spanning-tree mode rstp

The configuration of Switch B

#configure port Ethernet0/0/1 and port Ethernet0/0/2 for trunk port。
S-switch-B(config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-B(config-if-range)#switchport mode trunk
S-switch-B(config-if-range)#exit

Configure the bridge priority of S-switch-B to 4096, Ensure that S-switch-B is the designated bridge, and the path cost of configuration Ethernet0/0/1 and Ethernet0/0/2.
S-switch-B(config)#spanning-tree priority 4096
S-switch-B(config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-B(config-if-range)#spanning-tree cost 10
S-switch-B(config-if-range)#exit

Start RSTP globally.
S-switch-B(config)#spanning-tree
S-switch-B(config)#spanning-tree mode rstp

The configuration of Switch C

#configure Ethernet0/0/1 and Ethernet0/0/2 for trunk port.
S-switch-C(config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-C(config-if-range)#switchport mode trunk
S-switch-C(config-if-range)#exit

the path cost of configuration Ethernet0/0/1 and Ethernet0/0/2 is 10. Ensure that the link connecting S-switch-B and S-switch-C is the primary link.
S-switch-C(config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-C(config-if-range)#spanning-tree cost 10
S-switch-C(config-if-range)#exit

Start RSTP globally.
S-switch-C(config)#spanning-tree
S-switch-C(config)#spanning-tree mode rstp
S-switch-C(config)# spanning-tree priority 32768

Verify the configuration

Run the display command on Switch A and view the election result and port status of RSTP. The results are as follows:

```
S-switch-A(config)#show spanning-tree interface ethernet 0/0/1 ethernet 0/0/2
  The bridge is executing the IEEE Rapid Spanning Tree protocol
  The bridge has priority 0, MAC address: 000a.5a13.b13d
  Configured Hello Time 2 second(s), Max Age 20 second(s),
  Forward Delay 15 second(s)
  Root Bridge has priority 0, MAC address 000a.5a13.b13d
  Path cost to root bridge is 0
  Stp top change 3 times

  Port e0/0/1 of bridge is Forwarding
  Spanning tree protocol is enabled
  remote loop detect is disabled
  The port is a DesignatedPort
```

Port path cost 200000
Port priority 128
root guard enabled and port is not in root-inconsistent state
Designated bridge has priority 0, MAC address 000a.5a13.b13d
The Port is a non-edge port
Connected to a point-to-point LAN segment
Maximum transmission limit is 3 BPDUs per hello time
Times: Hello Time 2 second(s), Max Age 20 second(s)
Forward Delay 15 second(s), Message Age 0
sent BPDU: 54
TCN: 0, RST: 54, Config BPDU: 0
received BPDU: 10
TCN: 0, RST: 10, Config BPDU: 0

Port e0/0/2 of bridge is Forwarding
Spanning tree protocol is enabled
remote loop detect is disabled
The port is a Designated Port
Port path cost 200000
Port priority 128
root guard enabled and port is not in root-inconsistent state
Designated bridge has priority 0, MAC address 000a.5a13.b13d
The Port is a non-edge port
Connected to a point-to-point LAN segment
Maximum transmission limit is 3 BPDUs per hello time
Times: Hello Time 2 second(s), Max Age 20 second(s)
Forward Delay 15 second(s), Message Age 0
sent BPDU: 16
TCN: 0, RST: 17, Config BPDU: 0
received BPDU: 3
TCN: 0, RST: 3, Config BPDU: 0

S-switch-A is elected as the root bridge because S-switch-A has the highest priority in the entire network. Therefore, Ethernet0 / 0/1 and Ethernet0 / 0/2 of S-switch-A are designated ports. They are in the forwarding state.

Run the display command on S-switch-B and view the election result and port status of RSTP. The following information is displayed:

S-switch-B (config)#show spanning-tree interface ethernet 0/0/1 ethernet 0/0/2
The bridge is executing the IEEE Rapid Spanning Tree protocol
The bridge has priority 4096, MAC address: 0000.0077.8899
Configured Hello Time 2 second(s), Max Age 20 second(s),
Forward Delay 15 second(s)
Root Bridge has priority 0, MAC address 000a.5a13.b13d
Path cost to root bridge is 10
Stp top change 3 times

Port e0/0/1 of bridge is Forwarding
Spanning tree protocol is enabled
remote loop detect is disabled
The port is a RootPort
Port path cost 10
Port priority 128

root guard disabled and port is not in root-inconsistent state
Designated bridge has priority 0, MAC address 000a.5a13.b13d
The Port is a non-edge port
Connected to a point-to-point LAN segment
Maximum transmission limit is 3 BPDUs per hello time
Times: Hello Time 2 second(s), Max Age 20 second(s)
Forward Delay 15 second(s), Message Age 0
sent BPDU: 21
TCN: 0, RST: 12, Config BPDU: 9
received BPDU: 204
TCN: 0, RST: 202, Config BPDU: 2

Port e0/0/2 of bridge is Forwarding
Spanning tree protocol is enabled
remote loop detect is disabled
The port is a DesignatedPort
Port path cost 10
Port priority 128
root guard disabled and port is not in root-inconsistent state
Designated bridge has priority 4096, MAC address 0000.0077.8899
The Port is a non-edge port
Connected to a point-to-point LAN segment
Maximum transmission limit is 3 BPDUs per hello time
Times: Hello Time 2 second(s), Max Age 20 second(s)
Forward Delay 15 second(s), Message Age 1
sent BPDU: 191
TCN: 0, RST: 188, Config BPDU: 3
received BPDU: 13
TCN: 0, RST: 5, Config BPDU: 8

The priority of S-switch-B is lower than that of S-switch-A. Ethernet0 / 0/1 of S-switch-B is the root port and in the forwarding state. At the same time, because S-switch-B has a higher priority than S-switch-C, Ethernet 0/0/2 of S-switch-B is calculated as the specified port and in the forwarding state.

Run the display command on S-switch-C and view the election result and port status of RSTP. The following information is displayed:

S-switch-C(config)#show spanning-tree interface ethernet 0/0/1 ethernet 0/0/2
The bridge is executing the IEEE Rapid Spanning Tree protocol
The bridge has priority 32768, MAC address: 000a.5a13.f48e
Configured Hello Time 2 second(s), Max Age 20 second(s),
Forward Delay 15 second(s)
Root Bridge has priority 0, MAC address 000a.5a13.b13d
Path cost to root bridge is 20
Stp top change 3 times

Port e0/0/1 of bridge is Forwarding
Spanning tree protocol is enabled
remote loop detect is disabled
The port is a RootPort
Port path cost 10
Port priority 128
root guard disabled and port is not in root-inconsistent state
Designated bridge has priority 0, MAC address 000a.5a13.b13d
The Port is a non-edge port

Connected to a point-to-point LAN segment
Maximum transmission limit is 3 BPDUs per hello time
Times: Hello Time 2 second(s), Max Age 20 second(s)
Forward Delay 15 second(s), Message Age 0
sent BPDU: 3
TCN: 0, RST: 3, Config BPDU: 0
received BPDU: 396
TCN: 0, RST: 396, Config BPDU: 0

Port e0/0/2 of bridge is Discarding
Spanning tree protocol is enabled
remote loop detect is disabled
The port is a AlternatePort
Port path cost 10
Port priority 128
root guard disabled and port is not in root-inconsistent state
Designated bridge has priority 4096, MAC address 0000.0077.8899
The Port is a non-edge port
Connected to a point-to-point LAN segment
Maximum transmission limit is 3 BPDUs per hello time
Times: Hello Time 2 second(s), Max Age 20 second(s)
Forward Delay 15 second(s), Message Age 1
sent BPDU: 8
TCN: 0, RST: 8, Config BPDU: 0
received BPDU: 418
TCN: 0, RST: 418, Config BPDU: 0

The priority of S-switch-C is lower than that of S-switch-A and S-switch-B. The cost of the route from Ethernet 0/0/1 to the root bridge is lower than that of Ethernet 0/0/2. So Ethernet 0/0/ 1 is calculated as the root port and in the forwarding state. Ethernet0 / 0/2 is calculated as the alternate port and in the discarding state.

19. Multiple Spanning Tree Configuration

19.1 MSTP Overview

The Spanning Tree Protocol cannot migrate port state quickly. Even on the point-to-point link or edge port, it must wait for a delay of two times forwarding delay, the port can be transferred to the forwarding state.

RSTP (Rapid Spanning Tree Protocol) can quickly converge, but there are following shortcomings the same as STP: all the bridges of LAN share a spanning tree, and cannot block

redundant links according to VLAN, all VLAN messages are forwarded along a spanning tree.

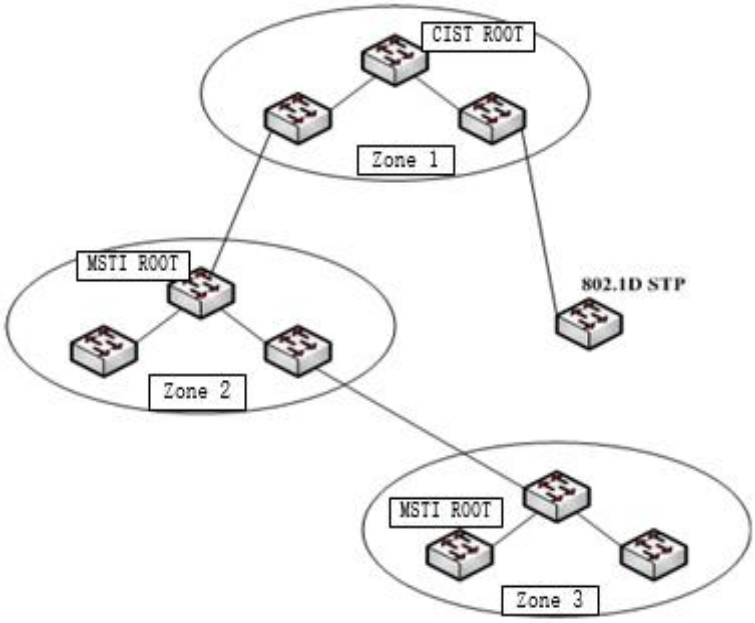
Multiple spanning tree protocol (MSTP) prunes the loop network into an acyclic tree network to avoid the proliferation and infinite looping of messages in the loop network. It also provides multiple redundant for data forwarding. Load balance of VLAN data is achieved during data forwarding.

MSTP is compatible with STP and RSTP, which can compensate for STP and RSTP defects. It can quickly converge and distribute traffic of different VLANs along its own path. Thus it can provide a better load sharing mechanism for redundant link.

19.1.1 Bridge Protocol Data Unit

MSTP uses BPDU to calculate the spanning tree as STP/RSTP. The BPDU carries MSTP configuration information on the switch.

19.1.2 Basic Concepts in MSTP



MSTP topology example

Above figure shows an example MSTP network, include three multi-spanning tree areas and a switch running 802.1D STP protocol.

1. MSY region

Above figure shows an example MSTP network. Area 1, 2, and 3 are three MST regions.

MST region is made up of multiple switches in the switched network and the network segments among them. These switches are MSTP-enabled and have same domain name, same VLAN-to-spanning-tree configuration, same MSTP revision level configuration, and physical link connection.

A switching network can have multiple MST regions. User can use the MSTP configuration

command to divide multiple switches into the same MST region.

2. CIST

Common and internal spanning tree is made up of all individual switches and connected LANs. These switches may belong to different multi-spanning tree areas, or may run a traditional STP or RSTP protocol. Switch running two protocols in a multi-spanning tree network is considered to be only in its own area.

After the network topology is stable, the entire CIST selects a CIST root bridge. In each area, the root bridge in the CIST area is elected as the shortest path from the intra-area to the CIST root.

3. CST

CST is short of common spanning tree. If each multi-spanning tree area is treated as a single switch, the CST is the spanning tree that connects all these "individual switches". As shown in Figure 1, zones 1, 2, 3, and STP switches together form the CST of the network.

4. IST

IST is short of internet spanning tree, which refers to the part of the CIST in a multi-spanning tree area, and can also be understood that IST and CST together form the CIST.

5. MSTI

MSTI is short of multiple spanning tree instances. MSTI protocol allows different VLANs to be divided into different spanning trees, thereby a plurality of spanning tree instances is established. Normally a spanning tree instance with the number 0 means CIST, which can be extended to the whole network. The spanning tree instance that starts from 1 is in the interior of a certain area. Each spanning tree instance can be assigned multiple VLANs. Initially, all VLANs are assigned in the CIST.

In a multi-spanning tree area, all MSTI are independent of each other. They can select different switches as their own roots. For example, in the area 3 of Figure 1, the root bridge of MSTI01 may be the switch in the lower left corner, and the MSTI00, that is, the root bridge of the CIST may be the switch in the middle position.

6. CIST Root Bridge

CIST root is the bridge with the highest priority bridge identity in the entire network.

7. CIST external root path cost

CIST external root path cost is the path cost between the bridge and the CIST root. Only changes across the MST region. The CIST external root path of all the bridges costs the same in the same MST region.

8. CIST regional root

CIST regional root is the least expensive bridge in the external root path of each CIST root. In fact, it is the root bridge of the IST, or the virtual bridge of the MST region. If the CIST root is in an MST region, the CIST root is also the root bridge of the CIST region in the MST region.

9. CIST internal root path cost

CIST internal root path cost is the root path cost of the bridge in the MST region to the CIST regional root bridge, which is valid only in this region.

10. CIST designated bridge

CIST designated bridge is same as STP designated bridge.

11. MSTI regional root

MSTI regional root is MSTI root bridge in each MST region. The root bridge of the MSTI may

not be the same for different MSTI.

12. MSTI internal root path cost

MSTI internal root path cost is the root path cost of the bridge in the MST region to the MSTI regional root, which is valid only in the region.

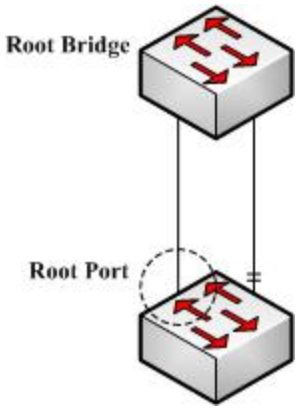
13. MSTI designated bridge

MSTI designated bridge is same as STP designated bridge.

19.1.3 Role of Port

The MSTP protocol has similar port role assignment as RSTP.

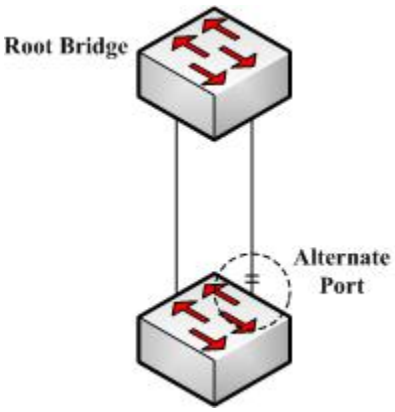
Root Port



Root port

The root port represents the path from the current switch to the network root bridge, which has the smallest root path cost.

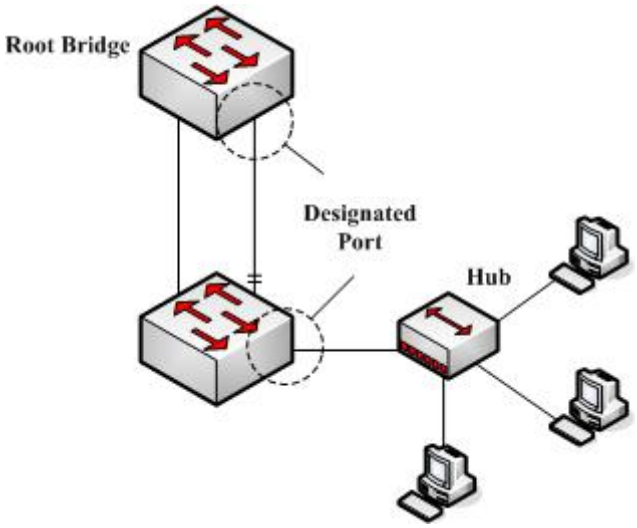
Alternate Port



Alternate port

The alternate port acts as the backup of the current switch to the root bridge of the network. When the root port fails, the alternate port can work immediately as the new root port.

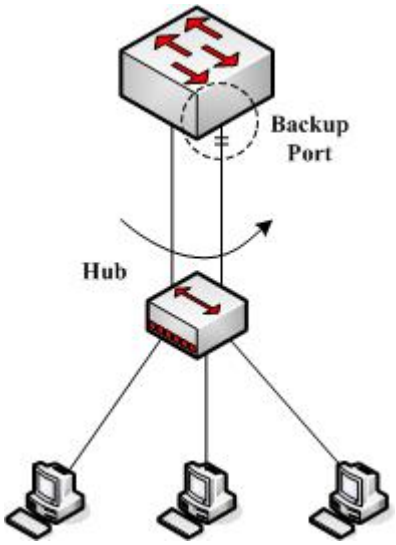
Designated Port



Designated port

The designated port can be connected with a downstream switch or a local area network (LAN) , which can act as the path of the LAN to the root bridge of the network.

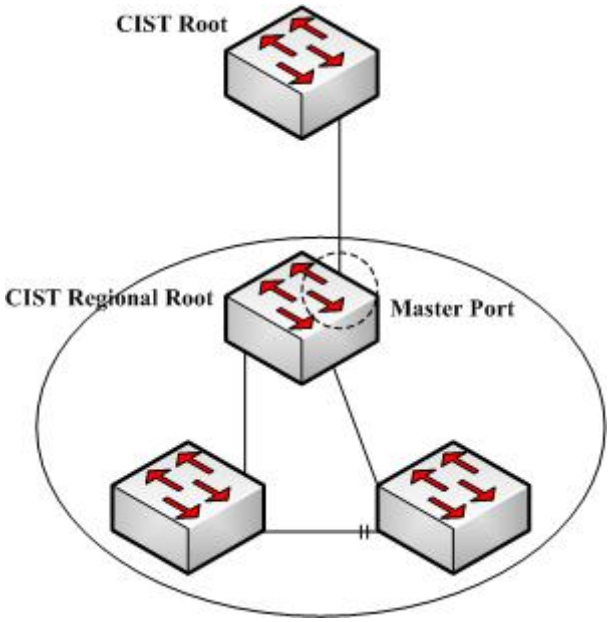
Backup Port



Backup port

When the two ports of the switch are directly connected or connect to the same LAN, the port with the lower priority becomes the backup port (the higher one becomes the designated port). If the specified port fails, the backup port turns to the designated port to start working.

Master port



Master port

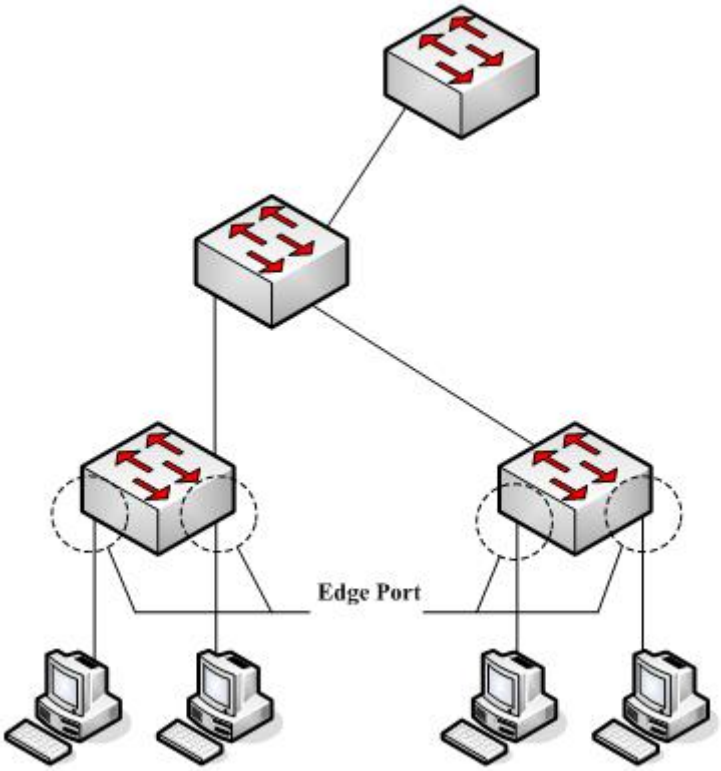
The master port acts as the shortest path connecting the CIST root bridge in multiple spanning tree regions. The master port is the root port of the root bridge in the CIST.

Boundary port

The concept of boundary port is slightly different in CIST with that in each MSTI. In the CIST, a boundary port represents a port that connects to another multi-spanning tree region. In the MSTI, a boundary port role indicates that the spanning tree instance is no longer extended at this port.

Edge Port

In the RSTP and MSTP protocols, edge port denotes that port directly connects to network host. These ports do not need to wait to come into the forwarding state, and do not cause a loop on the network.



Edge port

In the initial case, the MSTP (including RSTP) protocol considers all ports to be edge ports, thus ensure rapid establishment of network topology. If a port receives a BPDU from another switch, the port returns to the normal state from the edge state. If it receives an 802.1D STP BPDU, the port waits for a forward delay of two times to enter the forwarding state.

19.2 MSTP Election Calculation

19.2.1 MSTP Protocol Message

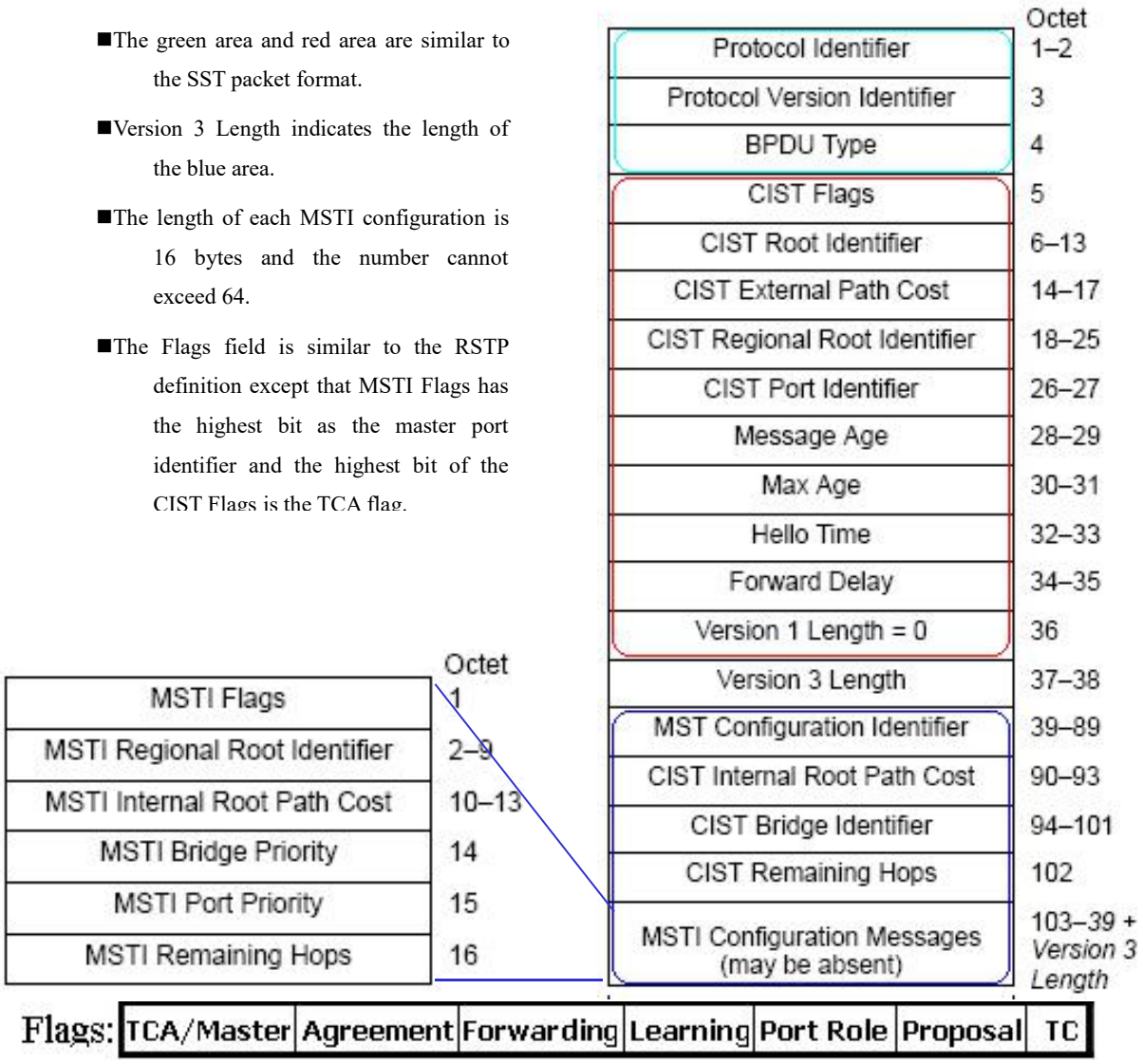
The message form of MST BPDU is shown in the following figure:

■The green area and red area are similar to the SST packet format.

■Version 3 Length indicates the length of the blue area.

■The length of each MSTI configuration is 16 bytes and the number cannot exceed 64.

■The Flags field is similar to the RSTP definition except that MSTI Flags has the highest bit as the master port identifier and the highest bit of the CIST Flags is the TCA flag.



The message form of MST PDU

Protocol identifier: 0 x0000. It identifies the spanning tree protocol (two bytes).

Protocol Version identifier: 0x03. It identifies the protocol version (one byte).

BPDU Type: 0x02. It identifies RST BPDU (one byte).

CIST Flags: it identifies the topology change acknowledgment, agreement, forwarding, learning, port role, proposal, topology change status (1 byte) of the CIST.

CIST Root Identifier: it is the unique identifier of the CIST root bridge, which is composed of the CIST root bridge priority and MAC address (8 bytes).

CIST External Root Path Cost: it changes only in the cross-domain, and propagates constant in the time domain (4 bytes).

CIST Regional Root Identifier： It consists of the CIST root bridge priority and the CIST root bridge's MAC address. It changes only in the cross-domain, and propagates constant in the time domain (8 bytes).

CIST Port Identifier： It is the port ID for sending MST BPDU message. It consists of the port priority and the port ID (2 bytes).

Message Age: the time of MST BPDU message generated from the CIST root bridge. It

changes only in the cross-domain, and propagates constant in the time domain (2 bytes).

Max Age: it is the validity time of the MST BPDU message, and set by the CIST root bridge (2 bytes).

Hello time: it is the validity interval of the MST BPDU message, and set by the CIST root bridge (2 bytes).

Forward Delay: it is set by the CIST root bridge (2 bytes), which has two functions:

It acts as the protocol timer of the port state transition (from Discarding to Learning, from Learning to Forwarding), and the aging time of the dynamic entries of the filtering database when the network topology changes.

Version 1 length: additional information, it is fixed to 0 (1 byte).

Version 3 length: It indicates the length from the MSTP configuration to the end of the BPDU message (2 bytes).

MST Configuration Identifier: it consists of a configuration selector, configuration name, revision level and configuration summary. It changes only in the cross-domain, and propagates constant in the time domain (51 bytes).

CIST Internal Root Path Cost: It is valid only in the MST region (4 bytes).

CIST Bridge Identifier: It is the bridge ID for sending MST BPDU, and consists of bridge priority and bridge MAC address (8 bytes).

CIST Remaining Hops: it is the remaining number of hops of MST BPDU in the CIST (1 byte).

MSTI Flag: It identifies the MSTI's primary port, agree, forward, learn, port role, offer, topology change status (1 byte).

MSTI Regional Root Identifier: it is the only identifier of MSTI regional root bridge. It is composed of the root bridge priority of the MSTI, MSTID, and the MAC address of the root bridge of the MSTI. The root bridges of the MSTI may be different for different MSTIs (8 bytes).

MSTI Internal Root Path Cost: it is valid only in the MST region (4 bytes).

MSTI Bridge Priority: It forms the sending bridge of the MSTI configuration information together with the MAC address in the CIST Bridge Identifier (1 byte).

MSTI Port Priority: It forms the transmission port of the MSTI configuration information together with the port ID in the CIST Port Identifier (1 byte).

MSTI Remaining Hops: It is the remaining hops of MST BPDU in the MSTI (1 byte).

19.2.2 CIST Priority Vector

CIST needs to use the following seven protocol parameters when it performs bridge role and port role calculation, called the CIST priority vector:

1）CIST root id
2）CIST external root path cost
3）CIST regional root id
4）CIST internal root path cost
5）CIST designated bridge id
6）CIST designated port id
7）CIST receiving port id

Between these parameters there is a priority, the higher the priority is, the more forward the position is.

19.2.3 MSTI Priority Vector

MSTI needs to use the following five protocol parameters when it performs bridge role and port role calculation, called the MSTI priority vector:

1）MSTI regional root id

2）MSTI internal root path cost

3）MSTI designated bridge

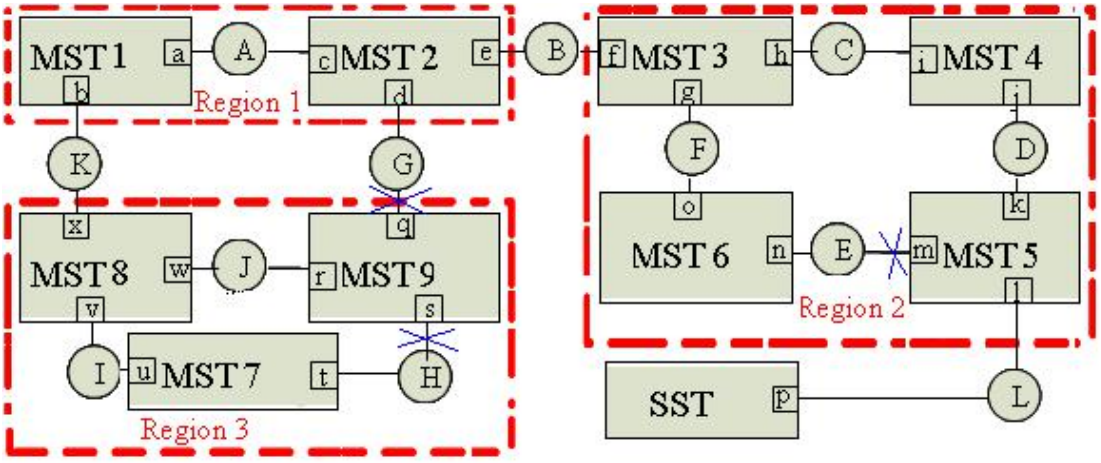
4）MSTI designated port

5）MSTI receiving port

Between these parameters there is a priority, the higher the priority is, the more forward the position is.

19.2.4 MSTP Election Process

The MSTP election process can be divided into two parts. First, the CIST election is performed and then the MSTI election is performed.



As shown in above figure, it is assumed that all ports of all bridges in the network have equal port path costs. The bridge identities of MST Bridge 1 and MST Bridge 9 are incremented at one time and the identity of the SST Bridge is the maximum.

19.2.5 CIST Role Selection

1. Select CIST Root Bridge, CIST root port

In the whole bridged LAN, MST Bridge 1 has the highest bridge identification priority and is elected as the CIST root bridge. Assume that the external root path of Region 2, Region 3 to the CIST root bridge costs 1. The bridge CIST priority vector of the MST bridge 8 is updated to (MST bridge1, 1, MST bridge8, 0, MST bridge8), the port x of MST bridge 8 is the CIST root port; the level vector is updated to (MST Bridge 1, 1, MST Bridge 9, 0, MST Bridge 9). Similarly, port f on MST 3 is the CIST root port.

2. Select the CIST root bridge (IST root bridge) and the CIST root port of each domain

After selecting the CIST root bridge, the CIST root bridge of each area is selected. Take Region 3 for example:

The port u of MST bridge 7 receives the CIST priority vector (MST bridge 1, MST bridge 8, 0, MST bridge 8, v) of the MST bridge 8. Compared with the port u itself (MST bridge 7, 0, MST bridge 7, 0, MST bridge 7, u), MST bridge 8 is better, so the information of port u is updated to (MST bridge 1, 1, MST bridge 8, 0, MST bridge 8, v), and the information of port t is updated to (MST bridge 1, 1, MST bridge 9, 0, MST bridge 9, s). Then MST bridge 7 port u and t again compare CIST priority vector and find that the information of port u is better, so select MST Bridge 8 as the CIST regional root bridge in Region 3. Port u of MST bridge 7 is the CIST root port. Assume that the CIST internal root path cost of MST bridge 7 is 1, and the information of port t is updated to (MST bridge 1, MST bridge 8, 1, MST bridge 7, t).

The port w of MST bridge 8 receives the CIST priority vector (MST bridge 1, 1, MST bridge 9, 0, MST bridge 9, r) of the MST bridge 9. Compared with the port w itself (MST bridge 1, 1, MST bridge 8, 0, MST bridge 8, w), itself is better. The information of port w is not updated. The information of port v receiving the CIST priority vector (MST bridge 7, 0, MST bridge 7, 0, MST bridge 7, u) of the MST bridge 7 is better. The information of port v is not updated. Then select MST Bridge 8 as the CIST regional root bridge in Region 3. Select MST bridge 8 as the CIST root port.

The port r of MST bridge 9 receives the CIST priority vector of MST bridge 8 (MST bridge 1, 1, MST bridge 8, 0, MST bridge 8, w), compared with port r itself (MST bridge 1, 1, MST bridge 9, 0, MST bridge 9, r), MST bridge 8 is better. The information of port r is updated to (MST bridge 1, 1, MST bridge 8, 0, MST bridge 8, w); the information of port s is better than received the CIST priority vector of MST bridge 7 (MST bridge 7, 0, MST bridge 7,0, MST bridge 7, u), the information of the port s is not updated. Compared CIST priority vector of the port r and s of MST bridge 9, for comparison, select MST Bridge 8 for CIST regional root bridge of Region 3, the port C of MST bridge 9 is the CIST root port. Assume that the CIST internal root path cost of MST bridge 9 is 1, and the information of port s is updated to (MST bridge 1, 1, MST bridge 8, 1, MST bridge 9, s).

MST bridge 3 is elected as the CIST root bridge in region 2. MST port 4 is the CIST root port. The port k of MST bridge 5 is the CIST root port. The port o of MST bridge 6 is the CIST root port.

Because the MST bridge 1 is CIST root bridge, MST bridge 1 is also the CIST region root bridge of region 1, the port c of MST bridge 2 is the CIST root port.

3. Elect IST designated bridge in each domain and the CIST designated port

After selecting the CIST regional root bridge, take Region 3 as an example:

MST bridge 8 is the CIST region root bridge. So both port w and port v are designated port. They are also the designated bridges of LAN I and LAN J.

MST bridge 9 receives the message priority vector at port s (MST bridge 1, 1, MST bridge 8, 1, MST bridge 7, t) better than MST bridge 9's own port priority vector (MST bridge 1, 1, MST bridge 8, 1, MST bridge 9, s). That is, the received CIST root bridge, CIST external root path cost, CIST regional root bridge and CIST internal root path cost are all equal, but the CIST designated bridge identifier is smaller than that of the CIST root bridge. Therefore, the MST Bridge 7 is CIST designated bridge of the local area network H. The port t of the MST bridge 7 becomes the designated port of the CIST. The port s of the MST bridge 9 is the replacement port and set to the Discarding state. Similarly, the port d of MST bridge 2 is designated port, MST bridge 2 is the

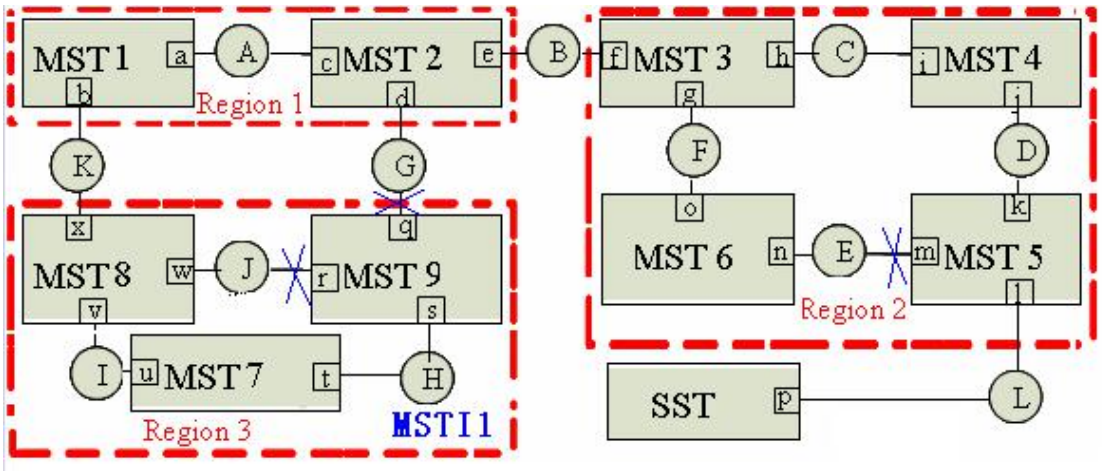
designated bridge of LAN G, port q of MST bridge 9 is the replacement port, and set to Discarding state.

Similarly, in Region 2, port j of MST bridge 4 is the CIST designated port , MST Bridge 4 is designated bridge of the LAN D; port n of MST bridge 6 is the CIST designated port, MST Bridge 6 is the local area network E designated bridge.

In Region 1, MST bridge 1 is the CIST region root bridge. Therefore, port a and port b are designated ports and the designated bridge of local area network A. Port e of MST bridge 2 is the designated port, and MST bridge 2 is local area network B designated bridge.

19.2.6 MSTI Role Selection

The MSTI election is similar to the single-spanning tree election process, which is performed by comparing and selecting MSTI priority vectors.



Take Region 3 as an example to introduce the formation process of MST1, as shown in above figure:

Assume bridge priority: MST bridge 9 <MST bridge 8 <MST bridge 7, the path cost of all ports is 1.

1. Select MSTI region root bridge

The bridge priority of MST bridge 7 is the highest, so it is elected as MSTI region root bridge.

2. Select MSTI root port of no-root bridge

MST bridge 8: Select port v as MSTI port. MSTI internal root path cost is 1.

MST bridge 9: Select port s as MSTI port. MSTI internal root path cost is 1.

3. Select MSTI designated port of each bridge

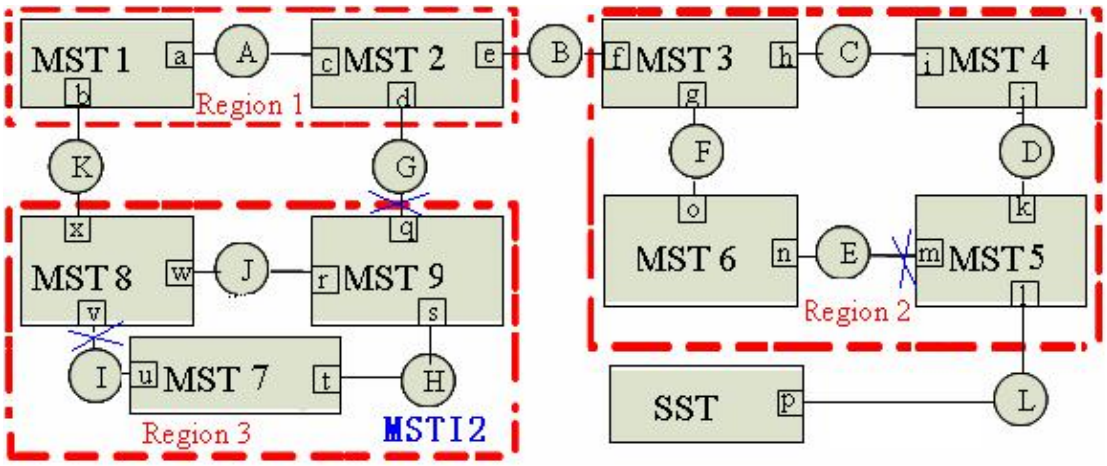
MST bridge 7: select designated bridge of LAN I and LAN H. The port u and port t are MSTI designated port.

MST bridge 8: select designated bridge of LAN J. The port w is MSTI designated port. Port x is the port where Region 3 communicates with the upstream port and is the MSTI master port of MSTI1.

MST bridge 9: Port r is the MSTI replacement port of port s; port q is the CIST replacement port of Region 3, and is designated as the MSTI replacement port of MSTI1.

LAN J Select MSTI to specify the process for the bridge and MSTI designated ports: MST

bridge 9 at port r received priority vector (MST bridge 7, 1, MST bridge 8, w) is superior to the port priority vector (MST bridge 7, 1, MST bridge 9, r) of the MST bridge 9. That is, the received MSTI region root bridge and the MSTI internal root path cost are all equal, but the MSTI designated bridge identity is smaller than itself. So select MST bridge 8 as MSTI designated bridge of LAN J, port w is MSTI designated port of LAN J, port r is configured for Discarding state.



Take the Region 3 as an example to introduce the formation process of MSTI2, as shown in above figure:

Assume bridge priority: MST bridge 8< MST bridge 7< MST bridge 9, the path cost of all ports is 1.

1. The bridge priority of the MST bridge 9 is the highest, which is elected for MSTI region root bridge.

2. The MSTI internal path cost of the MST bridge 7 and the MST bridge 8 is 1. The port t and w are the MSTI root port.

3. MST bridge 9 is designated bridge of LAN J and LAN H. The port r and s are MSTI designated port; MST bridge 7 is elected for the designated bridge of LAN I, the port u is the MSTI designated port; the port v of MST bridge 8 is elected MSTI replace port of port w.

4. the port x is the port of region 3 and upstream communication, and designed as MSTI master port of the MSTI2; the port q is the replacement port of region 3, and is specified as the MSTI replacement port of the MSTI2.

From the above can be seen: The port role in the MSTI is bounded by the CIST, if the CIST port role is CIST root port (root port of IST root bridge), it is the master port for all MSTIs; if the CIST port role is the CIST replacement port of the master port, it is the replacement port of all MSTIs. The same port may have different port states for different MSTIs (for example the port v is Forwarding state in the MSTI1, but discarding state in the MSTI2),

In addition, the setting of bridge priority, port priority and port path cost is unaffected for different MSTIs (for example, MSTI1 and MSTI2 can be configured with their own parameter values).

19.2.7 Topology Stable State

Run MSTP switch, the calculation and comparison operations are performed according to the

received BPDU, the network can achieve the following stable state:

(1) An switch is elected for the CIST root of the whole network.

(2) Each switch and LAN segment will determine the minimum cost of the CIST root path, to ensure the integrity of the connection and to prevent loop.

(3) Within each region, a switch is elected as CIST Regional Root, which has the minimum cost of the CIST root path.

(4) Each MSTI independently selects a switch as the MSTI region root.

(5) Each switch and LAN segment in the region can determine the minimum cost of MSTI root path.

(6) The CIST root port provides the minimum cost path through which the CIST region root (If the switch is not the CIST region root)reaches the CIST root (If the switch is not the CIST root).

(7) Alternate and Backup ports provide connectivity when a switch, port, or LAN fails or is removed.

(8) The MSTI root port provides the minimum cost path to the MSTI region root(if the switch is not in the MSTI root bridge).

(9) A master port provides the connection between the region and the out of the CIST region root bridge. In the region, the CIST root port of the CIST region root bridge can be master port of the MSTI.

19.2.8 Topology Change

The STP topology change propagation is similar to the RSTP.

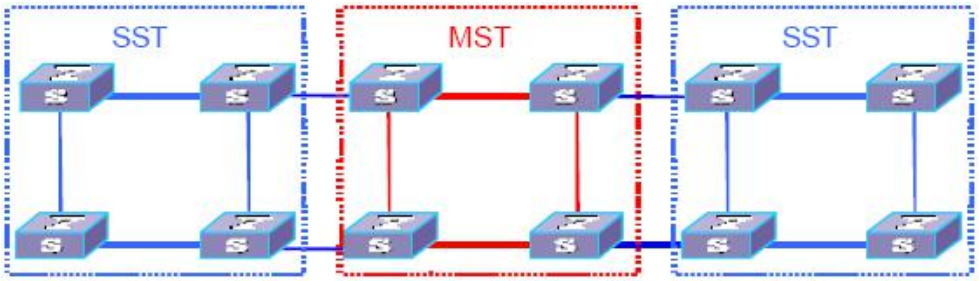
In the MSTP, only one situation is considered to be a topology change, that is, when a port is from the inactive port into an active port, which is considered a topology change. The port role of the backup port or backup port switch to the root port , designated port or the master port.

In addition, MSTP supports "offer / agree" mechanism and point-to-point link type the same as RSTP, which is used to quickly switch port state to forwarding state.

19.2.9 MST and SST compatibility

The STP protocol divides the supportive MSTP switch and the unsupportive MSTP switch into different regions, they are called as the MST region and the SST region respectively. Run the multi-instance spanning tree in the MST region and run the RSTP-compatible protocol at the edge of the MST region.

The following figure shows the working principle of STP:



The switch in the red MST region uses the MSTP BPDU to exchange topology information. The switch in the blue SST domain exchanges topology information with STP / RSTP BPDU.

The MSTP processing on the edge port between the MST region and the SST region is slightly more complicated

When the edge port receives STP BPDU from other switches, the port enters STP-compatible state and sends STP BPDU.

When an edge port receives an RSTP BPDU, the port enters the RSTP-compatible state, but still sends the MSTP BPDU.

Because RSTP is designed with expansion in mind, the opposite RSTP device can interpret MSTP messages as correct RSTP messages.

19.3 Configure MSTP

19.3.1 Start MSTP

After the global tree is automatically generated, all the ports take part in the calculation of the spanning tree topology. If the administrator wishes to exclude certain ports from the calculation of the spanning tree, use the no spanning-tree command to disable the spanning tree function in the configuration mode of the specified port.

- Start MSTP

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Select the spanning tree mode	**spanning-tree mode mstp**	required
Star the spanning tree globally	**spanning-tree**	required
Shut down the spanning tree globally	**no spanning-tree**	
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Port opens the spanning tree	**spanning-tree**	optional
Port closes the spanning tree	**no spanning-tree**	

19.3.2 Configure the MSTP Timer Parameter Value

MSTP timer includes forwarding delay, hello time, maximum age, and max hops. User can configure these four parameters to calculate spanning trees for MSTP on the switch.

- Configure the MSTP Timer Parameter Value

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure bridge forward delay	**spanning-tree mst forward-time** *forward-time*	optional
Configure bridge hello time	**spanning-tree mst hello-time** *hello-time*	optional
Configure bridge max age	**spanning-tree mst max-age** *max-age*	optional
Configure bridge max hops	**spanning-tree mst max-hops** *max-hops*	optional

Note:

An excessively long Hello Time value causes that the bridge considers a link failure and starts

to recalculate the spanning because of packet loss; a too short hello time value causes the bridge to send configuration information frequently, increasing the network and CPU load. Hello Time is in the range of 1 to 10 seconds. It is recommended to use the default value 2 seconds. The hello time must be less than or equal to that of forward delay-2.

If the forward delay is too small, a temporary redundant path may be introduced. If the forward delay is too large, the network may not resume communication for a long time. The value of forward delay is in the range of 4 to 30 seconds. It is recommended to use the default value 15 seconds. The time of forward delay must be greater than or equal to hello time + 2.

Max Age sets the maximum time interval for MSTP protocol message aging. If the timer expires, the packet is discarded. If the value is too small, the calculation of the spanning tree may be frequent. It is possible to misinterpret network congestion as a link failure. If the value is too large, it is not conducive to timely detection of link failure. Max Age is in the range of 6 to 40 seconds. The Max Age time value depends on the network diameter of the switched network. The default value of 20 seconds is recommended. Max Age must be greater than or equal to 2 * (Hello Time + 1) and less than or equal to 2 * (Forward Delay-1).

19.3.3 Configure the MSTP Configuration Identifier

The MSTP configuration identifier includes the MSTP configuration name, the MSTP revision level, and the mapping between MSTP instance and VLAN. MSTP treats the bridge with the same configuration identifier and interconnection each other logically as a virtual bridge.

- Configure the MSTP Configuration Identifier

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure the MSTP configuration name	**spanning-tree mst name** *name*	optional
Configure the MSTP revision level	**spanning-tree mst revision** *revision-level*	optional
Configure the mapping between MSTP instance and VLAN	**spanning-tree mst instance** *instance-num***vlan***vlan-list*	optional
Delete the mapping between MSTP instance and VLAN	**no spanning-tree mst instance** *instance-num***vlan***vlan-list*	

19.3.4 Configure MSTP Bridge Priority

In MSTP, the bridge priority is based on the parameters of each spanning tree. The bridge priority, along with the port priority and port path cost, determine the topology of each spanning tree instance, and form the basis for link load balancing.

The size of the switch bridge priority determines whether the switch can be selected as the root bridge of the spanning tree. By configuring a smaller bridge priority, user can specify that a switch becomes the root bridge of the spanning tree.

By default, the bridge priority of the switch is 32768.

- Configure MSTP Bridge Priority

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Configure the bridge priority in the MSTP instance	**spanning-tree mst instance** *instance-num***priority** *priority*	optional

19.3.5 Configure the Boundary Port Status of a Port

Border port refers to the port connecting to the host. These ports can enter the forwarding state within a short time after the linkup, but they will automatically switch to the non-edge port once they receive the spanning tree message.

- Configure the port as a boundary port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure the port as a boundary port	**spanning-tree mst portfast**	optional
Configure the port as a non-boundary port	**no spanning-tree mst portfast**	

19.3.6 Configure the Link Type of a Port

There are two link types: one is the shared media link type (connected via hub etc.), and the other is the point-to-point link type. The link type is mainly used in the proposal of rapid transition of port state--consent mechanism. Only the port with the link type of point-to-point can allow the fast transition of the port state.

In MSTP, the port quickly enter the forwarding state, which requires that the port must be a point-to-point link, not a shared media link, user can manually specify the link type of the port, or automatically determine the current link type of the port(Full-duplex port is automatically judged as point-to-point link, half-duplex port automatically determine non-point-to-point link).

- Configure the link type for the MSTP port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure the port link type	**spanning-tree mst link-type point-to-point {auto | forcetrue| forcefalse}**	optional

19.3.7 Configure the Path Cost of Port

The path cost of the port is divided into internal and external costs. The former is based on the configuration parameters of each MSTP instance. It is used to determine the topology of different instances in each MSTP region. The latter is an instance-independent parameter that determines the topology of the CST composed of regions.

By configuring the path cost of a port, user can make the port more easily to be a root port or a designated port.

The path cost of a port depends on the link rate of the port. The larger the link rate is, the smaller the parameter configuration is. The MSTP protocol automatically detects the link rate of the current port and translates it into the corresponding path cost.

Configuring the path cost of an Ethernet port will cause the spanning tree to recalculate. The cost of a port path ranges from 1 to 65,535. It is recommended to use the default value. Let the MSTP protocol calculate the path cost of the current port. By default, the path cost is determined based on the current rate of the port.

The default port path cost is based on the port speed. The default is 200,000 when the port speed is 10M. The default value is 200,000 when the port is 100M. When the port rate is not available, the path cost is 200,000 by default.

- Configure the path cost of the port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	optional
Configure the internal path cost for the port	**spanning-tree mst instance** *instance-num* **cost** *cost*	optional
Configure the external path cost of the port	**spanning-tree mst external cost** *cost*	optional

19.3.8 Configure Port Priority

In MSTP, port priority is based on the parameters of each spanning tree. By configuring the priority of a port, user can make a port more easily to be a root port.

The lower the priority value is, the higher the priority is. Changing the priority of an Ethernet port will cause the spanning tree to recalculate. The spanning tree priority of a port ranges from 0 to 240. It must be an integer multiple of 16. By default, the port spanning tree priority is 128.

- Configure Port Priority

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure Port Priority	**spanning-tree mst instance** *instance-num***port-priority** *priority*	optional

19.3.9 Configure Root Protection of a Port

Because of configuration error or malicious attack on the network, the legal root bridge of the network may receive a configuration message of a higher priority, so the current root bridge will lose the status of the root bridge, and cause the wrong changes of the network topology. Assume that the original traffic is forwarded through the high-speed link. This illegal change will lead to the traffic that is towed to the low-speed link via the original high-speed link, and network congestion. Root protection can prevent it from happening.

For a port with root protection enabled, the port role can only be the designated port. Once a high priority configuration information has been received by the port, the status of these ports will be set for Discarding and not be forwarded (The link connected to this port is disconnected). When the configuration information is not received within a sufficient period of time, the port returns the original state.

In MSTP, this function works on all instances.

- Configure the root guard function of the port

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Enable the root guard function of the port	**spanning-tree mst root-guard**	optional
Disable the root guard function of the port	**no spanning-tree mst root-guard**	

19.3.10 Configure the Digest Snooping Function

When the port of the switch connects with the switch of Cisco etc using a private spanning tree, because these vendors' switches are configured with spanning tree-related private protocols, even if the MST regions are configured the same, the switch cannot communicate with each other in the MST region. The digest snooping feature prevents this from happening. After the digest snooping function is enabled on a port connected with vendors' switch using spanning tree-related proprietary protocols, when receiving these BPDU from the vendor' switch, the switch considers them as messages from the same MST region and records the configuration BPDU. When BPDU messages are sent to these vendors' switches, the switch adds the configuration digests. In this way, the switch implements the interworking between these switches in the MST region.

- Configure the Digest Snooping Function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Enable the digest snooping function	**spanning-tree mst config-digest-snooping**	optional
Disenable the digest snooping function	**no spanning-tree mst config-digest-snooping**	

19.3.11 Configure Loop-guard Function

The loop-guard function: prevent a blocked port because of abnormal link from becoming a forwarding state after not receiving the BPDU configuration information. When the port is configured with this option, the port remains blocked even if BPDU configuration BPDU are not received.

- Configure loop-guard function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	
Enable the loop-guard function	**spanning-tree mst loop-guard**	Don't share with root-guard
disenable the loop-guard function	**no spanning-tree loop-guard**	

19.3.12 Configure BPDU Guard Function

For an access layer device, the access port usually directly connects to the user terminal or the file server. In this case, the access ports are set as edge ports to enable rapid transition of these ports. When these ports receive BPDU message, the system automatically sets these ports as non-edge ports and recalculates the spanning trees, and causes network topology changes. These ports should normally not receive BPDU message. If someone forges BPDU to attack the device, the network will become unstable.

The device provides the BPDU guard function to prevent these attacks: After the BPDU guard function is enabled on a device, if a port configured with edge port attributes receives BPDU

message, the device will SHUTDOWN the port and prompt the user with the syslog information.

- Configure bpdu-guard function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enable the bpdu guard function globally	(no) **spanning-tree mst bpdu-guard**	In global mode, this command is enabled on all ports
Disable the bpdu guard function	**no spanning-tree mst bpdu-guard**	
Enter the port configuration mode	**interface ethernet** *interface-num*	
Enable the bpdu guard function	**spanning-tree mst bpdu-guard**	Take effect on the specified port
Disable the bpdu guard function	**no spanning-tree mst bpdu-guard**	

Note: The BPDU guard function of a port takes effect only on the port configured with the edge port attribute. If the edge port attribute is configured, but because it receives the BPDU message from other port and re-becomes non-edge port. In this case, the BPDU guard function is enabled. The port can take effect only when it is restarted as an edge port.

19.3.13 Configure Bpdu-filter Function

If the BPDU is set on the edge port, the device will discard the received BPDU message and the port will not send BPDU message.

- Configure bpdu-filter function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enable bpdu-filter function	**spanning-tree mst bpdu-filter**	In global mode, this command is enabled on all ports
Disable bpdu-filter function	**no spanning-tree mst bpdu-filter**	
Enter the port configuration mode	**interface ethernet interface-num**	
Enable bpdu-filter function	**spanning-tree mst bpdu-filter**	Take effect on the specified port
Disable bpdu-filter function	**no spanning-tree mst bpdu-filter**	

19.3.14 Configure Mcheck Function

The switch running in MSTP mode can be connected to STP switch to ensure compatibility. However, after the neighbor changes the working mode for MSTP, the two connected ports still work in STP mode by default. The Mcheck function is used to force a port to send MSTP packets and determine whether the adjacent port can work in MSTP mode. If yes, the switch works in MSTP mode automatically.

- Configure mcheck function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** *interface-num*	-
Configure mcheck function	**spanning-tree mst mcheck**	optional

Note: The Mcheck function requires that the port send BPDU message only on the specified port.

19.3.15 Enable / Disable MSTP Instance

In order to control MSTP flexibly, user can enable the DISABLE feature of INSTANCE. The effect of disable instance is similar to that of executing no spanning-tree in stp mode. The port mapped to the vlan of the instance is forwarding on all the connections.

- Enable / Disable MSTP Instance

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Ignore publishing an MSTP instance	**spanning-tree mst disable instance** *instance-number*	optional
Restore publishing an MSTP instance	**no spanning-tree mst disable instance** *instance-number*	optional

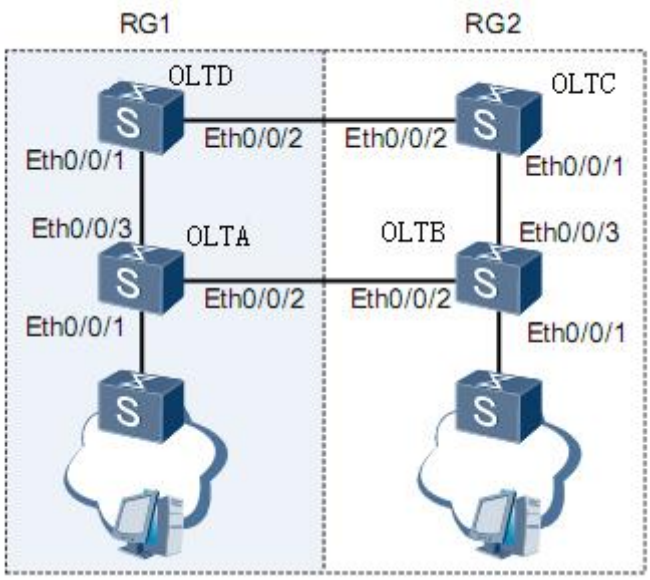
19.3.16 MSTP Display and Maintenance

After you complete the above configuration, you can use the following command to view the configuration.

- MSTP Display and Maintenance

operation	command	remark
Display MSTP configuration identifier information	**show spanning-tree mst config-id**	All modes are executable
Show all MSTI port information	**show spanning-tree mst instance brief** *id*	
Display a port MSTI information	**show spanning-tree mst instance** *id* **interface ethernet** *interface-list*	
Show posting examples ignored	**show spanning-tree mst disabled-instance**	

19.3.17 MSTP Configuration Example



1. Networking Requirements

In the network shown in above figure, s-switch A and S-switch C are configured in a domain with domain name RG1 and MSTI 1 is created. Configure S-switch B and S-switch D to another domain. The domain name is RG2, and MSTI1 and MSTI2 are created.

Configure S-switch C as the CIST root. In the RG1domain, S-switch C is the CIST domain root, and S-switch C is the domain root of MSTI1. Apply the root guard function to S-switch C ports Ethernet0/0/1 and Ethernet0/0/2. In the RG2 domain, S-switch D is the CIST regional root, S-switch B is the regional root of MSTI1, and S-switch D is the regional root of MSTI2.

The L2 switch connected to S-switch A and S-switch B does not support MSTP. Set Ethernet0/0/1 as the edge interface for switch A and switch B

2, Configuration steps

Configure S-switch C

Create VLAN from 1 to 20 and configure ports Ethernet 0/0/1 and Ethernet 0/0/2 as trunk ports and add them to VLAN 1 to 20

```
S-switch-C (config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-C (config-if-range)#switchport mode trunk
S-switch-C (config-if-range)#exit
S-switch-C (config)#vlan 1-20
S-switch-C (config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/2
S-switch-C (config-if-vlan)#exit
```

Configure the MST region of S-switch-C.

```
S-switch-C (config)#spanning-tree mst name RG1
S-switch-C (config)#spanning-tree mst instance 1 vlan 1-10
```

Set the priority of S-switch-C in MSTI 0 to 0, ensuring S-switch-C as the common root of the CIST.

```
S-switch-C(config)#spanning-tree mst instance 0 priority 0
```

Set the priority of S-switch-C in MSTI 1 to 0, ensuring S-switch-C as the common root of the MSTI1.

```
S-switch-C(config)#spanning-tree mst instance 1 priority 0
```

Enable root guard on port Ethernet0/0/1 and Ethernet0/0/2.

```
S-switch-C(config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-C(config-if-range)#spanning-tree mst root-guard
S-switch-C(config-if-range)#exit
```

Enable MSTP.

```
S-switch-C(config)#spanning-tree mode mstp
S-switch-C(config)#spanning-tree
```

Configure **S-switch-A**

Create VLAN from 1 to 20 and configure ports Ethernet 0/0/2 and Ethernet 0/0/3 as trunk ports and add them to VLAN 1 to 20.

```
S-switch-A (config)#interface range ethernet 0/0/2 ethernet 0/0/3
S-switch-A (config-if-range)#switchport mode trunk
S-switch-A (config-if-range)#exit
S-switch-A (config)#vlan 1-20
S-switch-A (config-if-vlan)#switchport ethernet 0/0/2 ethernet 0/0/3
S-switch-A (config-if-vlan)#exit
```

Configure the MST region of S-switch-A.

```
S-switch-A (config)#spanning-tree mst name RG1
S-switch-A (config)#spanning-tree mst instance 1 vlan 1-10
```

Configure Ethernet0/0/1 as an edge interface.

```
S-switch-A(config)#interface ethernet 0/0/1
S-switch-A(config-if-ethernet-0/0/1)#spanning-tree mst portfast
S-switch-A(config-if-ethernet-0/0/1)#exit
```

Enable MSTP.

```
S-switch-A(config)#spanning-tree mode mstp
S-switch-A(config)#spanning-tree
```

Configure **S-switch-D**

Create VLAN from 1 to 20 and configure ports Ethernet 0/0/1 and Ethernet 0/0/2 as trunk ports and add them to VLAN 1 to 20.

```
S-switch-D (config)#interface range ethernet 0/0/1 ethernet 0/0/2
S-switch-D (config-if-range)#switchport mode trunk
S-switch-D (config-if-range)#exit
S-switch-D (config)#vlan 1-20
S-switch-D (config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/2
S-switch-D (config-if-vlan)#exit
```

Configure the MST region of S-switch-D

```
S-switch-D(config)#spanning-tree mst name RG2
S-switch-D(config)#spanning-tree mst instance 1 vlan 1-10
S-switch-D(config)#spanning-tree mst instance 2 vlan 11-20
```

Set the priority of S-switch-D in MSTI0 to 4096. Ensure the S-switch-D as the CIST region root of RG2.

```
S-switch-D(config)#spanning-tree mst instance 0 priority 4096
```

Set the priority of S-switch-D in MSTI 2 to 0. Ensure the S-switch-D as the MSTI 2 region root.

```
S-switch-D(config)#spanning-tree mst instance 2 priority 0
```

Enable MSTP.

```
S-switch-D(config)#spanning-tree mode mstp
S-switch-D(config)#spanning-tree
```

Configure S-switch-B

Create VLAN from 1 to 20 and configure ports Ethernet 0/0/2 and Ethernet 0/0/3 as trunk ports and add them to VLAN 1 to 20.

```
S-switch-B (config)#interface range ethernet 0/0/2 ethernet 0/0/3
S-switch-B (config-if-range)#switchport mode trunk
S-switch-B (config-if-range)#exit
S-switch-B (config)#vlan 1-20
S-switch-B (config-if-vlan)#switchport ethernet 0/0/2 ethernet 0/0/3
S-switch-B (config-if-vlan)#exit
```

Configure the MST region of S-switch-B.
S-switch-B(config)#spanning-tree mst name RG2
S-switch-B(config)#spanning-tree mst instance 1 vlan 1-10
S-switch-B(config)#spanning-tree mst instance 2 vlan 11-20

Set the priority of S-switch-B in MSTI 1 to 0, ensure that S-switch-B as the MSTI domain root.
S-switch-B(config)#spanning-tree mst instance 1 priority 0

Configure Ethernet 0/0/1 as an edge interface.
S-switch-B(config)#interface ethernet 0/0/1
S-switch-B(config-if-ethernet-0/0/1)#spanning-tree mst portfast
S-switch-B(config-if-ethernet-0/0/1)#exit

Enable MSTP.
S-switch-B(config)#spanning-tree mode mstp
S-switch-B(config)#spanning-tree

Verify the configuration

Run the display command on the S-switch-C to view the election result and port status of multiple spanning trees. The results are as follows:

S-switch-C(config)#show spanning-tree mst instance 0 interface ethernet 0/0/1 ethernet 0/0/2
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 0-000a.5a13.f48e
Cist root is 0-000a.5a13.f48e,root port is
Region root is 0-000a.5a13.f48e,root port is
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 20
External root path cost is 0,internal root path cost is 0

Port e0/0/1 of instance 0 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard enable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 20
Received BPDUs:TCN 0,RST 18,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 137,Config BPDU 0

Port e0/0/2 of instance 0 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard enable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 20

Received BPDUs:TCN 0,RST 85,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 86,Config BPDU 0

S-switch-C(config)#show spanning-tree mst instance 1 interface ethernet 0/0/1 ethernet 0/0/2
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 0-000a.5a13.f48e
Cist root is 0-000a.5a13.f48e,root port is
Region root is 0-000a.5a13.f48e,root port is
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 20
External root path cost is 0,internal root path cost is 0

Port e0/0/1 of instance 1 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard enable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 18,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 182,Config BPDU 0

Port e0/0/2 of instance 1 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard enable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 130,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 131,Config BPDU 0

Because S-switch-C has the highest intra-CIST priority, S-switch-C is selected as the CIST common root, and it is also the regional root of RG1. The port Ethernet0/0/1 and port Ethernet0/0/2 of S-switch-C are designated ports on the CIST. They are in the forwarding state.

S-switch-C has the highest priority on MSTI1 in RG1domain, so S-switch-C is selected as the domain root of MSTI1. Ethernet0/0/1 and Ethernet0/0/2 are calculated as the designated ports on MSTI1. They are in the forwarding state.

Run the display command on the device S-switch-A to view the election result and port status of multiple spanning trees. The results are as follows:

S-switch-A(config)#show spanning-tree mst instance 0 interface ethernet 0/0/1 ethernet 0/0/2 ethernet 0/0/3
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 32768-000a.5a13.b13d
Cist root is 0-000a.5a13.f48e,root port is e0/0/3
Region root is 0-000a.5a13.f48e,root port is e0/0/3
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20

Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External root path cost is 0,internal root path cost is 200000

Port e0/0/1 of instance 0 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-000a.5a13.b13d,designated port is e0/0/1
Port is a(n) edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 19
Received BPDUs:TCN 0,RST 0,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 249,Config BPDU 0

Port e0/0/2 of instance 0 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-000a.5a13.b13d,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 19
Received BPDUs:TCN 0,RST 30,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 279,Config BPDU 0

Port e0/0/3 of instance 0 is forwarding
Port role is RootPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 19
Received BPDUs:TCN 0,RST 313,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 18,Config BPDU 0

S-switch-A(config)#show spanning-tree mst instance 1 interface ethernet 0/0/1 ethernet 0/0/2 ethernet 0/0/3
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 32768-000a.5a13.b13d
Cist root is 0-000a.5a13.f48e,root port is e0/0/3
Region root is 0-000a.5a13.f48e,root port is e0/0/3
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External root path cost is 0,internal root path cost is 200000

Port e0/0/1 of instance 1 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-000a.5a13.b13d,designated port is e0/0/1
Port is a(n) edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 0,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 273,Config BPDU 0

Port e0/0/2 of instance 1 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-000a.5a13.b13d,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 30,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 303,Config BPDU 0

Port e0/0/3 of instance 1 is forwarding
Port role is RootPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 337,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 18,Config BPDU 0

Ethernet0/0/3 of S-switch-A is the root port in CIST and MSTI1. Ethernet0/0/2 is the designated port in CIST and MSTI1. They are in the forwarding state. The Ethernet0/0/1 is the edge port and in the forwarding state.

Run the display command on S-switch-D to view the election result and port status of multiple spanning trees. The results are as follows:

S-switch-D(config)#show spanning-tree mst instance 0 interface ethernet 0/0/1 ethernet 0/0/2
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 4096-0000.0077.8899
Cist root is 0-000a.5a13.f48e,root port is e0/0/2
Region root is 4096-0000.0077.8899,root port is e0/0/2
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 20
External root path cost is 200000,internal root path cost is 0

Port e0/0/1 of instance 0 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 4096-0000.0077.8899,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 20
Received BPDUs:TCN 0,RST 663,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 58,Config BPDU 0

Port e0/0/2 of instance 0 is forwarding
Port role is RootPort, priority is 128
Port external path cost is 200000,internal path cost is 200000

Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-000a.5a13.f48e,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 20
Received BPDUs:TCN 0,RST 652,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 655,Config BPDU 0

S-switch-D(config)#show spanning-tree mst instance 1 interface ethernet 0/0/1 ethernet 0/0/2
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 32768-0000.0077.8899
Cist root is 0-000a.5a13.f48e,root port is e0/0/2
Region root is 0-0000.0a0a.0001,root port is e0/0/1
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 20
External root path cost is 200000,internal root path cost is 200000

Port e0/0/1 of instance 1 is forwarding
Port role is RootPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-0000.0a0a.0001,designated port is e0/0/3
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 973,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 370,Config BPDU 0

Port e0/0/2 of instance 1 is forwarding
Port role is MasterPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-0000.0077.8899,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 962,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 655,Config BPDU 0

S-switch-D(config)#show spanning-tree mst instance 2 interface ethernet 0/0/1 ethernet 0/0/2
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 0-0000.0077.8899
Cist root is 0-000a.5a13.f48e,root port is e0/0/2
Region root is 0-0000.0077.8899,root port is
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 20
External root path cost is 200000,internal root path cost is 0

Port e0/0/1 of instance 2 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state

Designated bridge is 0-0000.0077.8899,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 1003,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 401,Config BPDU 0

Port e0/0/2 of instance 2 is forwarding
Port role is MasterPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-0000.0077.8899,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 992,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 655,Config BPDU 0

The S-switch-D has a lower priority on the CIST than the S-switch-C, and Ethernet0/0/2 of the S-switch-D is calculated as the root port in the CIST. At the same time, because the S-switch-C and the S-switch-D do not belong to the same domain, Ethernet0/2 of the S-switch-D is calculated as the master port on the MSTI1 and the MSTI2. The S-switch-D takes precedence over the S-switch-B in the CIST, and Ethernet0/0/1 of the S-switch-D is calculated as the designated port in the CIST.

In the MSTI1, the S-switch-D has a lower priority than the S-switch-B. The S-switch-B is elected as the MSTI1 regional root. Therefore, Ethernet0/0/1 of the S-switch-D is calculated as the root port.

In the MSTI2, the S-switch-D takes precedence over the S-switch-B. The S-switch-D is elected as the MSTI2 regional root. Therefore, Ethernet0/0/1 of the S-switch-D is calculated as the designated port.

Run the display command on the S-switch-B to view the election result and port status of multiple spanning trees. The results are as follows:

S-switch-B(config)#show spanning-tree mst instance 0 interface ethernet 0/0/1 ethernet 0/0/2 ethernet 0/0/3
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 32768-0000.0a0a.0001
Cist root is 0-000a.5a13.f48e,root port is e0/0/3
Region root is 4096-0000.0077.8899,root port is e0/0/3
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External root path cost is 200000,internal root path cost is 200000

Port e0/0/1 of instance 0 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-0000.0a0a.0001,designated port is e0/0/1
Port is a(n) edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 19
Received BPDUs:TCN 0,RST 0,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 24,Config BPDU 0

Port e0/0/2 of instance 0 is discarding
Port role is AlternatedPort, priority is 128

Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-000a.5a13.b13d,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 20
Received BPDUs:TCN 0,RST 770,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 7,Config BPDU 0

Port e0/0/3 of instance 0 is forwarding
Port role is RootPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 4096-0000.0077.8899,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:HelloTime 2,MaxAge 20,FwdDelay 15,MsgAge 0,RemainingHops 19
Received BPDUs:TCN 0,RST 783,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 773,Config BPDU 0

S-switch-B(config)#show spanning-tree mst instance 1 interface ethernet 0/0/1 ethernet 0/0/2 ethernet 0/0/3
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 0-0000.0a0a.0001
Cist root is 0-000a.5a13.f48e,root port is e0/0/3
Region root is 0-0000.0a0a.0001,root port is
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External root path cost is 200000,internal root path cost is 0

Port e0/0/1 of instance 1 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-0000.0a0a.0001,designated port is e0/0/1
Port is a(n) edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 0,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 96,Config BPDU 0

Port e0/0/2 of instance 1 is discarding
Port role is AlternatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-0000.0a0a.0001,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 842,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 7,Config BPDU 0

Port e0/0/3 of instance 1 is forwarding
Port role is DesignatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state

Designated bridge is 0-0000.0a0a.0001,designated port is e0/0/3
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 20
Received BPDUs:TCN 0,RST 855,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 846,Config BPDU 0

S-switch-B(config)#show spanning-tree mst instance 2 interface ethernet 0/0/1 ethernet 0/0/2 ethernet 0/0/3
Current spanning tree protocol is MSTP
Spanning tree protocol is enable
Bridge id is 32768-0000.0a0a.0001
Cist root is 0-000a.5a13.f48e,root port is e0/0/3
Region root is 0-0000.0077.8899,root port is e0/0/3
Bridge time:HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time:HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External root path cost is 200000,internal root path cost is 200000
Port e0/0/1 is not a member of instance 2

Port e0/0/2 of instance 2 is discarding
Port role is AlternatedPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 32768-0000.0a0a.0001,designated port is e0/0/2
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 858,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 7,Config BPDU 0

Port e0/0/3 of instance 2 is forwarding
Port role is RootPort, priority is 128
Port external path cost is 200000,internal path cost is 200000
Root guard disable and port is not in root-inconsistent state
Designated bridge is 0-0000.0077.8899,designated port is e0/0/1
Port is a(n) non-edge port,link type is point-to-point
Port time:RemainingHops 19
Received BPDUs:TCN 0,RST 871,Config BPDU 0
Transmitted BPDUs:TCN 0,RST 861,Config BPDU 0

The S-switch-D has a higher priority in the CIST than the S-switch-B. Ethernet0/0/2 of the S-switch-B is calculated as an alternate port. Because the S-switch-B is not in the same domain as S-switch-A, the port Ethernet0/0/2 of the S-switch-B is calculated as alternate ports on the MSTI1 and MSTI2.

In the MSTI1, the priority of the S-switch-D is lower than that of the S-switch-B. The S-switch-B is elected as the root of the MSTI1. Therefore, Ethernet0/0/3 of the S-switch-B is calculated as the designated port.

In the MSTI2, the priority of the S-switch-D is higher than that of the S-switch-B. The S-switch-D is elected as the root of the MST2. Therefore, Ethernet0/0/3 of the S-switch-B is calculated as the root port.

The port Ethernet0/0/1 of the S-switch-B is an edge port and contained only in the MSTI0 and MSTI1. It is not included in the MSTI2. Therefore, it is in forwarding state in the MSTI0 and MSTI1 and not displayed in the MSTI2.

20.GSTP

20.1 GSTP Overview

The switch is connected with the client. If there is a loop in the client network, which will affect the entire network. GSTP is to solve this problem. After the GSTP is enabled on the switch port, the switch periodically sends a detection message. If the client network has a loop, the switch receives the detection message from the switch. In this case, the switch considers that the client network exists loop, and the port connected to the client port according to the treatment strategy placed discarding or shutdown.

Some people may ask, the spanning tree can also be remote loop detection, why need GSTP? This is because if the client network also has equipment to open spanning tree, the client network topology change easily affects the network of the room. The general networking is to connect the client port which does not open the spanning tree, with GSTP alternative.

20.2 GSTP Configuration

20.2.1 Enable Configuration

Enable all ports

operation	command	remark
Enter the global configuration mode	configure terminal	required
Enable all ports	[no] spanning-tree remote-loop-detect interface	required
View the configuration information	show spanning-tree remote-loop-detect interface	optional

Enable designed port

operation	command	remark
Enter the global configuration mode	configure terminal	required
Enable designed ports	[no] spanning-tree remote-loop-detect interface ethernet ***port-id***	required
Enter the port mode	interface ethernet ***port-id***	required
Enable port	[no]spanning-tree remote-loop-detect	required
View the configuration information	show spanning-tree remote-loop-detect interface [ethernet ***port-id***]	optional

📖 Note:

Enabling a Specified Port has two ways to configure a designated port: 1. Enter the specified port and enable GSTP. 2. Enter the specified port when the port is enabled globally. The same effect, only need to configure one..

20.2.2 Configure the Processing Policy

When GSTP detects the existence of loop, there are two ways: one is discarding the port, the other is the port shutdown, and then periodically restores the port; the default use discarding.

Configure the processing policy

operation	command	remark
Enter the global configuration mode	configure terminal	required
Configure the processing policy	spanning-tree remote-loop-detect action { shutdown | discarding }	optional

20.2.3 20.2.3 Configure the Recovery Timer

When GSTP detects that a loop exists and the shutdown command is used, the shutdown port periodically recovers the corresponding port. The default recovery period is 20 seconds and can be modified as needed. If it is configured as 60s, it means that it will not be automatically restored. User needs to manually run the shutdown / no shutdown command on the port. The port can re-linkup.

Configure the recovery timer

operation	command	remark
Enter the global configuration mode	configure terminal	required
Configure the shutdown processing policy	spanning-tree remote-loop-detect action shutdown	required
Configure the recovery time of the port	spanning-tree remote-loop-detect recover-time ***value***	optional
View the configuration information	show spanning-tree remote-loop-detect interface	optional

20.2.4 Configure the Detection Period

After the GSTP function is enabled, GSTP detection messages are periodically sent from the corresponding port. If the DST receives a GSTP message from itself, it considers that there is a loop and processes it according to the processing policy. The detection time is 5s by default, which allows user to modify the transmission time.

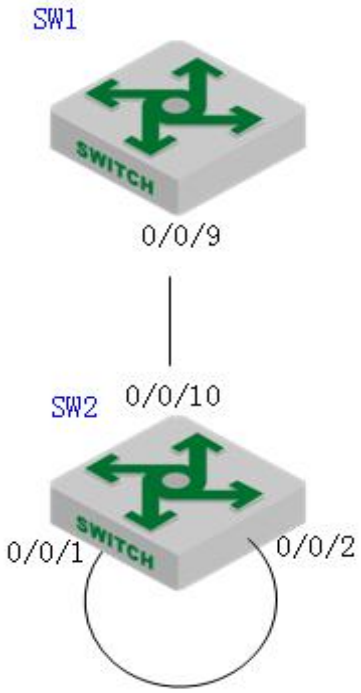
Configure the detection period

operation	command	remark
Enter global configuration mode	configure terminal	required
Configure the detection period	spanning-tree remote-loop-detect interval-time ***value***	optional
View configuration information	show spanning-tree remote-loop-detect interface	optional

20.2.5 GSTP Configuration Example

1. Networking Requirements

As shown in the figure, the port 9 of the SW1 enables GSTP and the SW1 is connected to the switch SW2. When there is a loop on the SW2, the SW1 detects that there is a loop under port 9, port 9 is discarding by default.



GSTP

2. Configuration procedure

#SW1 configuration：Enable the GSTP function of the port 9；

SW1(config)# spanning-tree remote-loop-detect interface ethernet 0/0/9

SW1(config)#interface range ethernet 0/0/9

SW1(config-if- 0/0/9)#no spanning-tree

SW2 connects port 1 and port 2 to form a single loop, SW1 GSTP is displayed as shown below：

SW1(config)#show spanning-tree remote-loop-detect interface ethernet 0/0/9

Loopback-detection action is Discarding

The interval time is 5 second(s)

The recovery time is 20 second(s)

Port Information:

port loopback status

e0/0/9 Enable Discarding

After the GSTP processing policy is changed to shutdown, the following is displayed：

SW1(config)#spanning-tree remote-loop-detect action shutdown

SW1(config)#show spanning-tree remote-loop-detect interface ethernet 0/0/9

Loopback-detection action is Shutdown
The interval time is 5 second(s)
The recovery time is 20 second(s)
Port Information:
port loopback status
e0/0/9 Enable Shutdown

21. ERRP Configuration

21.1 ERRP Overview

Ethernet Redundant Ring Protocol is a link layer protocol specifically designed for Ethernet ring. It prevents broadcast storms caused by data loops when the Ethernet ring is complete; when a link on the Ethernet ring is disconnected, the communication path between the nodes on the ring network can be quickly restored. Compared with STP, ERRP has the characteristics of fast topological convergence speed and convergence time independent of the number of nodes on the ring network.

In order to avoid conflict between ERRP and STP in calculating port congestion / release status, ERRP and STP are mutually exclusive on the enabled port. That is, the STP protocol cannot be enabled by the two ports connected to the ERRP ring, and STP can be enabled by the other ports.

21.1.1 Concept Introduction

ERRP region

The ERRP region is identified by an integer ID. A set of switch groups configured with the same domain ID, control VLAN and connected to each other form an ERRP domain. An ERRP domain has the following constituent elements:

- ERRP loop
- VLAN controlled by ERRP
- Master node
- Transport node
- Edge node and assistant edge node

ERRP loop

The ERRP ring is also identified by an integer ID, and an ERRP ring physically corresponds to a ring-connected Ethernet topology. An ERRP domain consists of an ERRP ring or multiple ERRP rings that are connected to each other. One of them is the master ring and the other ring is a sub-ring. The master ring and the sub-ring are distinguished by the specified level at the time of configuration. The level of the primary ring is 0 and the level of the sub-ring is 1.

The ERRP ring has two states:

Health state: All links of the ring are normal and the physical link of the ring is connected.

Fault state: The link on the ERRP ring is faulty. One or many physical links of the ring network are down.

Node role

The node on the ERRP ring is divided into the master node and the transit node. The node

role is specified by the user. The master node is the decision-making and control node for ring protection. Each ERRP ring must specify only one master node. All nodes except the master node are called transit nodes.

If more than one ERRP ring intersects, one of the intersecting nodes is designated as an edge node and the other intersecting node is designated as an assistant edge node. The role of the two nodes on the master ring is the transit node. The two nodes role of the sub-ring is the edge node and the assistant edge node. The specific role of the sub-ring can be specified by the user. There is no special requirement, mainly to distinguish the two nodes.

Port role

Each node of an ERRP ring has two ports connected to a ring. User can specify one of the ports as the primary port and the other port as the secondary port. The master port of the master node is used to send health detection message (hello message), received from the secondary port of the main node. The master port and secondary port of the transit node are functionally indistinguishable. To prevent the loop from causing broadcast storms, if the ERRP ring is normal, the secondary port of the master node is blocked and all the other ports are in the forwarding state.

If multiple ERRP rings intersect, the ports in the intersecting nodes that access both the primary ring and the sub-ring (that is, the port of the primary ring and the sub-ring common link) are called common ports at the same time. Only the ports that access the sub-rings are called edge ports. Conceptually, a public port is not considered to be a port of a sub-ring, it is regarded as part of the main ring, that is, the public link is the link of the primary ring, not the link of the sub-ring. The state change of the public link is only reported to the master node of the primary ring. The master node of the sub-ring does not need to know.

Control VLAN

Control VLAN is relative to the data VLAN, the data VLAN is used to transmit data messages, control VLAN is used to transmit ERRP protocol messages.

Each ERRP region has two control VLANs, called the primary control VLAN and the sub-control VLAN. The protocol message of the primary ring is propagated in the master control VLAN, and the protocol message of the sub-ring is propagated in the sub-control VLAN. User need to specify the primary control VLAN. The VLAN that is one greater than the master control VLAN ID, is used as the sub-control VLAN.

Only port (ERRP port) connecting the Ethernet of each switch belongs to the control VLAN, and the other ports cannot join the control VLAN. The ERRP port of the primary ring belongs to both the primary control VLAN and the sub-control VLAN. The ERRP port of the sub-ring belongs to the sub-control VLAN. The data VLAN can contain ERRP ports or non-ERRP ports. The primary ring is regarded as a logical node of the sub-ring. The protocol messages of the sub-ring are transmitted through the primary ring and processed in the primary ring as data messages. The protocol messages of the primary ring are transmitted only within the primary ring. Don’t enter sub-rings.

Query Solicit function

ERRP is used in conjunction with IGMP Snooping, if the topology of the ERRP changes, the

forwarding state of the port will be changed. If the multicast state is not updated through the IGMP Snooping module after the port state changes, the multicast forwarding may become abnormal. To introduce the query solicit function. When a topology change occurs in the ERRP, the device sends a query solicit message or a general IGMP query message to all the ports so that the member port re-initiates an IGMP report to update the multicast entry.

21.1.2 Protocol Message

HELLO message

The hello message is initiated by the master node, and detects loop integrity of the network. The master node periodically sends HELLO message from its primary port, and the transit node forwards the message to the next node, which is then received by the secondary port of the master node. Periodically send, and the sending period is Hello timer.

LINK_UP message

The LINK_UP message is initiated by the transit node, edge node, or assistant edge node that recovers the link. It informs the master node that there is link recovery on the loop. Trigger to send.

LINK_DOWN message

The LINK_DOWN message is initiated by the transit node, edge node, or assistant edge node that fails the link. It informs the master node that there is link failure on the loop, and the physical loop disappears. Trigger to send.

COMMON_FLUSH_FDB message

It is initiated by the master node, and informs the transit node, the edge node and the assistant edge node to update their respective MAC address forwarding tables. Trigger on link failure or link recovery.

COMPLETE_FLUSH_FDB message

It is initiated by the master node, and informs the transit node, the edge node and the assistant edge node to update their respective MAC address forwarding tables, and informs the transit node to release the blocked state of the port temporarily blocking the data VLAN. It is sent when the link recovery (That is, the secondary port of the master node receives Hello packets) is complete.

EDGE_HELLO message

The EDGE_HELLO message is initiated by the edge node of the sub-ring to check the loop integrity of the major ring in the domain.

Edge nodes send EDGE_HELLO messages periodically from the two ports connected to the primary ring. The nodes in the primary ring process the message as data message and receive them from the assistant edge nodes on the same sub-ring. Periodically send, sending cycle is the

Edge Hello timer.

MAJOR_FAULT message

The MAJOR_FAULT message is originated by the assistant edge node and reports to the edge node that the primary ring of the domain is faulty. When the assistant edge node of the sun-ring cannot receive the EDGE_HELLO message from the edge node in the specified time, the assistant edge node sends a MAJOR_FAULT message from its edge port. After the sub-ring node receives the message, it forwards the message directly to the next node, and finally the edge node of same sub-ring receives. Periodically send after triggering, the sending period is Edge Hello timer.

21.1.3 Operate Principle

Health status

The master node periodically sends the hello message from its primary port, which in turn travels through the transit nodes of the ring. If the secondary port of the master node receives a hello message before it times out, it considers that the ERRP ring is health status. The status of the master node reflects the health of the ring. When the ring network is in a healthy state, the master node blocks its secondary port in order to prevent the data message from forming a broadcast loop.

Link failure

Two mechanisms are provided for detecting link failures:

(1) LINK_DOWN escalation and processing:

When an ERRP port of the transit node detects a port Link Down, the node sends a LINK_DOWN message to the master node from the ERRP PORT in the up state that is paired with the faulty port.

After the master node receives the LINK_DOWN message, the node state is immediately changed for failed state. Disable the blocking state of the secondary port. The FDB table is refreshed and a COMMON_FLUSH_FDB message is sent from the primary and secondary ports to notify all transit nodes to refresh their respective FDB tables.

After receiving the COMMON_FLUSH_FDB message, the transit node immediately refreshes the FDB table and starts learning the new topology.

(2) Polling mechanism:

The fault reporting mechanism is initiated by the transit node. In order to prevent the LINK_DOWN message from losing during transmission, the master node implements the Polling mechanism. The Polling mechanism is the mechanism that the master node of the ERRP ring actively detects the health status of the ring network. The master node periodically sends HELLO message from its master port, and then transmits it through the transmission nodes.

If the master node can receive the HELLO message from the secondary port in time, it indicates that the ring network is complete and the master node will keep the secondary port blocked. If the secondary port of the master node cannot receive HELLO message in the specified time, it is considered that a link fault has occurred on the ring network. The fault handling process is the same as the LINK_DOWN process mechanism.

Link recovery

There are two situations to deal with:

(1) LINK_UP escalation and processing

After the ports of the transit node that belong to the ERRP region are re-up, the master node may find loop recovery after a certain period of time. In the time, the network may form a temporary loop, which makes data VLAN produce a broadcast storm.

In order to prevent the generation of the temporary loop, the transit node moves to the Preforwarding state and immediately blocks the port that has just been recovered, after it finds the port accessing the ring network re-up. At the same time, the transmitting node that has recovered the link sends a LINK_UP message to the master node from ERRP port that is paired with the recovery port in the UP state. After receiving the LINK_UP message from the transmitting node, the master node sends a COMMON_FLUSH_FDB message from the primary port and the secondary port to notify all transit nodes to refresh the FDB table. The port recovered by the transit node only releases the blocked state after receiving the COMPLETE_FLUSH_FDB packet sent by the master node or the Preforward timer expires.

The response of the master node to the LINK_UP message does not represent the response processing to the ring network recovery. If multiple links on the ring network fail and then one of the links is restored, the LINK_UP reporting mechanism and the response mechanism of the master node are introduced to quickly refresh the FDB tables of the nodes on the ring.

(2) Ring network recovery processing:

Ring network recovery processing is initiated by the main node. The master node sends the Hello messages periodically from the master port. After the faulty link on the ring network is restored, the master node will receive its own test messages from the secondary port. After receiving the HELLO message from the host, the master node first moves the state back to the complete state, blocks the secondary port, and then sends the COMPLETE_FLUSH_FDB message from the primary port. After receiving the COMPLETE_FLUSH_FDB message, the transit node moves back to the Link_Up state, releases the temporarily blocked port, and refreshes the FDB table.

If the COMPLETE_FLUSH_FDB message is lost during transmission, a backup mechanism is adopted to recover the temporarily blocked port of the transit node. The transmission node is in the Pre-forwarding state, if the COMPLETE_FLUSH_FDB message from the master node is not received in the specified time, Self-release temporary blocking port, restore data communication.

21.1.4 Multi-loop Intersection Processing

Multi-ring and single-ring is almost the same, The difference between a multi-ring and a single ring is that multiple rings are introduced the sub-ring protocol message channel state detection mechanism in the main ring, after the channel is interrupted, the edge port of the edge node is blocked before the secondary port of the master node of the sub-ring is released to prevent the data loop from forming between the sub-ring. For details, see Sub-channel Protocol Channel Status Check Mechanism on the Main Ring.

In addition, when a node on the master ring receives a COMMON-FLUSH-FDB or

COMPLETE-FLUSH-FDB message from the sub-ring, it will refresh the FDB table. The COMPLETE-FLUSH-FDB of the sub-ring does not cause the sub ring transit node to release the temporarily blocked port. The COMPLETE-FLUSH-FDB message of the primary ring does not do so.

🕮 Note:

Before the interface starts IGMP, the multicast protocol must be enabled. In addition, if you need to cooperate with the PIM, you must configure the PIM protocol on this interface at the same time. See the PIM-DM/SM configuration guide.

21.2 ERRP Configuration

21.2.1 Enable/disable ERRP

By default, ERRP is disabled and need to be configured in global mode.

Enable/disable ERRP

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enable/disable ERRP	[no] errp	required

21.2.2 Configuration Domain

When creating an errp domain, user needs to specify the domain ID, and must configure the same domain ID on all the nodes in the same domain. Create up to 16 domains on a device.

Configuration domain

operation	command	remark
Enter the global configuration mode	configure terminal	-
Create and enter the domain configuration mode	errp domain ***domain-id***	required
Exit domain mode	**exit**	optional
Delete the domain	no errp domain ***domain-id***	optional
View the domain information	show errp domain ***domain-id***	optional

21.2.3 Configure Control VLAN

Control VLAN is relative to the data VLAN, the data VLAN is used to transmit data message, control VLAN is used to transmit ERRP protocol message.

Each ERRP domain has two control VLANs, called the primary control VLAN and the sub-control VLAN. The protocol messages of the primary ring are propagated in the master control VLAN, and the protocol messages of the sub-ring are propagated in the sub-control VLANs. User needs to specify only the primary control VLAN and a VLAN with the maximum control VLAN ID of 1 as the sub-control VLAN.

When an ERRP port sends protocol messages, it always takes control VLAN tags, regardless of whether the ERRP port is in trunk mode.

Configure control VLAN

operation	command	remark
Enter the domain configuration mode	errp domain ***domain-id***	-
Configure control vlan	[no] control-vlan ***vlan-id***	required
View control vlan	show errp control-vlan	optional

21.2.4 Configure the Ring

To avoid conflict between ERRP and STP in calculating port blocking / releasing status, ERRP and STP are mutually exclusive on the port. Before specifying an ERRP port, user must disable STP on the port.

If a device is on multiple ERRP rings of the same ERRP domain, only one master ring can exist. The node role of the device on other sub-rings can be only the edge node or assistant edge node.

The ERRP field takes effect only when both the ERRP protocol and the ERRP ring enable. To enable the ring, user must first configure the control VLAN.

ERRP ring is divided into the main ring and sub-ring. Respectively use 0,1 .

Configure loop

operation	command	remark
Enter the domain configuration mode	errp domain *domain-num*	-
Configure ring and ring levels	ring ***ring-id*** role [master | transit] primary-port [ethernet port | channel-group ***lacp-id***] secondary-port [ethernet port | channel-group ***lacp-id***] level ***level-number***	required
Enable the ring	ring ***ring-id*** enable	required
Close the ring	ring ***ring-id*** disable	optional
Delete the ring	no ring ***ring-id***	optional
View the ring information	show errp domain ***domain-id*** ring ***ring-id***	optional

21.2.5 Configure Node Role

Configure node role

operation	command	remark
Enter the domain configuration mode	errp domain ***domain-id***	-
configure node role	ring ***ring-id*** role { **master | transit** } primary-port [ethernet ***port-id*** | channel-group ***lacp-id***] secondary-port [ethernet ***port-id*** | channel-group ***lacp-id***] level *level-number*	required
	ring ***ring-id*** role { ***edge | assistant-edge*** } common-port [ethernet ***port-id*** | channel-group ***lacp-id***] edge-port [ethernet ***port-id*** | channel-group ***lacp-id***]	

21.2.6 Configure Port Role

Configure port role

operation	command	remark
Enter the domain configuration mode	errp domain *domain-num*	-
Configure the port role	ring ***ring-id*** role **{** master | transit **}** **primary-port [ethernet *port-id* |** channel-group ***lacp-id*] secondary-port [ethernet *port-id* | channel-group *lacp-numb*]** level *level-number*	required
	ring ***ring-id*** role **{** edge | assistant-edge **}** **common-port [ethernet *port-id* | channel-group *lacp-id*]** edge-port **[ethernet *port-id*** | channel-group ***lacp-id*]**	

21.2.7 Configure Work Mode

In order to connect with other vendors' device, user can modify the work mode in the ERRP domain, and configure multiple ERRP domains on the same device. Each domain can be configured with different work modes. All the nodes in the same ERRP domain must work in the same mode.

By default, it works in standard mode. Support compatible with EIPS and RRPP.

Configure work mode

operation	command	remark
Enter the domain configuration mode	errp domain ***domain-id***	-
Configure the standard mode	work-mode standard	optional
Run compatible EIPS mode	work-mode eips-subring	optional
Run compatible RRPP mode	work-mode huawei	optional

21.2.8 Configure Query Solicit

This function is used to cooperate with IGMP SNOOPING. When the topology of the ERRP ring network changes, it immediately notifies the IGMP querier to resend the IGMP general query to update the IGMP SNOOPING multicast database in time. Currently, there is not related standard. The query solicit message is private and the IGMP type is 0xff.

Specific implementation is as follows:

1. The default Query solicitation function is enabled on the master node, the transit node closes Query solicitation function.
2. The master node topology change is determined by: The master node status is from Health to Fault or from Fault to Health.
3. Other nodes topology changes are determined by: The primary and secondary port status is from forwarding to non-forwarding (block/disable) or from non-forwarding to forwarding

(block/disable).

4. When the node detects a topology change: If the node itself is the IGMP querier, it immediately sends a General Query message to all the ports. Otherwise, immediately send a Query Solicit message to all ports;

5. After the IGMP querier receives the Query Solicit message: Respond immediately to the receiving port a General Query message.

Configure Query Solicit

operation	command	remark
Enter the domain configuration mode	errp domain ***domain-id***	-
Configure the query function	[no] ring ***ring-id*** query-solicit	optional

21.2.9 Configure Time Parameter

User can modify the ERRP timer parameters as requirement, but make sure that the timer parameters are the same on all nodes. Ensure that the value of the Failed timer is not less than 3 times the Hello timer value.

Configure time parameter

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the health message timer	errp hello-timer ***value***	optional
Configure the information timeout timer	errp fail-timer ***value***	optional
Configure the recovery delay timer	errp preup-timer ***value***	optional

21.2.10 Configure the Topology Discovery Function

Configure the topology discovery

operation	command	remark
Enter the domain configuration mode	errp domain ***domain-id***	-
Enable topology discovery	[no]topo-collect	required
Topology Information View	show errp topology [domain ***domain-id*** | summary]	optional

21.2.11 Clear Protocol Message Statistic

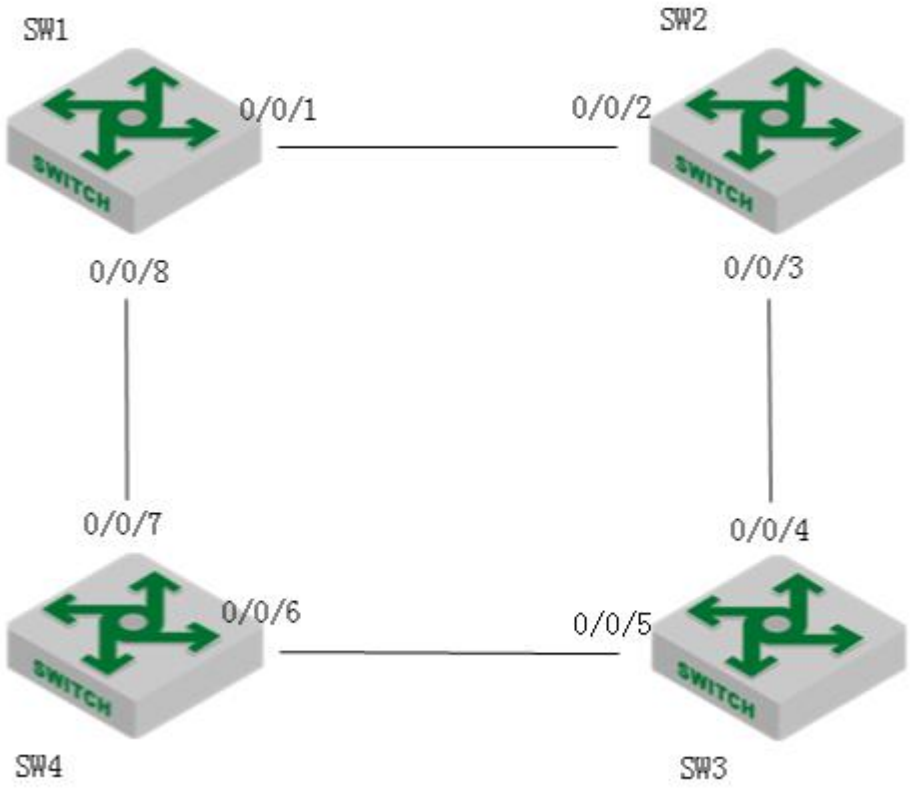
Clear protocol message statistic

operation	command	remark
Enter the global configuration mode	configure terminal	-
Clear the statistics	clear errp [domain ***domain-id*** [ring ***ring-id***]]	optional

21.3 ERRP Configuration Example

I. Network requirements

As shown in the following figure, four DUTs form a single ring and run ERRP.



ERRP

2. Configuration procedure

The ERRP configuration on SW1 is as follows:

SW1(config)# interface range ethernet 0/0/1 ethernet 0/0/8

SW1(config-if-range)#no spanning-tree

SW1(config-if-range)#exit

SW1(config)#errp domain 0

SW1(config-errp-domain-0)#ring 0 role master primary-port ethernet 0/0/1 secondary-port ethernet 0/0/8 level 0

SW1(config-errp-domain-0)#control-vlan 100

SW1(config-errp-domain-0)#ring 0 enable

SW1(config-errp-domain-0)#exit

SW1(config)#errp

The ERRP configuration on SW2 is as follows:

SW2(config)# interface range ethernet 0/0/2 ethernet 0/0/3

SW2(config-if-range)#no spanning-tree

SW2(config-if-range)#exit

SW2(config)#errp domain 0
SW2(config-errp-domain-0)#ring 0 role transit primary-port ethernet 0/0/2 secondary-port ethernet 0/0/3 level 0
SW2(config-errp-domain-0)#control-vlan 100
SW2(config-errp-domain-0)#ring 0 enable
SW2(config-errp-domain-0)#exit
SW2(config)#errp

The ERRP configuration on SW3 is as follows:
SW3(config)# interface range ethernet 0/0/4 ethernet 0/0/5
SW3(config-if-range)#no spanning-tree
SW3(config-if-range)#exit
SW3(config)#errp domain 0
SW3(config-errp-domain-0)#ring 0 role transit primary-port ethernet 0/0/4 secondary-port ethernet 0/0/5 level 0
SW3(config-errp-domain-0)#control-vlan 100
SW3(config-errp-domain-0)#ring 0 enable
SW3(config-errp-domain-0)#exit
SW3(config)#errp

The ERRP configuration on SW4 is as follows:
SW4(config)# interface range ethernet 0/0/6 ethernet 0/0/7
SW4(config-if-range)#no spanning-tree
SW4(config-if-range)#exit
SW4(config)#errp domain 0
SW4(config-errp-domain-0)#ring 0 role transit primary-port ethernet 0/0/6 secondary-port ethernet 0/0/7 level 0
SW4(config-errp-domain-0)#control-vlan 100
SW4(config-errp-domain-0)#ring 0 enable
SW4(config-errp-domain-0)#exit
SW4(config)#errp

3. Validation results
View the control VLAN configuration
SW1(config)#show errp control-vlan
VLAN name : ERRP domain 0 primary-control-vlan
VLAN ID : 100
VLAN status : ERRP used only
VLAN member : e0/0/1,e0/0/8.
Static tagged ports : e0/0/1,e0/0/8.
Static untagged ports :

VLAN name : ERRP domain 0 sub-control-vlan
VLAN ID : 101

VLAN status : ERRP used only
VLAN member : e0/0/1,e0/0/8.
Static tagged ports : e0/0/1,e0/0/8.
Static untagged ports :

Total entries: 2 vlan.

View the status of the ERRP loop as follows:
SW1(config)#show errp domain 0
ERRP state: enable
Time value: hlth 1, hlthFl 6, mjrFlt 5, preFwd 6, preup 0
domain 0 info: control-vlan 100, work-mode standard, topo-collect disable
ring 0 info:
status: active
role : master
level : 0
stm : COMPLETE
query solicit: enable
primary/common port: e0/0/1 forwarding
rcv-pkts: 0hlth,0comn,0cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
snd-pkts: 6hlth,0comn,1cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
secondary/edge port: e0/0/8 blocking
rcv-pkts: 6hlth,0comn,1cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
snd-pkts: 0hlth,0comn,0cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
Total 1 ring(s).

The status of the loop fault is as follows:
SW1(config)#show errp domain 0
ERRP state: enable
Time value: hlth 1, hlthFl 6, mjrFlt 5, preFwd 6, preup 0
domain 0 info: control-vlan 100, work-mode standard, topo-collect disable
ring 0 info:
status: active
role : master
level : 0
stm : FAULT
query solicit: enable
primary/common port: e0/0/1 forwarding
rcv-pkts: 0hlth,0comn,0cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
snd-pkts: 99hlth,1comn,1cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
secondary/edge port: e0/0/8 forwarding
rcv-pkts: 95hlth,0comn,1cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt
snd-pkts: 0hlth,0comn,0cmplt,0lnkdn,0lnkUp,0edgHlo,0mjrFlt

22. ERPS Configuration

22.1 ERPS

22.1.1 ERPS Overview

ERPS (Ethernet Ring Protection Switching) is released by ITU-T with the convergence rate of telecommunication level. If all devices inside the ring support this agreement, it can achieve intercommunication.

22.1.2 ERPS Basic Conception

ERPS mainly includes ERPS ring, node, port role and port status.

1. ERPS Example

EPRS instance is formed by the same instance ID, control VLAN and interconnected Switches.

2. Control VLAN

Control VLAN is the transmission VLAN of ERPS protocol, and the protocol packet will carry corresponding VLAN tag.

3. RPL

RPL (Ring Protection Link), Link designated by mechanism that is blocked during Idle state to prevent loop on Bridged ring

4. ERPS ring

ERPS ring is EPRS basic unit. It composed by a set of the same control VLAN and the interlinked L2 Switch equipment.

5. Node

The L2 Switch equipment added in ERPS ring are called nodes. Each node cannot be added to more than two ports in the same ERPS ring. The nodes are divided into RPL Owner, Neighbor, Next Neighbor, and Common.

6. Port Role

In ERPS, port roles include: RPL Owner, Neighbor, Next Neighbor, and Common:

① RPL Owner: An ERPS ring has only one RPL Owner port configured by the user and it prevents loops in the ERPS ring via blocking the RPL Owner port. The node that owns the RPL Owner port becomes the RPL Owner node.

② RPL Neighbour: An ERPS ring has only one RPL Neighbor port configured by the user and it must be a port connected to the RPL Owner port. If the network is normal, it will block together with the RPL Owner port to prevent loops in the ERPS ring. The node with the RPL Neighbor port becomes the RPL Neighbor node.

③ RPL Next Neighbour: An ERPS ring can have up to two RPL Next Neighbor ports configured by the user. It must be the port connecting the RPL Owner node or the RPL Neighbor node. To become the RPL Next Neighbor node, the RPL Next Neighbor port should own the node of RPL Next Neighbor port.

Note: RPL Next Neighbour nodes are not much different from ordinary nodes. They can be replaced by Common nodes.

④ Common: The common port. The ports except RPL owner, Neighbor and Neighbor port are common ports. If the node has only the Common port, this node will become the Common node.

7. Port Status

In the ERPS ring, the port status of the ERPS protocol is divided into three types.

① Forwarding: In *Forwarding* status, the port forwards user traffic and receives / forwards R-APS packets. Moreover, it forwards R-APS packets from other nodes.

② Discarding: In the *Discarding* status, the port can only receive / forward R-APS packets and cannot forward R-APS packets from other nodes.

③Disable: port in *Linkdown* status.

8. Wrok Mode: ERPS operating mode

Work mode includes: revertive and non-revertive.

① Revertive: When the link fails, the RPL link is in the release protection state and the RPL link is re-protected after the faulty link is restored to prevent loops.

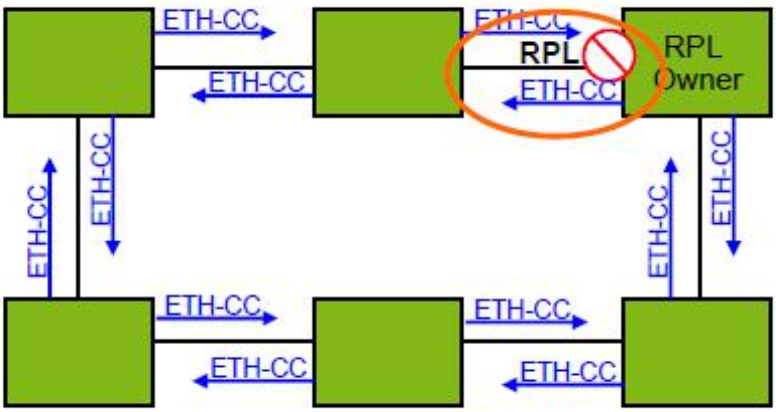
Non-revertive: After the fault is rectified, the faulty node remains faulty (without entering Forwarding) and the RPL link remains in the release protection state

22.1.3 ERPS Ring Protection Mechanism

ERPS uses ETH CFM for link monitoring. When the network is normal, a blocking link is set on the ring network to prevent the ring network from ringing. If a fault occurs in the network, a blocked backup link is opened to ensure uninterrupted link between each node. The general process is as follows:

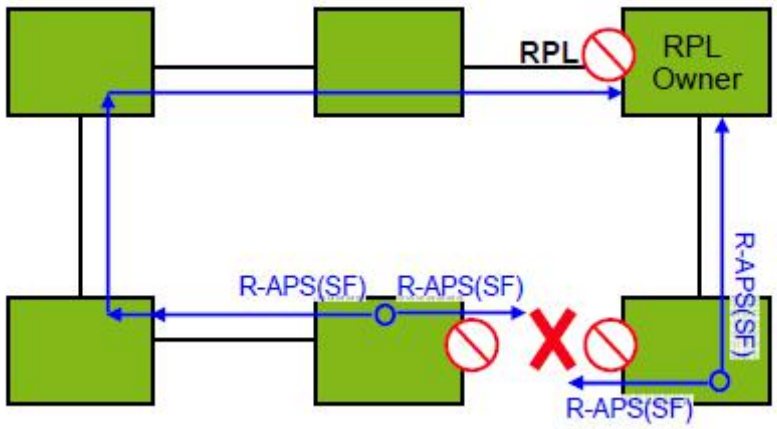
As shown in Figure 1-1, when six devices are connected in a ring and the link is in the IDLE state, the loop is removed via setting the RPL link and locking the port (RPL Owner port).

Figure 1-1 loop diagram in IDLE state



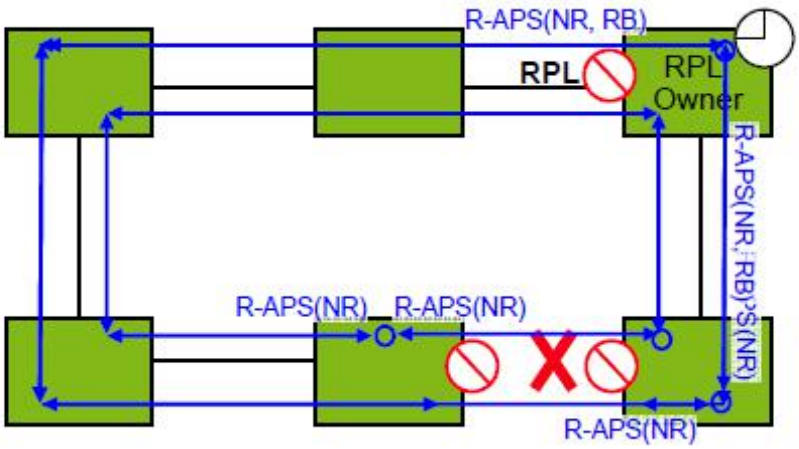
② When a node on the link detects a fault, it immediately blocks the faulty node and reports the fault message (R-APS (SF)) to all the other devices in the ring. After receiving the message, all other nodes refresh the FDB. The RPL owner port receives the fault message, and the recovery port is in the forwarding state. The ERPS ring enters the protection state. As shown in Figure 1-2:

Figure 1-2 loop diagram in ring network protection fault (link fault)



③As shown in Figure 1-3, when the link of the faulty device recovers, it sends RAPS (NR) packets to other devices in the ring to inform them that there is no local request. When the RPL owner receives the packet, it will block the port and send the R-APS (NR, RB) message again after some time. After receiving the packet, the other nodes will refresh the FDB entry. Later, the port of the faulty node will be restored to the forwarding state, and the ring will revert to the IDLE state (Figure 1-1)

Figure 1-3 loop diagram in loop protection (link recovery) state



22.2 ERPS Configuration

22.2.1 Enable ERPS

Enable ERPS

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable ERPS	**erps**	required
Disable ERPS	**no erps**	

22.2.2 ERPS Configuration Example

ERPS Configuration Example

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure erps instance	**erps instance** *instance-id*	required
Configure control-vlan	**control-vlan** *vlan id*	required
Configure work-mode	**work-mode {***revertive*|*non revertive***}**	optional
Configure ring id	**ring** *ring id*	optional
Configure ring level	**ring** *level*	optional， 0 refers to main ring, 1 refers to subring.
Configure ring port	**{Port0|port1} ethernet** *interface-num* **{**neighbour | next-neighbour |owner }	required
Enable ring	**ring enable/disable**	Required. Before enabling Ring, you should configure port and control VLAN.

Note:

About Ring ID: ERPS ring ID, the last byte of the DMAC in the R-APS message is Ring Id. From G.8032 can be learned that the ERPS ring ID can be the same, and the control VLAN needs to be different. The reverse is also true. The ring ID of each instance can be 1 to 239, and the control VLAN does not allow duplication.

To configure ERPS port, you must disable the spanning tree.

22.2.3 Configure Connectivity Detection of ERRP Link

In ERPS, there is no HELLO packet to monitor link connectivity in real time. Instead, it uses the CC function in ETH CFM to detect the link connectivity by sending ETH-CC messages between the two ports. Therefore, you need to configure the CFM CC for the ports in the ERPS. In the ERRP instance, you need to configure the MEL (MEG level, which must be consistent with the CFM configuration).

For more information about CFM, please refer to the CFM User Manual.

Configure connectivity detection of ERRP link

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure erps instance and then enter instance configuration mode	**erps instance** *instance-id*	optional
Configure MEL	**mel** *level*	Optional; MEG level; It needs to be the same as CFM configuration

🕮 Note:

CFM ETH-CC function is to detect non-LINKDOWN link failure, such as single-pass. If you do not use the CFM ETH-CC function, ERPS can also work normally while it cannot detect single-pass failure.

22.2.4 Configure ERPS Related Timers

ERPS has two timers: WTR timer and Guard timer.

① ①WTR timer: When the RPL owner port is restored to the Forwarding state due to another device or link failure, if the fault is restored and some ports may not have been

changed from the Down state to the Up state, it starts the WTR timer when the RPL owner port receives the fault-free RAPS packet from a port to prevent the shock of blocking point; If the fault is received before the timer expires, the WTR timer is disabled. If a faulty RAPS packet from another port is received before the timer times out, the WTR timer will be disabled. If the WTR timer does not receive any faulty RAPS packets from other ports, it will block the RPL Owner port and send RPL blocking RAPS packets after timed out. After receiving the packet, the other ports set the forwarding state of its own port as *Forwarding* state.

② Guard timer: After the failure recovery, the equipment involved in link failure or node failure will send R-APS packet to the other devices and it will start the Guard Timer at the same time. The device does not process RAPS packets until the timer times out with the purpose to prevent the receipt of outdated faulty R-APS packets. If the device receives the faulty RAPS packet from another port after the timer times out, the port forwarding state will turn to *Forwarding*.

Configure ERPS Related Timers

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure erps instance and then enter instance configuration mode	**erps instance** *instance-id*	required
Configure wtr-timer timer	**wtr-timer** *timer value*	Optional, 5min by default, 1~12min
Configure guard-timer timer	**guard-timer** *timer value*	Optional; By default, 0～2s for 500ms

22.2.5 ERPS Display and Maintenance

After you completed the above configurations, you can use the following commands to view the configurations.

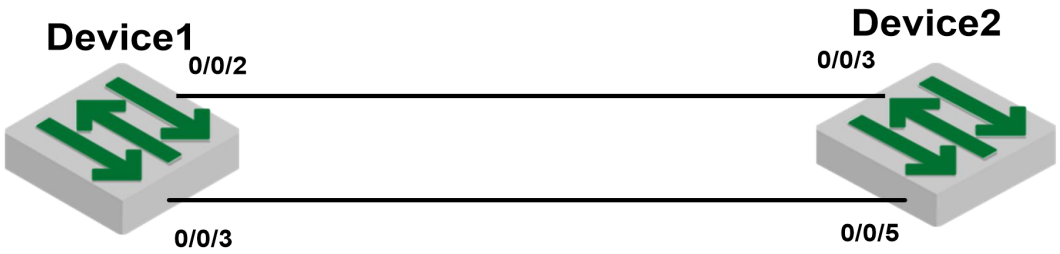
ERPS display and debug

Operation	Command	Remarks
Display ERPS information	**show erps [***instance instance id***]**	
Display the sending and receiving packets	**show erps [***instance instance id***] statistics**	
Clear the sending and receiving packets	**clear erps [***instance instance id***] statistics**	

22.3 ERPS Configuration Example

22.3.1 Demand and networking

The two switches form a single ring. The ETH CFM detects the link fault and eliminates the loop through ERPS. The network diagram is as follows:



network diagram of ERPS configuration

22.3.2 Configuration on ERPS Ring Network Protection

1) Configure ETH CFM on DEVICE1 and DEVICE2

#DEVICE1 CFM configuration

DEVICE1(config)#s run cfm

![CFM]

cfm md 1

cfm md format string name test Erps level 1

cfm ma 1

cfm ma format string name test Erps primary-vlan 100

cfm mep 1 direction down interface ethernet 0/0/3

cfm mep 1 state enable

cfm mep 1 cc enable

cfm rmep 2 mep 1

exit

exit

#DEVICE2 CFM configuration

DEVICE2(config)#s run cfm

![CFM]

cfm md 1

cfm md format string name test Erps level 1

cfm ma 1

cfm ma format string name test Erps primary-vlan 100

cfm mep 2 direction down interface ethernet 0/0/2

cfm mep 2 state enable

cfm mep 2 cc enable

cfm rmep 1 mep 2

exit

exit

2) Configure ERPS on DEVICE1 and DEVICE2

ERPS configuration of DEVICE1

DEVICE1(config)#s running-config erps

![ERPS]

erps

erps instance 1

mel 1

ring level 1

control-vlan 100

port0 ethernet 0/0/2 owner

port1 ethernet 0/0/3

ring enable

exit

ERPS configuration of DEVICE2

DEVICE2 (config)#s run erps

![ERPS]

erps

erps instance 1

mel 1

ring 1

ring level 1

control-vlan 100

port0 ethernet 0/0/3 neighbour

port1 ethernet 0/0/5

ring enable

exit

🕮 Note:

- cfm md format string name test Erps level 1

Here, the “level” refers to MEG Level, that is, the MEL in ERPS needs to be configured into the same.

- cfm ma format string name test Erps primary-vlan 100

Here, the primary-vlan 100 should be the same as the ERPS control VLAN.

- mep and rmep need one-to-one correspondence

22.3.3 Result Verification

The 0/0/2 port of Device 1 and the 0/0/3 port of Device 2 are blocked and the loop is removed.

```
DEVICE1 (config)#show erps
ERPS state: enable

Instance Id : 0
Mel              : 1
Work-mode      : revertive
Time value    : WTR 5 min, guard timer 500 ms, holdoff timer 0 s
Ring 1 info:
Control vlan: 100
Status           : enable
Node Role      : owner
Level            : 1
Stm              : Idle
          portId    role      state          nodeId                  BPR
port0    e0/0/2    owner    Blocking      00:01:7f:00:00:11    0
port1    e0/0/3    Common   Forwarding    10:7b:ef:fd:4b:cd    0
```

23.Static Routing Configuration

23.1 Static Routing Overview

Layer3 Switch is a kind of gigabit intelligent routing Switches based on ASIC technology. There is a layer-3 forward routing table maintained in system, applied to specify the next-hop address of some certain destinations and related information. All of these routings can be learned dynamically via certain routing protocol, of course, manually adding is OK. Static routing refers to the routing which specified to one or a certain field by manual.

23.2 Detailed Configuration of Static Routing Table

23.2.1 Add/Delete Static Routing Table

static routing basic configurations

Operation	Command	Remarks
Enter global configuration mode	**ip route** *dst-ip mask gate-ip*	
Delete specified static routing table	**no ip route** dst-ip mask [gate-ip]	
Delete all static routing table	**no ip route static all**	

Note:

gate-ip: the next-hop address of static routing. Moreover, it is dot-decimal notation format.

dst-ip: the destination address of static routing which you are going to add. Moreover, it is dot-decimal notation format.

mask: the mask of destination address. Moreover, it is dot-decimal notation format.

23.2.2 Add/ Delete Static Routing Backup Table

There are multiple static routers forwarded to a certain ip address or network segment. It forms a primary router and multiple backup routers according to the priorities. It will switch to backup router which possesses the highest priority if the primary router loses efficacy. However, it will switch to the primary router again when the primary router renews.

static routing basic configurations

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure static routing backup	**ip route** *dst-ip mask gate-ip* **priority** *priority*	
Delete static routing backup	**no ip route** *dst-ip mask gate-ip* **priority** *priority*	

Note: The smaller the value is, the higher priority it will be. And the priority with smallest value will be the primary route.

23.2.3 Display Routing Table Information

This command is used to display the related information of specified routing table, including the next-hop address, routing type and so forth. It can be displayed the information of specified destination address routings, all static routings and all routings. If you do not input the parameter, it will display all routings information.

Display Routing Table Information

Operation	Command	Remarks
Display routing table	**show ip route** [*ip-address* [*mask*] | static | rip |ospf]	
Display static routing backup table	**show ip route priority static**	

Note:

ip-address: the destination address. In addition, it is dot-decimal notation format.

mask: the destination network segment presented with IP address. In addition, it is dot-decimal notation format.

static: display all static routing tables

rip: display all rip routing tables

ospf: display all ospf routing tables

priority: the priority of the static routing table. The smaller the value is, the higher priority it will be.

23.3 Configuration Example

```
! add a network routing to 192.168.0.100, set 10.11.0.254 as next-hop
Switch(config)#ip route 192.168.0.100 255.255.0.0 10.11.0.254
! delete a network routing to 192.168.0.100
Switch(config)#no ip route 192.168.0.100 255.255.0.0
! delete all static routing
Switch(config)#no ip route static all
! display the routing information of 192.168.0.100
Switch(config)#show ip route 192.168.0.100
! display all routing information
Switch(config)#show ip route
! display all rip information
Switch(config)#show ip route rip
! display all ospf information
Switch(config)#show ip route ospf
```

24.RIP Configuration

24.1 RIP Overview

Routing Information Protocol (RIP) is a routing protocol based on the Distance-Vector (D-V) algorithm and has seen wide deployment. It exchanges routing information by sending route update packets over the User Datagram Protocol (UDP) every 30 seconds. If having not received a route update packet from the peer router within 180 seconds, the local router marks all the routes from the peer router as unreachable. If no update packet is received from the peer router

yet in 120 seconds after a route is marked as unreachable, the local router deletes the route from its routing table.

RIP uses Hop Count as a routing metric to measure the distance from a destination host. In a RIP network, Hop Count is 0 if a router is directly connected with a network and 1 if a route needs to traverse a router before reaching the destination network, and so on. To restrain the route convergence time, RIP stipulates that Hop Count is an integer ranging from 0 to 15. The distance is considered infinite if Hop Count is larger than or equal to 16. In this case, the destination network or host is unreachable.

RIP has two versions: RIP-1 and RIP-2 (support for plaintext authentication).

To improve routing performance and avoid routing loops, RIP presents the concepts of Split Horizon and Poison Reverse.

Each RIP router manages one routing database, which contains all the destination reachable routing entries on a network. These routing entries include the following information:

Destination address: IP address of a host or network;

Next-hop address: address of a next router on the route to a destination;

Outbound interface: interface from which packets are forwarded;

Metric value: cost of a route from the local router to a destination, which is an integer from 0 to 16.

Timer: time counted from the last modification of a routing entry. The timer is zeroed every time the routing entry is modified.

The RIP startup and operation procedure is described as follows:

(1) Upon RIP startup on a router, the router broadcasts a request packet to its neighboring routers. After receiving the request packet, the neighboring routers (with RIP started) return a response packet which contains the information about their respective local routing tables.

(2) Upon receipt of the response packets, the router that sends the request packet modifies its local routing table.

(3) RIP broadcasts or multicasts the local routing table to its neighboring routers every

30s. The neighboring routers maintain their local routes to select a best route and then broadcast or multicast the modification to their respective neighboring networks, so that the routing update will eventually take effect globally. RIP employs a timeout mechanism to process expired routes, ensuring that the routes are latest and valid. As an interior routing protocol, RIP helps acquaint routers with the network-wide routing information because of these mechanisms.

RIP has been accepted as one of the standards which regulates the route transmission between a router and a host. Layer3 Switches forward IP packets across a LAN the same way as routers. Therefore, RIP is also widely deployed on L3 Switches. It is applicable to most campus networks and regional networks with a simple structure and good continuity but not recommended in complex large networks.

24.2 RIP Configuration

24.2.1 Enable /Disable RIP Mode

enable /disable RIP mode

Operation	Command	Remarks
Enable RIP mode	**router rip**	-
Disable RIP mode	**no router rip**	-

24.2.2 Specify the IP Network Segment to Run RIP

By default, an interface does not send or receive RIP packets until the IP network segment to run RIP is specified by the administrator even if RIP is enabled on the interface.

Specify the IP network segment to run RIP

Operation	Command	Remarks
Specify the IP network segment to enable RIP	**network** *ip-address*	The ip should be interface ip
Specify the IP network	**no network** *ip-address*	-

segment to disable RIP		

24.2.3 Specify the RIP Operation State for an Interface

In interface configuration mode, the RIP operation state can be specified for an interface, for example, whether to run RIP on the interface (whether to enable the interface to send and receive RIP update packets) and you can also specify to send (or receive) RIP update packets individually.

Specify the RIP operation state for an interface

Operation	Command	Remarks
Specify the interface to enable RIP	**ip rip work**	-
Specify the interface to disable RIP	**no ip rip work**	-
Allow the interface to receive RIP packets.	**ip rip input**	
Forbid the interface to receive RIP packets.	**no ip rip input**	
Allow the interface to send RIP packets.	**ip rip output**	
Forbid the interface to send RIP packets.	**no ip rip output**	

24.2.4 Specify the RIP Version for an Interface

RIP has two versions: RIP-1 and RIP-2. You can specify the version of the RIP packets to be processed by an interface.

RIP-1 packets are transmitted in broadcast mode. RIP-2 packets may be transmitted in either broadcast or multicast mode. The multicast mode is used by default. In RIP-2, the multicast address is 224.0.0.9.

When the multicast mode is used, non-RIP hosts in the same network will not receive RIP broadcast packets and RIP-1 hosts will not receive or process the RIP-2 routes with a subnet mask. A RIP-2 interface can also adopt broadcast and receive the RIP-1 broadcast packets.

Specify the RIP version for an interface

Operation	Command	Remarks
Perform the following configuration in interface configuration mode	**ip rip version 1**	-
Set the RIP operation mode to RIP-2 multicast.	**ip rip version 2 mcast**	-
Set the RIP operation mode to RIP-2 broadcast.	**ip rip version 2 bcast**	
Delete the RIP version and uses RIP-1 by default.	**no ip rip version**	

Description:

A RIP-1 interface can send and receive RIP-1 broadcast packets. A RIP-2 broadcast interface can receive RIP-1 packets and RIP-2 broadcast packets but not RIP-2 multicast packets. A RIP-2 multicast interface can send and receive RIP-2 multicast packets.

24.2.5 Enable the Host Route Function

The RIP packets received by a route may sometimes contain host route entries, which are not conducive to routing and addressing but occupy a great amount of network resource. This function is designed to determine whether a Switch receives the host route entries in RIP packets.

Enable the Host Route Function

Operation	Command	Remarks
Enter RIP mode	**router rip**	
Receive host routing	**host-route**	-
Not receive host routing	**no host-route**	-

24.2.6 Enable the Route Aggregation Function

Route aggregation consolidates the routes on different subnets of a natural network segment into one route with a natural mask and sends the route to another network segment. This function minimizes both the number of entries in a routing table and the amount of information that needs to be exchanged.

RIP-1 sends only the routes with a natural mask, that is, aggregate routes. RIP-2 supports the subnet mask. To broadcast all the subnet routes, you should disable the route aggregation function of RIP-2.

Enable the Route Aggregation Function

Operation	Command	Remarks
Enter RIP mode	**router rip**	
Perform routing aggregation dynamically	**auto-summary**	-
Do not perform routing aggregation	**no auto-summary**	-

24.2.7 Configure RIP Packet Authentication

RIP-1 does not support packet authentication. A RIP-2 interface, however, can be configured with packet authentication in plaintext or MD5.

Configure RIP Packet Authentication

Operation	Command	Remarks
Configure plaintext authentication under the interface configuration mode	**ip rip authentication simple** *password*	-
Configure MD5 authentication under the interface configuration mode	**ip rip authentication md5 key-id** *key-id* **key-string** *key-string*	
Disable authentication function	**no ip rip authentication**	-

24.2.8 Configure Split Horizon

Split horizon is designed to prevent the routes learned on an interface from being sent through the interface, which avoids routing loops. This function must be disabled in some special situations to ensure correct route advertisement at the cost of advertisement efficiency. By default, split horizon can be enabled on an interface.

Configure Split Horizon

Operation	Command	Remarks
Enter interface configuration mode	**Interface { vlan-interface *vid* | supervlan-interface *id}***	

Enable split horizon	**ip rip split**	-
Disable split horizon	**no ip rip split**	-

24.2.9 Set an Additional Routing Metric

The additional can routing metric value is added to RIP routes on an inbound or outbound interface. It does not change the routing metric value of routes in the routing table but adds a designated metric value to the routes to be sent or received by an interface.

Set an Additional Routing Metric

Operation	Command	Remarks
Configure an additional routing metric value for routes when receives the RIP packets	**ip rip metricin** *value*	-
Forbid an additional routing metric value for routes when receives the RIP packets	**no ip rip metricin**	-
Configure an additional routing metric value for routes in the RIP packets to be sent.	**ip rip metricout *value***	
Forbid the interface to set an additional routing metric value for routes in the RIP packets to be sent.	**no ip rip metricout**	

24.2.10 Define a Prefix List

A prefix list is identified by a prefix list name, and may contain multiple entries, each of which corresponds to a network prefix identified by a sequence number. The sequence number indicates the matching sequence of a network prefix.

During prefix matching, the Switch checks the entries in ascending order of sequence numbers. If an entry is matched, it is permitted by the current prefix list and will not be matched next time.

Note: By default, if more than one prefix list entry has been defined, at least one permit

entry should be available. The deny entries can be defined in advance so that the routes that do not meet the condition are filtered quickly. However, if all the entries are prefixed by deny, no route will be permitted by the address prefix list. You are advised to define an entry permit 0.0.0.0/0 after defining multiple deny entries, so that all the routes meeting the condition are permitted.

Alternatively, you can run the ip prefix-list default command to change the default configuration. For details, see the description of this command in a command line manual.

Define a Prefix List

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Configure prefix-lis	**ip prefix-list** *name* **seq** *seq_number* **{permit | deny}** *ip-address Ip prefix* **ge** *Ip prefix* **le** *Ip prefix*	-
Delete specified prefix-lis	**no ip prefix-list** *name* **{permit | deny}**	-
Delete all prefix-lis	**no ip prefix-list all**	
Configure the matching mode when there is no prefix list or no matching item	**ip prefix-list default{ [entry-rule | tab-rule][permit | deny]}**	
Configure the matching mode to be default value when there is no prefix list or no matching item	**no ip prefix-list default [entry-rule | tab-rule]**	

24.2.11 Configure Route Redistribution

Routes of protocols other than RIP can be imported into RIP.

In an Ethernet Switch, connected, static, and OSPF routes can be imported into RIP.

Configure Route Redistribution

Operation	Command	Remarks
Enter RIP mode	router rip	
Perform the following configuration under RIP mode	**redistribute {** connected | static | ospf } metric *metric* prefix-list *list*	-
Perform the following configuration under RIP mode	**no redistribute {** connected | static | ospf }	Delete route redistribution

24.2.12 Configure External Route Aggregation

This configuration can be able to aggregate the external route which is imported by route redistribution.

Configure external route aggregation

Operation	Command	Remarks
Enter RIP mode	router rip	
Perform the following configuration under RIP mode	**summary-address** *address mask*	-
Perform the following configuration under RIP mode	**no summary-address** *address mask*	Delete route aggregation

24.2.13 Display RIP Configurations

Display RIP Configurations

Operation	Command	Remarks
Display the RIP packet statistics information.	**show ip rip**	-
Display the RIP interface configuration, such as the version and authentication information.	**show ip rip interface**	-
Display RIP routing tables.	**show ip route rip**	

24.3 Examples

24.3.1 Configuration Examples

！ To configure RIP to deny host routes, run the following command:
Switch(config-router-rip)#no host-route

！To configure plaintext authentication on VLAN interface 3 and set keyword to be Switch, run the following command:

Switch(config-if-vlanInterface-3)#ip rip authentication simple Switch

！ To forbid VLAN interface 3 to receive RIP packets, run the following command:
Switch(config-if-vlanInterface-3)#no ip rip input

！To set the additional routing metric value to 1 for RIP packets received by VLAN interface 3, run the following command:
Switch(config-if-vlanInterface-3)#ip rip metricin 1

！To set the additional routing metric value to 1 for RIP packets sent by VLAN interface 3, run the following command:
Switch(config-if-vlanInterface-3)#ip rip metricout 1

！To forbid VLAN interface 3 to send RIP packets, run the following command:
Switch(config-if-vlanInterface-3)#no ip rip output

！To enable split horizon on VLAN interface 3 towards RIP packet sending, run the following command:
Switch(config-if-vlanInterface-3)#ip rip split

！To configure VLAN interface 3 to run RIP-2 multicast, run the following command:
Switch(config-if-vlanInterface-3)#ip rip version 2 mcast

！To allow VLAN interface 3 to send and receive RIP packets, run the following command:
Switch(config-if-vlanInterface-3)#ip rip work

！To specify that RIP runs on the network segment 192.1.1.1/24, run the following command:
Switch(config-router-rip)#network 192.1.1.1

！To enable RIP, run the following command:
Switch(config)#router rip

！To disable RIP, run the following command:
Switch(config)#no router rip

！To deny all the routes (including subnet routes) destined for 192.168.1.0/24 by configuring the prefix list, run the following command:
Switch(config)#ip prefix-list pflst001 deny 192.168.1.0 24

Switch(config)#ip prefix-list pflst001 permit 0.0.0.0 0

！To configure the matching mode to permit if no matching entry exists in the prefix list, run the following command:
Switch(config)#ip prefix-list default entry-rule permit

！To display the entire prefix lists currently available, run the following command:
Switch(config)#show ip prefix-list

！To import OSPF routes to RIP, run the following command:
Switch(config-router-rip)#redistribute ospf

！To display the RIP statistics on L3 interfaces, run the following command:
Switch(config)#show ip rip

！To display the RIP configuration of L3 interfaces, run the following command:
Switch(config-router-rip)#show ip rip interface

！To display the RIP configuration of L3 interface 1, run the following command:
Switch(config-router-rip)#show ip rip interface vlan-interface 1

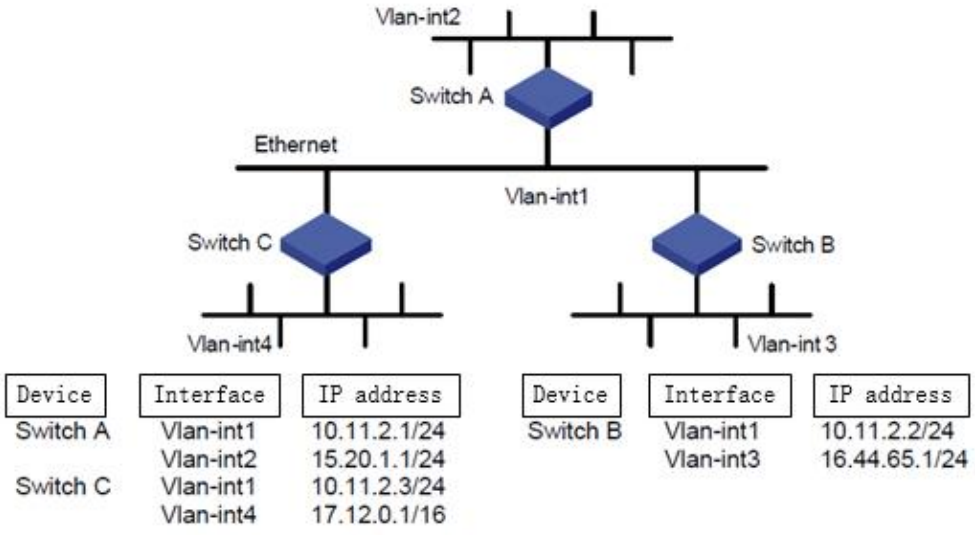
24.3.2 Application Examples

1. Network requirements

As shown in the following figure: SwitchC connects with subnet 17.12.0.0 via Ethernet interface. The Ethernet interface of SwitchA and SwitchB connect with network 15.20.1.0 and 16.44.65.0 respectively. SwitchC connects with SwitchA and SwitchB via Ethernet 10.11.2.0.

Configure RIP router protocol correctly to ensure the network that SwitchC connects with SwitchA and SwitchB can be able to interwork.

Network diagram



Network diagram of RIP typical configuration

3. Configuration steps

Note:

The following configuration only lists the operations which relate to RIP. Please ensure that the Ethernet link layer can work normally and the IP of each VLAN interface has been configured before performing the following configuration.

(1) Configure SwitchA

Switch#

Switch#configure terminal

Switch(config)#router rip

Switch(config-router-rip)#network 10.11.2.1

Switch(config-router-rip)#network 15.20.1.1

Switch(config-router-rip)#

(2) Configure SwitchB

Switch#

Switch#configure terminal

Switch(config)#router rip

Switch(config-router-rip)#network 10.11.2.2

Switch(config-router-rip)#network 16.44.65.1

Switch(config-router-rip)#

(3) Configure SwitchC

Switch#

Switch#configure terminal

Switch(config)#router rip

Switch(config-router-rip)#network 10.11.2.3

Switch(config-router-rip)#network 17.12.0.1

Switch(config-router-rip)#

25.OSPF Configuration

25.1 OSPF Overview

Open Shortest Path First (OSPF) is an interior routing protocol, which is developed by IETF based on the link state detection and shortest path first technologies. In an IP network, OSPF dynamically discovers and advertise routes by collecting and transmitting the link states of autonomous systems (ASs). It supports interface-based packet authentication for purposes of route calculation security and employs IP multicast to send and receive packets.

Each OSPF router maintains a database that describes the topological structure of an AS. The database is a collection of link-state advertisements (LSAs) of all the routers. Every router always broadcasts the local state information across the entire AS. If two or more routers exist in a multi-access network, a designated router (DR) and a backup designated router (BDR) must be elected. The DR is responsible for broadcasting the LSAs of the network. With a DR, a multi-address access network may require less neighbor relationships to be established between routers. OSPF allows an AS to be divided into areas, between which routing information is further abstracted. As a result, smaller network bandwidth will be occupied.

OSPF uses four types of routes, which are listed in order of priority as follows:

Intra-area routes

Inter-area routes

Type 1 external routes

Type 2 external routes

Intra-area and inter-area routes describe the network structure of an AS, while external routes depict how routes are distributed to destinations outside an AS. Generally, type 1 external routes are based on the information imported by OSPF from other interior routing protocols and comparable to OSPF routes in routing cost; type 2 external routes are based on the information imported by OSPF from exterior routing protocols and the costs of such routes are far greater than those of OSPF routes. Therefore, route calculation only takes the external costs into consideration.

Based on the link state database (LSDB), each router builds a shortest path tree with itself as the root, which presents the routes to every node in an AS. An external route emerges as a leaf node and can also be marked by the router that broadcasts the external route so that additional information about an AS is recorded.

All the OSPF areas are connected to the backbone area, which is identified by 0.0.0.0. OSPF areas must be logically continuous. To achieve this end, virtual connection is introduced to the backbone area to ensure the logical connectivity of areas even if they are physically separated.

All the routers in an area must accept the parameter settings of the area. Therefore, the configuration of routers in the same area must be performed in consideration of the parameter settings of the area. A configuration error may lead to the failure of information transfer between adjacent routers and even routing failures or routing loops.

25.2 OSPF Detailed Configuration

25.2.1 Enable/Disable OSPF Configuration

OSPF basic configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	
Enable ospf	**router ospf**	-
Disable ospf	**no router ospf**	-

25.2.2 Configure the ID of a Router

The router ID is a 32-bit unsigned integer and the only one label of a router in AS. Therefore, user should configure the router ID. When configure the router ID manually, you should ensure that any two router's IDs in AS are different. Generally, the router ID is in line with the IP address of this router's certain interface.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Configure router id	**router id** *router-id*	-

When performing the network planning, you should choose the router ID and manually configure it so as to ensure the stability of OSPF.

The smallest IP address in the interface IP will be selected to be the router ID by default.

25.2.3 Specify an Interface and Area ID

OSPF divides an AS into different areas, based on which routers are logically classified into different groups. Area border routers (ABRs) may belong to different areas. A network segment belongs to only one area, that is, the homing area of an OSPF interface must be specified. An area is identified by an area ID. Routes between areas are transmitted by ABRs.

In addition, all the routers in an area must unanimously accept the parameter settings of the area. Therefore, the configuration of routers in the same area must be performed in consideration of the parameter settings of the area. A configuration error may lead to the failure of information transfer between adjacent routers and even routing failures or routing loops.

OSPF related parameter configuration

Operation	Command	Remarks
Run the command in OSPF configuration mode.	**router ospf**	
Specify area for IP interface	**network address wildcard-mask area area-id**	
Cancel to specify area for IP	**no network address wildcard-mask area area-id**	

interface		
Configure the authentication type for an area.	**area *area-id* authentication [message-digest]**	
Restore the authentication type of an interface to no authentication.	**no area *area-id* authentication**	

25.2.4 Configure the Authentication Type for an Area

When in the same area, the authentication types of all routers should be the same. (MD5 authentication and plaintext authentication are supported.)

Operation	Command	Remarks
Run the command in OSPF configuration mode.	**router ospf**	
Configure the authentication type for an area.	**area *area-id* authentication [message-digest]**	
Restore the authentication type of an interface to no authentication.	**no area *area-id* authentication**	

25.2.5 Set a Password for Packet Authentication

OSPF supports MD5 authentication or plaintext authentication between two adjacent routers.

Operation	Command	Remarks
Enter interface configuration mode	interface {**vlan-interface** *vid* | **supervlan-interface** *id* }	
Set a password for plaintext authentication.	**ip ospf authentication-key *password***	
Disable plaintext authentication.	**no ip ospf authentication-key**	
Set a password for MD5 authentication.	**ip ospf message-digest-key *key-id* md5 *key***	
Disable MD5 authentication.	**no ip ospf message-digest-key**	

25.2.6 Configure OSPF Interface Type

OSPF calculates routes based on the topological structure of the network adjacent to the local router. Each router describes the topology of its adjacent network and transmits it to the other routers. According to the link layer protocol, OSPF classifies networks into the following four types:

Broadcast networks: When Ethernet or FDDI is used as the link layer protocol, OSPF considers that the network type is broadcast by default.

Non Broadcast MultiAccess (NBMA) networks: When ATM is used as the link layer protocol, OSPF considers that the network type is NBMA by default.

Point-to-Multipoint networks: This network type will be considered as default in no case. It is always a substitute of other network types through forcible change. An NBMA network that is not fully meshed is often changed to a point-to-multipoint network.

Point-to-Point networks: When PPP, LAPB, or POS is used as the link layer protocol, OSPF considers that the network type is Point-to-Point by default.

The ATM network is a typical NBMA network. A polling interval can be configured to specify the interval of sending Hello packets before a router establishes a neighbor relationship with its neighboring router.

On a broadcast network incapable of multi-address access, you can configure the interface type to nonbroadcast.

If some routers are not directly reachable on an NBMA network, you can configure the interface type to point-to-multipoint.

If a router has only one peer router on an NBMA network, you can set the interface type to point-to-point.

The differences between an NBMA network and a point-to-multipoint network are as

follows:

In OSPF, an NBMA network refers to a non-broadcast multi-access network that is fully meshed. A point-to-multipoint network may not be fully meshed.

A DR and a BDR must be elected on an NBMA network but are not involved on a point-to-multipoint network.

NBMA is a default network type. For example, if the link layer protocol is ATM, OSPF considers that the network type is NBMA by default no matter whether the network is fully meshed. Point-to-multipoint is not a default network type. No link layer protocol is viewed as a point-to-multipoint protocol. You can use this network type through a forcible change. An NBMA network that is not fully meshed is often changed to a point-to-multipoint network.

On an NBMA network, packets are transmitted in unicast mode, which requires you to configure neighbor relationship manually. On a point-to-multipoint network, packets are transmitted in multicast mode.

An Ethernet Switch uses Ethernet as the link layer protocol, so OSPF regards that the network type is broadcast. Do not change the network type of an Ethernet Switch at discretion.

Related configuration of OSPF interface

Operation	Command	Remarks
Enter interface configuration mode	interface {**vlan-interface** *vid* | **supervlan-interface** *id* }	
Set the network type of an interface.	**ip ospf network { broadcast | non-broadcast | point-to-multipoint | point-to-point }**	-
Restore the network type of an interface to the default value.	**no ip ospf network**	-
Set the cost of sending packets through a VLAN interface.	**ip ospf cost *cost***	
Restore the packet sending cost of a VLAN interface to the default value.	**no ip ospf cost**	
Set the priority of an interface in DR election.	**ip ospf priority *value***	
Restore the default priority of an interface.	**no ip ospf priority**	
Set the interval of sending Hello packets for an interface.	**ip ospf hello-interval *seconds***	

Restore the interval of sending Hello packets for an interface to the default value.	**no ip ospf hello-interval**	
Set the timeout time of the neighboring router.	**ip ospf dead-interval *seconds***	
Restore the timeout time of the neighboring router to the default value.	**no ip ospf dead-interval**	
Set the interval of LSA retransmission between two adjacent routers.	**ip ospf retransmit-interval *seconds***	
Restore the interval of LSA retransmission between two adjacent routers to the default value.	**no ip ospf retransmit-interval**	
Set the time for sending a link state update packet.	**ip ospf transmit-delay *seconds***	
Restore the time for sending a link state update packet to the default value.	**no ip ospf transmit-delay**	
Set a password for plaintext authentication.	**ip ospf authentication-key *password***	
Disable plaintext authentication.	**no ip ospf authentication-key**	
Set a password for MD5 authentication.	**ip ospf message-digest-key *key-id* md5 *key***	
Disable MD5 authentication.	**no ip ospf message-digest-key**	

25.2.7 OSPF Area Related Configuration

A stub area is a special LSA area in which ABRs do not distribute the external routes they have received. In stub areas, both the size of routing tables and the amount of the routing information are drastically reduced.

Any area that meets certain conditions can be configured into a stub area. Generally, a stub area is located at the border of an AS. It may be a non-backbone area with only one ABR or a non-backbone area with multiple ABRs between which no virtual connection is configured.

To make a stub area reachable for other ASs, the ABR in the stub area generates a default route (0.0.0.0) and advertises it to non-ABR routers in this area.

When configuring a stub area, note the following points:

A backbone area cannot be a stub area and a virtual connection is not allowed in a stub area.

All the routers in a stub area must be configured to indicate that they are located in a stub area.

No ASBR is allowed in a stub area, that is, routes from outside the AS where the stub area resides cannot be advertised within the stub area.

OSPF area related configuration

Operation	Command	Remarks
Enter OSPF protocol configuration mode	**router ospf**	
Configure a stub area.	**area *area-id* stub [no-summary]**	-
Cancel the stub area configuration.	**no area *area-id* stub**	-
Configure the cost of the default route to a stub area.	**area *area-id* default-cost *cost***	
Cancel the cost configuration for the default route to a stub area.	**no area *area-id* default-cost**	
Configure an NSSA area.	**area *area-id* nssa [no-summary]**	
Cancel the NSSA area configuration.	**no area *area-id* nssa**	
Configure the cost of the default route to an NSSA area.	**area *area-id* default-cost *cost***	
Cancel the cost configuration for the default route to an NSSA area.	**no area *area-id* default-cost**	
Configure route aggregation in an OSPF area.	**area *area-id* range *address mask* [advertise | notadvertise]**	
Remove route aggregation in an OSPF area.	**no area area-id range address mask**	
Configure external router aggregation and broadcast this aggregation.	**summary-address** *address mask* **tag** *value*	
Configure external router	**summary-address** *address mask* **notadvertise tag** *value*	

aggregation but not broadcast this aggregation.		
Cancel external router aggregation	**no summary-address** *address mask*	
Create and configures a virtual connection.	**area** ***area-id*** **virtual-link** ***router-id*** **[{ hello-interval** ***seconds*** **| retransmit- interval** ***seconds*** **| transmit-delay** ***seconds*** **| dead-interval** ***seconds*** **| { authentication-key** ***password*** **| message-digest-key** ***keyid*** **md5** ***key*** **} } *]**	
Cancel a virtual connection.	**no area** ***area-id*** **virtual-link** ***router-id***	
Import routes of other protocols into OSPF.	**redistribute** ***protocol*** **[metric** ***metric*** **] [type { 1 | 2 }] [tag** ***tag-value*** **][prefix-list** ***prefix-list-name*****]**	
Disable the import of routes of other protocols into OSPF.	**no redistribute** ***protocol***	
Import the default route to OSPF.	**default-information originate [always] [metric** ***metric-value*** **] [type** ***type-value*** **]**	
Disable the import of the default route.	**no default-information originate**	
Configures a default metric value for reception of external routes.	**default redistribute metric** ***metric***	
Cancel the default metric value configuration for reception of external routes.	**no default redistribute metric**	
Configure the default type of external routes to be received.	**default redistribute type { 1 | 2 }**	
Cancel the default type configuration for the external routes to be received.	**no default redistribute type**	By default, the metric value is 1 and type is 2 for the external routes to be received by an OSPF router.
Filter the learned routes.	**ip ospf distribute-list prefix-list** ***prefix-list-name*** **in** **no ip ospf distribute-list prefix-list in**	interface configuration mode
Filter the advertised routes.	**ip ospf distribute-list prefix-list** ***prefix-list-name*** **out** **no ip ospf distribute-list prefix-list out**	interface configuration mode
Receive the routes from a specified neighboring Ethernet router.	**ip ospf distribute-list gateway** ***prefix-list-name*** **in** **no ip ospf distribute-list gateway in**	interface configuration mode
Enable BFD for link state	**ip ospf bfd**	interface

monitoring.		configuration mode
Disable BFD.	**no ip ospf bfd**	interface configuration mode

25.3 Configuration Example

！ To configure MD5 authentication for OSPF area 1, run the following command:

Switch(config-router-ospf)#area 0.0.0.1 authentication message-digest

！ To configure the cost of the default route 192.168.0.100 to 10, run the following command:

Switch(config-router-ospf)#area 192.168.0.100 default-cost 10

！ To aggregate the routes 202.38.160.0/24 and 202.38.180.0/24 into a route 202.38.0.0/16, run the following command:

Switch(config-router-ospf)#network 202.38.160.3 0.0.0.255 area 1.1.1.1

Switch(config-router-ospf)#network 202.38.180.3 0.0.0.255 area 1.1.1.1

Switch(config-router-ospf)#area 1.1.1.1 range 202.38.0.0 0.0.255.255

！ To configure area 1.1.1.1 as a stub area, run the following command:

Switch(config-router-ospf)#area 1.1.1.1 stub

！ To configure area 1.1.1.1 as an NSSA area, run the following command:

Switch(config-router-ospf)#area 1.1.1.1 nssa

！ To configure a virtual connection, for which the transit area is 1.1.1.1 and router-id of the peer router is 10.11.5.2, run the following command:

Switch(config-router-ospf)#area 1.1.1.1 virtual-link 10.11.5.2

！To configure that an ASE LSA is generated for a default route if any, run the following command:

Switch(config-router-ospf)#default-information originate

！ To configure that an ASE LSA is generated for the default route and advertised to OSPF routing domains even if no default route exists, run the following command:

Switch(config-router-ospf)#default-information originate always

！ To set the default routing metric value of an imported external route to 10, run the following command:

Switch(config-router-ospf)#default redistribute metric 10

！ To configure that OSPF imports type 1 external routes by default, run the following command:

Switch(config-router-ospf)#default redistribute type 1

！ To set the password for plaintext authentication on VLAN interface 3 to abc123, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf authentication-key abc123

！ To set the cost of running OSPF on VLAN interface 3 to 10, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf cost 10

！ To set the timeout time of the neighboring router on VLAN interface 3 to 60s, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf dead-interval 60

！ To set the interval of transmitting OSPF Hello packets for VLAN interface 3 to 15s, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf hello-interval 15

！ To set the password for MD5 authentication on VLAN interface 3 to abc123, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf message-digest-key 12 md5 abc123

！ To configure VLAN interface 2 as a non-broadcast interface, run the following command:

Switch(config-if-vlanInterface-2)#ip ospf network non-broadcast

！ To set the priority of VLAN interface 3 to 100, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf priority 100

！ To set the interval of retransmitting LSAs between VLAN interface 3 and its adjacent routers to 8s, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf retransmit-interval 8

！ To set the LSA transmission delay of VLAN interface 3 to 3s, run the following command:

Switch(config-if-vlanInterface-3)#ip ospf transmit-delay 3

！ To run OSPF on the interfaces whose primary IP address is 192.168.0.100 and inverse mask is 0.0.0.255 and set the ID of the area where these interfaces reside to 1.1.1.1, run the following command:

Switch(config-router-ospf)#network 192.168.0.100 0.0.0.255 area 1.1.1.1

！ To import RIP routes into OSPF, run the following command:

Switch(config-router-ospf)#redistribute rip

！ To use address prefix lists to match the routes learned by VLAN interface 2, run the following command:

Switch(config-if-vlanInterface-2)#ip ospf distribute-list prefix-list check in

！ To enable OSPF BFD on VLAN interface 2, run the following command:

Switch(config-if-vlanInterface-2)#ip ospf bfd

！ To set the router ID of a switch to 192.168.0.100, run the following command:

Switch(config)#router id 192.168.0.100

！ To enable OSPF on a switch, run the following command:

Switch(config)#router ospf

！ To disable OSPF on a switch, run the following command:

Switch(config)#no router ospf

！ To display OSPF information, run the following command:

Switch(config)#show ip ospf

！ To display the information of OSPF border routers, run the following command:

Switch(config)#show ip ospf border-routers

Switch(config-if-vlanInterface-2)#show ip ospf cumulative

！ To display the information of OSPF LSDBs, run the following command:

Switch(config)#show ip ospf database

！ To display the information of OSPF errors, run the following command:

Switch(config-if-vlanInterface-2)#show ip ospf error

！ To display the information of OSPF interfaces, run the following command:

Switch(config)#show ip ospf interface

！ To display the information of all the OSPF neighbors, run the following command:

Switch(config)#show ip ospf neighbor

！ To display the information of an OSPF request list, run the following command:

Switch(config)#show ip ospf request-list

！ To display the information of an OSPF retransmission list, run the following command:

Switch(config)#show ip ospf retrans-list

！ To display the information of an OSPF virtual connection, run the following command:

Switch(config)#show ip ospf virtual-link

！ To display the router ID, run the following command:

Switch(config)#show router id

！ To display the configured prefix lists, run the following command:

Switch(config)#show ip ospf distribute-link

25.4 Application Example

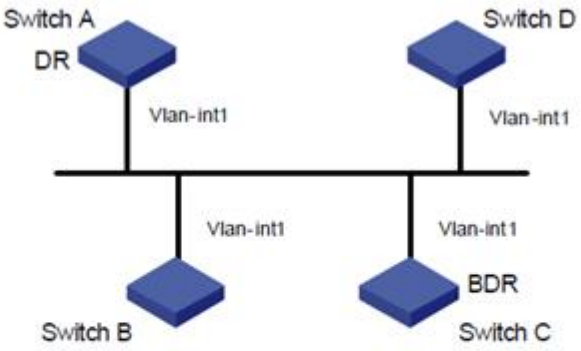
A) Configure the DR selection of OSPF priority

1.Network requirements

SwitchA, SwitchB, SwitchC and SwitchD run OSPF protocol in the same network segment, as shown below. Make sure to perform the right configuration to ensure SwitchA to be DR, SwitchC to be BDR (SwitchA possesses the highest priority of 100, so it is selected as DR; SwitchC

possesses the second highest priority of 2, so it is selected as BDR; the priority of SwitchB is 0, so it cannot be DR; SwitchD has no priority, so its value is 1 by default).

2. Network diagram



Device	Interface	IP address	Router ID	Priority
Switch A	Vlan-int1	192.168.1.1/24	1.1.1.1	100
Switch B	Vlan-int1	192.168.1.2/24	2.2.2.2	0
Switch C	Vlan-int1	192.168.1.3/24	3.3.3.3	2
Switch D	Vlan-int1	192.168.1.4/24	4.4.4.4	1

Network diagram for DR selection of OSPF priority

3. Configuration steps

(1). Configure SwitchA

Switch#configure terminal

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 192.168.1.1 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#router id 1.1.1.1

Switch(config)#router ospf

Switch(config-router-ospf)#network 192.168.1.1 0.0.0.255 area 0

Switch(config-router-ospf)#exit

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip ospf priority 100

Switch(config-if-vlanInterface-1)#

(2). Configure SwitchB

Switch#configure terminal

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 192.168.1.2 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#router id 2.2.2.2

Switch(config)#router ospf

Switch(config-router-ospf)#network 192.168.1.2 0.0.0.255 area 0

Switch(config-router-ospf)#exit

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip ospf priority 0

Switch(config-if-vlanInterface-1)#

(3). Configure SwitchC

Switch#configure terminal

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 192.168.1.3 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#router id 3.3.3.3

Switch(config)#router ospf

Switch(config-router-ospf)#network 192.168.1.3 0.0.0.255 area 0

Switch(config-router-ospf)#exit

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip ospf priority 2

Switch(config-if-vlanInterface-1)#

(4). Configure SwitchD

Switch#configure terminal

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 192.168.1.4 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#router id 4.4.4.4

Switch(config)#router ospf

Switch(config-router-ospf)#network 192.168.1.4 0.0.0.255 area 0

Switch(config-router-ospf)#exit

Switch(config)#

Run the command of show ip ospf neighbor in SwitchA to display OSPF neighbors. If the entire neighbor's state is "full", it means that SwitchA has adjoined with all its neighbors. SwitchA is DR, whereas SwitchC is BDR and all the other neighbors are DRother (DRother means that they are neither DR or BDR.)

Change the priority of SwitchB to be 200

Switch#configure terminal

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip ospf priority 0

Run the command of show ip ospf neighbor in SwitchA to display OSPF neighbors. The priority of SwitchB will become 200, but it is not the DR. Turn off SwitchA, and then run the command of show ip ospf neighbor in SwitchD to display OSPF neighbors. You will find SwitchC turns into DR while SwitchB becomes BDR.

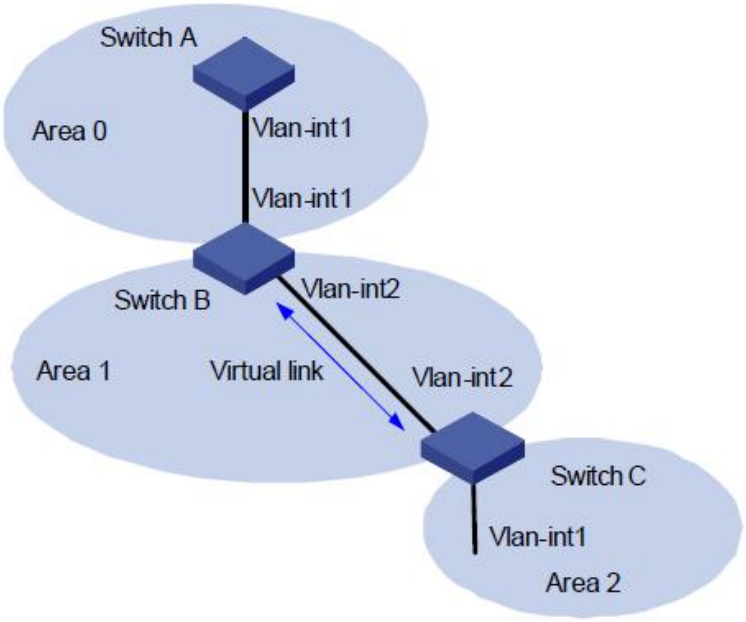
If all switches are moved away firstly and then added again, SwitchB will be selected to be DR with the priority of 200, and SwitchA will be selected to be BDR with the priority of 100. If you turn off all switches and then reboot the switches, it will have a new selection of DR/BDR.

B) Configure OSPF virtual connection

1. Network requirements

There is no direct connection between area 2 and area 0. Area 1 is asked to act as a transportation zone to connect area 2 with area 0. Therefore, there should be a virtual link between SwitchB and SwitchC in area 1.

2. Network diagram



设备	接口	IP 地址	Router ID
Switch A	Vlan-int1	192.168.1.1/24	1.1.1.1
Switch B	Vlan-int1	192.168.1.2/24	2.2.2.2
	Vlan-int2	193.168.1.2/24	
Switch C	Vlan-int1	155.100.1.1/24	3.3.3.3
	Vlan-int2	193.168.1.1/24	

Network diagram of OSPF virtual link

3. Configuration steps

(1). Configure SwitchA

Switch#configure terminal

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 192.168.1.1 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#router id 1.1.1.1

Switch(config)#router ospf

Switch(config-router-ospf)#network 192.168.1.1 0.0.0.255 area 0

Switch(config-router-ospf)#exit

Switch(config)#

(2) Configure SwitchB

Switch#configure terminal

Switch(config)#vlan 2

Switch(config-if-vlan)#exit

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 192.168.1.2 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#interface vlan-interface 2

Switch(config-if-vlanInterface-2)#ip address 193.168.1.2 255.255.255.0

Switch(config-if-vlanInterface-2)#exit

Switch(config)#router id 2.2.2.2

Switch(config)#router ospf

Switch(config-router-ospf)#network 192.168.1.2 0.0.0.255 area 0

Switch(config-router-ospf)#network 193.168.1.2 0.0.0.255 area 1

Switch(config-router-ospf)#area 1 virtual-link 3.3.3.3

Switch(config-router-ospf)#exit

Switch(config)#

(3) Configure SwitchC

Switch#configure terminal

Switch(config)#vlan 2

Switch(config-if-vlan)#exit

Switch(config)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 155.100.1.1 255.255.255.0

Switch(config-if-vlanInterface-1)#exit

Switch(config)#interface vlan-interface 2

Switch(config-if-vlanInterface-2)#ip address 193.168.1.1 255.255.255.0

Switch(config-if-vlanInterface-2)#exit

Switch(config)#router id 3.3.3.3

Switch(config)#router ospf

Switch(config-router-ospf)#network 155.100.1.1 0.0.0.255 area 2

Switch(config-router-ospf)#network 193.168.1.1 0.0.0.255 area 1

Switch(config-router-ospf)#area 1 virtual-link 2.2.2.2

Switch(config-router-ospf)#exit

Switch(config)#

26.BGP Configuration

26.1 BGP Overview

Border Gateway Protocol (BGP) is a kind of dynamic routing protocol between autonomous systems. Its basic function is to automatically exchange loop-free routing information between autonomous systems. By exchanging reachable information with the attributes of autonomous system (AS), BGP constructs the topology of the autonomous system.

BGP Protocol

BGP has the following features:

As an exterior gateway protocol, BGP focuses on the control of route distribution and select the best route among the ASs.

BGP adopts TCP as its transport layer protocol (Listening port number is 179), improving the protocol reliability.

BGP performing the routing selection in inter-domain, so it asks for high protocol stability. Therefore, it adopts TCP protocol to ensure the stability of BGP protocol.

BGP peers must be logically connected, and they should connect with TCP. Port number is 179 and the local port number is optional.

BGP supports CIDR（Classless Inter-Domain Routing）.

When BGP routes are updated, BGP only sends updated routes, which greatly reduces the bandwidth occupied by BGP routes. It is suitable for spreading a large amount of routing information on the Internet.

BGP is a kind of Distance-Vector routing protocol.

The design of BGP avoids loop circuit:

Inter-AS: BGP identifies the AS with AS -path information. The routing with the local AS number will be discarded so as to avoid loop circuit.

AS interior: The routing which BGP learned from AS interior no longer inform BGP neighbor who in AS interior so as to avoid loop circuit.

BGP provides various routing strategies, implementing flexible filtering and choices for routing.

BGP provides prevent routing oscillation mechanism, effectively improving the stability of the Internet.

BGP is apt to expand and it adapts to the new development network.

BGP operation mode

BGP runs on a router in either of the following modes:

IBGP（Internal BGP）

EBGP（External BGP）

BGP is regarded as IBGP when deployed within an AS and as EBGP when deployed between

ASs.

BGP roles:

Speaker: The router sending a BGP message is called the BGP speaker, which constantly receives or generates new routing information and advertises it to other BGP speakers. After receiving a new route advertisement from another AS, the BGP speaker distributes the route advertisement to all the other BGP speakers in the same AS if the route is better than the current one or has not been received ever.

Peer: If two BGP speakers are exchanging messages, they call each other the peer. Some peers may constitute peers group.

BGP message:

BGP running is driven by messages, which are classified as Open, Update, Notification, Keepalive and Route-Refresh.

Open message:

An Open message is the first message to be sent after setup of a TCP connection and used to establish a BGP peer relationship. BGP peer will sent Keepalive to confirm and save the connection effectiveness after receiving the Open message and performing successful negotiation. After the confirmation, BGP peer can be able to perform the exchange among the Update message, Notification message, Keepalive message and Route-Refresh message.

Update message:

An Update message is transmitted between BGP peers for routing information exchange. Update message can release multiple reachable routing information of same attribute as well as repeal multiple unreachable routings information.

a) An Update message can release multiple reachable routings of same attribute, and those routings can be able to share one set of routing attribute. All the routing attributes included in a given Update message is suitable for all destinations (With IP prefix represent) of the Network Layer Reachability Information field.

b) An Update message can revoke multiple unreachable routing. Each routing destination clearly defines the routing which is notified by BGP Speaker.

c) An Update message can be only used to revoke routing. In this case, it is unnecessary to contain path attributes or NLRI (Network Layer Reachability Information). On the contrary, an Update message can be only used to notify reachable routing. In this case, it is unnecessary to carry withdrawn routing information.

Notification message:

BGP will forward Notification message to peer if BGP detected error status, and then BGP connection will be interrupted immediately.

Keepalive message:

BGP will periodically forward Keepalive message to the peer in order to keep the connection validity.

Route-Refresh message:

Route-Refresh message is used to notify the peer what the Route-Refresh capabilities it supports. Under the situation that all BGP enable Route-Refresh capability, the local BGP equipment will release Route-Refresh to the peer if the routing strategies in the BGP entrance have changed. The peer will resend the routing message to local BGP equipment after receiving this message. In this case, it can perform dynamic refresh on BGP routing and apply new routing strategies under the situation without interrupting BGP connection.

BGP state machine:

There are six state in BGP finite state machine: Idle、Connect、Active、OpenSent、OpenConfirm and Established.

1. In Idle state, BGP is in the initial state and refuse to admit any connection request.

When BGP receives the start event, BGP launches the TCP connection with the peer, starting to connect ConnectRetry Timer, monitoring the TCP message from the peer and turning to Connect state.

2. Under Connect state, BGP decides the later operations after establishing the TCP connection.

If TCP connection is established successfully, BGP will stop connecting ConnectRetry Timer and then forward an Open message to the peer. In addition, it will turn to Opensent state.

If the TCP connection fails, BGP will reset ConnectRetry Timer, monitoring the TCP connection which sponsored by the peer, and it will turn to Active state.

If ConnectRetry Timer times out, BGP will restart ConnectRetry Timer and try to establish TCP connection with the peer. This moment, BGP will keep the Connect state.

3. Under Active state, BGP will try to establish TCP connection.

If TCP connection is established successfully, BGP will reset ConnectRetry Timer and then forward an Open message to the peer. Moreover, it will turn to Opensent state.

If ConnectRetry Timer times out, BGP will restart ConnectRetry Timer, and turn to Connect state.

If BGP try to establish TCP session with a unknown IP address, TCP connection will fail and ConnectRetry Timer will be reset. Moreover, BGP will keep the Active state.

4.OpenSent: under this state, BGP has already forwarded an Open message to the peer and now is waiting for peer's Open message.

If BGP receives the correct Open message, it will turn to OpenConfirm state.

If BGP receives the wrong Open message, it will forward a Notification message to peer and

then turn to Idle state.

If BGP receives the message of TCP connection break, BGP will reset ConnectRetry Timer, monitoring the TCP connection which peer sponsors. Moreover, BGP will turn to Active state.

5.OpenConfirm: in this state, BGP is waiting for a Notification message or Keepalive message.

If BGP receives the Notification message or the message of TCP connection break, it will turn to Idle state.

If BGP receives the Keepalive message, it will turn to Established state.

6.Established: The two sides of BGP peer can be able to exchange Update message, Notification message and Keepalive message.

If BGP receives Update message or Keepalive message, it will keep the Established state.

If BGP receives Notification message, it will turn to Idle state.

There are three common states in the process of BGP peer establishing: Idle、 Active、 Established.

Only when both sides of BGP peers BGP are in Established state can BGP neighbor relationship be established. The two sides exchange routing information through the Update message.

BGP processing procedures:

It adopts the following strategies when released BGP routing:

BGP Speaker only releases the optimal routing to peer if there exists more than one valid routing.

BGP Speaker will release the routing obtained from EBGP to all its BGP peers (including EBGP peer and IBGP peer).

BGP Speaker will not release the routing obtained from IBGP to its IBGP peer.

BGP Speaker will release the routing obtained from IBGP to its EBGP peer.

Once the connection established, BGP Speaker will release all his BGP routing to new peer.

Strategies of BGP routing release:

It asks to build the TCP connection between peers before establishing the BGP peer because BGP transport layer protocol is TCP protocol. BGP neighbor will negotiate related parameters via Open message to build up BGP peer relationship.

BGP neighbors will perform the exchange of the BGP routing table after the connection established. BGP protocol will not regularly update the routing table but it will incrementally update the routing table via Update message when the BGP routing changes.

BGP will forward Keepalive message to keep the BGP connection among the neighbor. If BGP finds that there are wrong states exists in the network (for example: receiving the error message), it will forward Notification message to report errors, and the BGP connection will be interrupted.

BGP Attribute:

BGP routing attribute is a set of parameters. It performs the further description on specified routing so as to make filtration and selection. In fact, the BGP routing attribute can be divided into 4 types as follow:

Well-known mandatory: It can be identified by all BGP devices and should be in Update message or the routing message will make errors.

Well-known discretionary: It can be identified by all BGP devices and it is not required to exist in Update message.

Optional transitive: It has the transferability attribute among the AS. BGP device is optional to support this attribute while it will receive this attribute and then transfer to other peers.

Optional non-transitive: If BGP device does not support this attribute, the corresponding attribute will be passed and it will not transfer to other peers.

Here are several common BGP routing attributes:

1) Origin attribute

Origin attribute is used to define the origin of path information, marking how a routing becomes a BGP routing. There are three types of origin attributes:

IGP: It possesses the highest priority, obtaining the routing information via IGP.

EGP: It possesses the next highest priority, obtaining routing information via EGP. EGP is its Origin attribute.

Incomplete: It possesses the lowest priority, obtaining the routing information via other ways learning.

2) AS_Path attribute

AS_Path attribute records all AS numbers that a certain routing goes through from the local address to the destination address according to vector sequence.

If BGP Speaker local notifies a routing:

When BGP Speaker notifies this routing to other AS, it will add the local AS number to AS_Path list and then notify to neighbor devices via Update message.

When BGP Speaker notifies this routing to local AS, it will create an empty AS_Path list in Update message.

If BGP Speaker propagates the routing learning from the Update message of other BGP Speaker:

When BGP Speaker notifies this routing to other AS, it will add the local AS number to the front of the AS_Path list (that is the leftmost of the AS_Path list). The BGP device which receives this routing can be able to know what the ASs it will go through during the way it goes to the destination address. The adjacent AS number who is closest to the local AS ranges ahead and the other AS numbers range in sequence.

If BGP Speaker notifies this routing to local AS, it will not change the relative attributes of this routing.

There are four types of AS_Path: AS_Sequence, AS_Set, AS_Confed_Sequence and AS_Confed_Set.

AS_Sequence: It is the ordered set recording the entire ASs which routings pass by.

AS_Set: It is the AS number of unordered set. AS_Set usually applies to the scene of routing aggregation. It has no other choice but to use AS_Set to perform unordered record due to system cannot make out an ordered set after routing aggregation. No matter how many AS numbers are included in AS_Set, BGP takes their length as 1 when selecting the routing.

AS_Confed_Sequence: It is an ordered set of sub-AS.

AS_Confed_Set: It is an unordered set of sub-AS, mainly applying to the scene of route aggregation.

The function of AS_Confed_Sequence and AS_Confed_Set is not only to avoid AS routing loops but also to perform routing selection.

3) Next_Hop attribute

The attribute of BGPNext_Hop is different from IGP. It does not always the IP address of neighbor device. Generally, Next_Hop attributes conform to the following rules:

When BGP Speaker releases a certain routing to EBGP peer, it will configure the next-hop attribute of this routing to be the interface address of the BGP neighbor relationship.

When BGP Speaker releases the local starting routing to IBGP peers, it will configure the next-hop attribute of this routing to be the interface address of the BGP neighbor relationship.

When BGP Speaker releases the routing learning from EBGP peer to IBGP peer, it does not change the next-hop attribute.

4) MED

MED （Multi-Exit-Discriminator） attribute is only transferred between two adjacent ASs. Neither of these two AS will notify this attribute to any other third party AS.

MED attribute is equal to IGP Metrics, applying to judge the optimum routing when enter AS. When a device which runs BGP obtains multiple routing that has the same destination addresses but different next-hop addresses via different EBGP peers, it will select the one with the smallest MED value to be optimum routing.

5)Local_Pref attribute

Local_Pref attribute is only valid among IBGP peers, and it cannot notify other ASs. It refers to the BGP priority of the devices.

Local_Pref attribute is applied to judge the optimum routing when the traffic leaves AS. If BGP device obtains multiple routing with the same destination address but different next-hop via different IBGP peers, it will give preference to the routing which has the highest Local_Pref attribute value.

EBGP synchronizes with IBGP:

Synchronize refers to the synchronization between IBGP and IGP so as to avoid misleading external AS device. If non-BGP device exists in AS to offer transfer service, the IP message might be discarded because of the unreachable destination address.

If the synchronizing characteristics have been configured, it will check IGP routing table firstly before adding the routing table to IBGP routing and then releasing it to the peer. Only when IGP also knows this IBGP routing can it be added to routing table and then release to EBGP peer.

26.2 BGP Configuration

26.2.1 Basic Configuration

All the BGP configurations should be performed under BGP configuration mode. The first time you configure the local AS, establish the AS and then enter BGP configuration mode. The

next time you configure, just enter BGP configuration mode.
BGPR outer ID uses 0.0.0.0 by default. Therefore, it asks to configure by manual.

basic configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the local AS and then enter BGP configuration mode	router bgp ***as-num***	required
Configure bgp router id	[no] bgp router-id ***router-id***	required
Exit bgp configuration mode	exit	optional
Delete AS	no router bgp ***as-num***	optional
Show BGP information	show ip bgp	optional

26.2.2 Configure the BGP Neighbor

If AS neighbor is the same as the local AS, it is called IBGP; if not, it is called EBGP. If necessary, you can configure different related parameter for different neighbors. It adopts the global parameter by default, for example, Hold time=180s Keepalive time= 30s. You can make a strategies limit on neighbor. After finishing the configuration of BGP neighbors, DUT will try to create session with the neighbor via the same network interface ip.

configure neighbor and parameter

Operation	Command	Remarks
BGP configuration mode	---	-
Configure BGP neighbor	[no] neighbor ***ip-address*** remote-as ***AS-num***	required
Configure the parameter of keepalve-time and hold-time	[no] neighbor ***ip-address*** timers ***keepalive-time hold-time***	optional
Configure non-	[no] neighbor ***ip-address*** ebgp-multihop [TTL]	required

directly connected neighbor		
Configure the interval time of routing renew message	[no] neighbor ***ip-address*** advertisement-interval ***time***	optional
Configure the self-IP address to be the next hop	[no] neighbor ***ip-address*** next-hop-self	optional
Configure source ip address of update message	[no] neighbor ***ip-address*** update-source loopback-interface ***loopback-ip***	optional

26.2.3 Configure the Timer

After BGP peer established neighbor session, it will maintain the session via keepalive message. If it cannot receive keepalive message in the hold time limit, the connection will be broken. You can configure the interval of keepalive message and hold time according to the needs, and it will work on every neighbor. By default, keepalive=30s, hold time=180s. You can configure different parameters for different neighbors. For more details, please refer to the content of last chapter.

Configure the timer

Operation	Command	Remarks
BGP configuration mode	--	-
Configure the timer	[no] timers bgp ***keepalive-interval hold-time***	optional

26.2.4 Import the Routing

BGP protocol cannot discover routing. Therefore, it needs to introduce the routing of other protocols(for example, IGP or static routing) into the BGP routing table so as to transfer these routing among AS.

It supports *Redistribute* and *Network* when BGP introduces the routing:

Redistribute means introducing the certain routing from RIP routing, OSPF routing, static

routing or connected routing into BGP routing table according to protocol type.

Network is more accurate than *Redistribute*. *Network* introduces a certain routing with the specific prefix and mask into BGP routing table. There must be configured routing or the BGP table will not be allocated.

The above two approaches will not introduce the acquiescent routing, and it has the command to perform separate configurations.

Import the routing

Operation	Command	Remarks
BGP configuration mode	--	-
Import the routing	[no] redistribute **{** connected **|**ospf **|**Rip |Static **}**	required
Import the routing and modify med	redistribute **{** connected |ospf |Rip |Static **}** metric ***metric-value***	required
Import specific routing	network { ***a.0.0.0 |a.b.0.0***	***a.b.c.0***} [med ***value***]	optional
	network ***a.b.c.d /mask-length*** [med ***value***]	optional
	network ***a.b.c.d*** mask ***a.b.c.d*** [med ***value***]	optional
Delete the import specific routing	no network ***a.b.c.d*** [mask ***a.b.c.d*** [med ***value***]]	required
Configure the med of import-default routing	[no]default-metric ***metric-value***	optional
Import default routing	[no]default-information originate	required

Note:

"redistribute { connected |ospf |Rip |Static } metric metric-value" can realize to import routing and modify med at the same time, so there is no need for you to configure "redistribute { connected |ospf |Rip |Static }". Configure either one will be OK.

26.2.5 Configure Routing Aggregation

In large and medium-sized BGP networks, BGP routing table will become very huge. Storage routing table takes up a lot of router memory resources. Moreover, routing information transmission and processing takes up a lot of network resources. In this case, using routes aggregation can greatly reduce the size of the routing table and the burden of network transfer routing information; in addition, hiding a certain specified routing via routes aggregation can reduce routing oscillation on the effects of the network.

In fact, routing aggregation is a process that combing multiple specific routing which has the same prefix into one brief routing. It supports manual aggregation. Specific information is as follows:

Configure routing aggregation

Operation	Command	Remarks
BGP configuration mode	--	-
Configure routing aggregation	[no] aggregate-address { ***ip mask*** | ***ip/mask-length*** }	required
Notice routing aggregation only	[no] aggregate-address { ***ip mask*** | ***ip/mask-length*** } summary-only	optional

Note:

When configure "aggregate-address { ip mask | ip/mask-length } summary-only", you should delete "aggregate-address { ip mask | ip/mask-length}" if it had existed.

26.2.6 Configure the Local Priority

Configure the local priority

Operation	Command	Remarks
BGP configuration mode	--	-
Configure the local priority	[no] bgp default local-preference ***value***	optional

26.2.7 Configure MED

MED（Multi-Exit-Discriminator）can only transmit the attribute between two neighboring AS. Either of the ASs who receives this attribute will not forward the attribute's notice to third-party AS.

MED attributes are equal to IGP Metrics, applying to judge the optimum routing when enter AS. When a device which runs BGP obtains multiple routing with the same destination address but different next-hop addresses via different EBGP peers, it will select the one with the smallest MED value to be optimum routing.

Configure MED

Operation	Command	Remarks
BGP configuration mode	--	-
Configure MED	[no] bgp always-compare-med	optional

26.2.8 Configure Routing Strategy

The operation object of the routing strategy is routing information. Under the normal routing protocol, routing strategy changes the content of the routing table according to a sort rules.

When the routing releasing, receiving and introducing the routing information, it needs to implement some strategies based on actual network so as to filter and change the routing attribute of the routing information. For example:

Control the release of routing: Only release the routing information which meets certain conditions.

Control the receiving of routing: Only receive the necessary and legal routing information so as to control the content of the routing table as well as to improve the network security.

Filter and control the introducing routing: when the routing protocol introduces the routing information from other routing protocol to enrich its routing information, it only introduces the routing information which meets certain conditions and then modifies certain routing attributes of the introducing routing to make it meet the requirements of this agreement.

Set specific routing attributes: Modify the routing attributes which goes through the filter of routing strategy so as to satisfy its needs.

Routing strategy is divided into two steps:

definition rule: Firstly, you should define the characteristic of routing information which is going to bring routing strategy into effect, that is, you should define a set of matching rules. You can set the matching rules according to different attributes of the routing information.

application rules: And then apply the matching rule to the process of routing strategies, such as routing release, receiving and introducing.

The core content of routing strategy is the filter. It can define a set of matching rule by using the filter. The supported patterns are ACL, IP-Prefix List and AS_Path-Filter. Specific information is

as follows:

Access Control List（ACL）

ACLs are sets of rules of filtering rule (or sets of permit or deny statements) that decide what packets can pass and what should be rejected based on matching criteria such as source MAC address, destination MAC address, source IP address, destination IP address, and port number. After that, system classify the packets which has arrived the router and decide what packets can pass and what should be rejected according to filtering rule.

IP-Prefix List

IP-Prefix List is a kind of filter what includes a set of routing filtering rule, and users can be able to define the range of prefix and mask to match the routing destination or next-hop address. IP-Prefix List can be applied to various kinds of dynamic routing protocols, filtering the routing release and routing acceptance.

AS_Path-Filter

AS is a set of rules filtering AS_Path of BGP routing. The BGP routing information includes AS_Path attribute. AS_Path records all the AS number BGP routing that will pass from local area to destination address based on vector sequence, so defining a certain filtering rule based on AS_Path can be able to filter the BGP routing information.

Configure routing strategy

Operation	Command	Remarks
BGP configuration mode	--	-
Configure distribute-list rule	[no] ip distribute-list ***list-num*** { permit | deny } ***ip-address mask***	required
Apply distribute-list	[no]neighbor ***ip-address*** distribute-list ***list-num*** { in |out }	required
Check distribute-list	show ip distribute-list [***distribute-list-number***]	optional
Configure AS_path filter table	[no] ip as-path access-list ***as-list-num*** { permit | deny } ***reg-expression***	required
Apply AS_path filter	[no] neighbor ***ip-address*** filter-list ***as-list-num*** { in | out }	required
Check as-path filter table	show ip as-path access-list [***as-path-list-number***]	optional

26.2.9 Check BGP Information

Check BGP Information

Operation	Command	Remarks
Any mode	----	-

Check BGP routing information	show ip bgp [***ip-address*** | ***A.B.C.D/M***]	optional
Check BGP neighbor information	show ip bgp neighbors [***neighbor-address***]	optional
Check the summary information of neighbor information	show ip bgp summary	optional
Check hardware list of routing forward	show ip fdb [***ip-address*** [***mask***]]	optional
Enter BGP configuration mode	---	
Check AS-path filter information	show ip as-path access-list [***as-path-list-number***]	optional
Check distribute-list information	show ip distribute-list [***distribute-list-number***]	optional

26.3 Example for BGP Configuration

1.Network requirements

SW1 connects SW2 via IF1, establishing EBGP. Add a static router to SW2. Then check SW1 peer state and BGP routing table.

sketch map of BGP

2.Configuration steps

Enter ONU mode

Basic configuration

SW1(config)#interface vlan-interface 1

SW1(config-if-vlanInterface-1)#ip address 192.168.1.20 255.255.255.0

SW1(config-if-vlanInterface-1)#exit

SW2(config)#interface vlan-interface 1

SW2(config-if-vlanInterface-1)#ip address 192.168.1.10 255.255.255.0

SW2(config-if-vlanInterface-1)#exit

SW2(config)#ip route 200.200.200.200 255.255.255.255 192.168.1.200

#BGP configuration

SW2(config)#router bgp 100

SW2(config-router-bgp)#neighbor 192.168.1.20 remote-as 200

SW2(config-router-bgp)#redistribute static

SW2(config-router-bgp)#ex

SW1(config)#router bgp 200

SW1(config-router-bgp)#bgp router-id 20.20.20.20

SW1(config-router-bgp)#neighbor 192.168.1.10 remote-as 100

3.Result validation

check BGP neighbor summary

SW1(config-router-bgp)#show ip bgp summary

Neighbor	VR	V	AS	MsgRcvd	MsgSent	Up/Down	State/PfxRcd
192.168.1.10	0	4	100	1	0	00:00:03	Established

Total number of neighbors 1

import an static routing to SW2:

SW2(config-router-bgp)#network 200.200.200.200 mask 255.255.255.255

#check BGP routing table on SW1: learn the import-route from SW2

SW1(config-router-bgp)#show ip bgp

Autonomous System number 200, local router ID 20.20.20.20

Status codes: s suppressed, * valid, > best, i internal

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	NextHop	Metric	LocalPref	Path

*> 200.200.200.200 192.168.1.200 100 i

Total number of best entries 1

27.Other Routing Configurations

27.1 IP-Def-CPU Overview

Equipment supports two types of forwarding pattern： 1） flow forwarding； 2） network topology forwarding.

In flow forwarding, it forwards the failure routing or the unreachable packet to CPU for further processing; in network topology forwarding, it discards these packets directly. So you should be careful when you adopt this pattern.

It adopts flow forwarding by default.

27.1.1 Configure IP-Def-CPU

Configure IP-Def-CPU

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
transmit the unknown packet to cpu	[no] IP def cpu	optional
transmit the unknown packet of vlan interface to cpu	[no] ip def cpu vlan ***vlan-id***	optional
Display the information	show ip def cpu	optional

Note:

1. It enables IP-Def-cpu by default even if flow forwarding is adopted；

2. All vlan interfaces allow unknown packet to transmit to cpu by default. In addition, the corresponding layer-3 interface of the vlan should be existed when configuring.

27.1.2 Configuration Example

1.Network requirements

Only permit the unknown packet of vlan 100 interface to transmit to cpu. Do not transmit the unknown packet of other vlan to cpu；

2.Configuration steps

disable the function of transmitting the unknown packet to cpu
SW(config)#no ip def cpu

permit the unknown packet of vlan 100 to transmit to cpu.
SW(config)#vlan 100
SW(config-if-vlan)#interface vlan-interface 100
SW(config-if-vlanInterface-100)#ex
SW(config)#ip def cpu vlan 100

Display the configuration
SW(config)#show ip def cpu
Routing def routes and def hosts to CPU: : FALSE
The IP destination of packet belonged to vlan interface can be send to CPU:
VLAN 100

27.2 URPF

URPF (Unicast Reverse Path Forwarding) is to prevent the network attack based on source address spoofing. URPF obtains the source address and ingress interface of the packet, taking the source address as the destination address, and then searching the route which corresponds to the source address in the routing table. If the route exists, it will be forwarded, otherwise it will be discarded.

URPF has two modes:

1.strict mode: That is, when searching the routing table for reverse path detection, it must match the source address existing in the routing table, and the egress interface to the source address of the packet is the same as the ingress interface of the packet.

2.loose mode: That is, when searching the routing table for reverse path detection, only to detect whether the source address of the packet is existent in the unicast routing table. If the routing table exists, it can pass the detection.

Under interface mode, configure the URPF function of the corresponding interface.

27.2.1 Configure URPF

Configure URPF

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enter interface	interface { vlan-interface | supervlan-interface } ***if-id***	required

configuration mode		
Enable and configure the mode	urpf { strict | loose }	required
Disable the function	no urpf	optional
Display the configurations	show urpf [interface { vlan-interface | supervlan-interface } ***if-id***]	optional

27.2.2 URPF Configuration Example

1.Network requirements

There are two Layer3 interfaces: if-1, if-2.

If-1 enables urpf for strict mode. There is a default route and next hop points to the ip in if-2 interface. Ixia A sends an if-1 packet to a Layer 3 packet whose sip is not equal to that of the if-1 network segment and then validates the urpf process.

sketch map of URPF

2.Configuration steps

Enable urpf on VLAN-interface 1 and use strict mode

SW(config)#interface vlan 1

SW(config-if-vlanInterface-1)#ip address 192.168.1.27 255.255.255.0

This ipaddress will be the primary ipaddress of this interface.

Config ipaddress successfully!

SW(config-if-vlanInterface-1)#urpf strict

Configure URPF strict mode successfully.

Configure VLAN 2 interface, and configure the default route as VLAN 2 interface

SW(config-if-vlanInterface-1)#vlan 2

SW(config-if-vlan)#interface ethernet 0/0/2

SW(config-if-ethernet-0/0/2)#switchport default vlan 2

SW(config-if-ethernet-0/0/2)#switchport mode access

SW(config-if-ethernet-0/0/2)#interface vlan-interface 2

Create vlan-interface successfully!

SW(config-if-vlanInterface-2)#ip address 192.168.2.27 255.255.255.0

This ipaddress will be the primary ipaddress of this interface.
Config ipaddress successfully!

SW(config-if-vlanInterface-2)#exit
SW(config)#ip route 0.0.0.0 0.0.0.0 192.168.2.100
Config static route successfully!

SW(config)#logging monitor 0
SW(config)#debug urpf

3.Result validation
(1) Display urpf configuration information.
SW(config-if-vlanInterface-1)#exit
SW(config)#show urpf
Interface URPF Status
VLAN-IF1 Strict Mode

(2) Ixia A forwards a Layer 3 packet to the if-1 interface whose sip is not equal to that of the if-1 network segment, the following log information will be printed:
SW(config)#00:06:21: SW: %URPF-7-urpf:
VLAN 1:--6-- strict mode, route exists, interface is different, packet dropped

SW(config)#00:06:50: SW: %URPF-7-urpf:
VLAN 1:--6-- strict mode, route exists, interface is different, packet dropped

(3) If-1 interface runs loose mode, forwarding packets in the same way: Ixia A forwards a Layer 3 packet to the if-1 interface whose sip is not equal to that of the if-1 network segment, the following log information will be printed:
SW(config)#interface vlan-interface 1
SW(config-if-vlanInterface-1)#urpf loose
Configure URPF loose mode successfully.

SW(config-if-vlanInterface-1)#00:37:25: SW: %URPF-7-urpf:
VLAN 1:--7-- loose mode, route exists, packet allowed

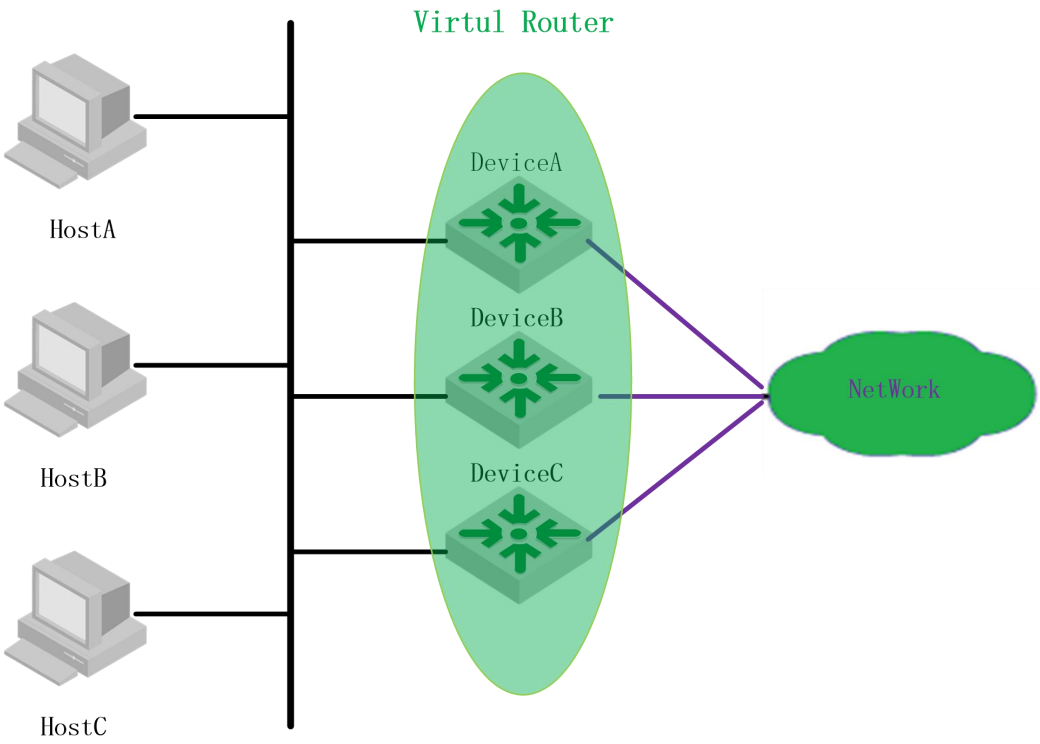
28.VRRP Configuration

28.1 VRRP Overview

VRRP is usually used to provide redundancy for gateway devices. VRRP is implemented in VRRPv2 and VRRPv3. VRRPv2 is based on IPv4, while VRRPv3 is based on IPv6.

VRRP can add a group of routers that are responsible for gateway functions to a backup group (or Layer 3 switch), and form a virtual router (or Layer 3 switch). As shown in Figure 1-1, Device A, Device B, and Device C form a virtual router. The virtual router has its own IP address. The hosts in the LAN set the virtual router as the default gateway. The router with the highest priority in Device A, Device B, and Device C functions as the master router and functions as the gateway. The other two routers act as backup routers. When the master fails, VRRP re-elects a new master to ensure that traffic forwarding is not interrupted.

VRRP Networking diagram



VRRP related concept

1. VRRP switch and virtual switch

VRRP switch refers to a Layer 3 switch that runs VRRP. It is a physical entity. A virtual switch refers that is created by the VRRP protocol. It is a logical concept. A group of VRRP switches work together, and form a virtual switch. The virtual switch is represented as a logical router with a unique fixed IP address and MAC address.

2. The Master switch and backup switch.

A switch in the same VRRP backup group has two mutually exclusive roles: the master switch and the backup switch. In a VRRP group, there is only one switch in the master role. User can have one or more switches in the backup role. VRRP uses a selection policy to select one from the switch group as the master, responsible for ARP responses and forwarding IP messages. The role of the other switches in the group as a backup is on standby. When the master router fails for some reason, the backup switch can be upgraded to the master switch after a delay of several seconds. Since this handover is very fast and does not require changing the IP address and MAC address, it is transparent to the end-user system.

28.2 VRRP Basic Configuration

28.2.1 Configure the Virtual IP of the VRRP Backup Group

Configure the IP address of the VRRP backup group (virtual switch). The hosts of the LAN must set their own gateway (the next hop address of the default route). The hosts communicate with the external network through this virtual router in the network.

Configure the virtual IP address of the VRRP group

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter vlan or superVLAN interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	-
Assign an IP address vip to a virtual SWITCH (backup group) numbered vrid,	**ip vrrp** *vrid vip*	optional

Vrid: Backup group number, vip: virtual ip

🕮 **Note:**

1. The backup group number ranges from 1 to 255. The virtual address can be an unassigned IP address in the network segment where the VRRP group resides or the IP address of the interface that belongs to the VRRP group.
2. You can configure up to 255 VRRP groups. Each VRRP group can be configured up to eight IP addresses.
3. If the virtual IP and switch ip are same, called the switch for an address owner (IP Address Owner), generally do not recommend such a configuration.
4. When you specify the first IP address to a VRRP group, the system creates the VRRP group. When you specify a virtual IP address for the VRRP group later, the system adds the IP address to the virtual IP address list of the VRRP group.
5. After the last virtual IP address in the VRRP group is deleted, the VRRP group is deleted at the same time. That is, the backup group no longer exists on this interface. The configurations of the

VRRP group are no longer valid.

28.2.2 Configure the Priority of the Switch in the VRRP group

A switch of a VRRP group determines its role in the VRRP group based on its priority. The switch with the highest priority becomes Master switch and the lower priority becomes Backup switch.

Configure the priority of the Switch in the VRRP group

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter vlan or superVLAN interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	-
Configure the virtual SWITCH priority as vrid	**vrrp priority** *vrid priority*	Optional. The default value is 100.

Note:

1. The priority ranges from 0 to 255 (the higher the value indicates the higher the priority), but the configurable range is from 1 to 254. Priority 0 is for the system reserved for special use to use, 255 is the system reserved for the IP address owner.
2. When the IP address owner exists in the VRRP group, the master router works as long as it works normally.

28.2.3 Configure the Work State of the Switch in the VRRP Group

A switch of a VRRP group works as the following two modes:

• Non-preemptive mode: If the switch in the VRRP group works in non-preemptive mode, as long as Master switch does not fail, Backup switch does not become Master switch even if it is configured with a higher priority.

• Preemptive mode: If a switch of the VRRP group works in preemptive mode, it will send VRRP advertisement messages if it finds that its priority is higher than that of the current master, which causes the switch in the backup group to re-elect the Master switch and replace the original Master switch. Accordingly, the original Master switch will become Backup switch.

Configure the work mode of the switch in the VRRP group

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter vlan or superVLAN interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	-

Enable the preemption function for the VRRP group	Enable preemption	**vrrp preempt** *vrid*	Optional, The default is preemption mode
	Non-preemptive mode	**no vrrp preempt** *vrid*	optional

28.2.4 Configure the Preemption Delay of the Backup Group

In the preemption mode, user can set the delay time for backup group. This allows Backup to delay a period of time to become a master. The purpose of setting the delay time is as follows: In a network with insufficient performance, if Backup does not receive the messages from the master on time, it becomes the master (The cause that Backup cannot receive messages is caused by network congestion. It is not caused by the failure of the master to work properly) and waits for a period of time to receive the message from the Master, thus avoiding the frequent state transition.

Set the preemption delay for the switch in the VRRP group

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter vlan or superVLAN interface configuration mode	**Interface {vlan-interface|** **supervlan-interface]}** *vlan-id*	-
Set the preemption delay for the SWITCH in the VRRP group	**vrrp preempt** *vrid* **delay** *delay*	Optional, the delay is 0

28.2.5 Configure the Switch Advertisement Interval in the Backup Group

Master switch periodically (The time interval is adver_interval) sends VRRP advertisements to notify other switches in the backup group that they work properly. Backup starts the timer to wait for the advertisement message to arrive. If the backup does not receive any VRRP message from the master for a specified period of time (The time interval is master_down_interval), it considers that it does not work normally. At the same time their status will change into Master.

Configure the switch advertisement interval in the VRRP group

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter vlan or superVLAN interface	**Interface {vlan-interface|**	-

configuration mode	**supervlan-interface]}** *vlan-id*	
Configure the switch advertisement interval in the VRRP group	**vrrp timer** *vrid adver-interval*	Optional, the default is 1S

Note:

1. The interval of the master_down_interval of the Backup switch is 3 times of the adver_interval.
2. If the network traffic is too large or different timer on the SWITCH has difference and other factors, they will lead to the master_down_interval exception to time out and state transition. In this case, it can be done by extending the adver_interval and setting the delay time.

28.2.6 Configure VRRP Tack Function

The VRRP track function monitors the uplink interface or uplink status and changes the priority of the router based on the status of the uplink.

When the uplink fails, the hosts cannot access the external network through the switch of the LAN, the status of the monitored track entry is down, and the priority of the router is automatically reduced by a certain value. So the priority of other switches in the backup group is higher than that of the switch, and becomes Master switch. This ensures that the communication between hosts in the LAN and the external network is not interrupted.

Configure VRRP Track function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enter vlan or superVLAN interface configuration mode	**Interface {vlan-interface| supervlan-interface]}** *vlan-id*	-
Configuration specified monitoring interface of backup group SWITCH	**vrrp track** *vrid* **{ supervlan-if** *<vid>* **| vlan-if** *<vid>* **}**	optional
Reduce the priority when the interface is down	**vrrp track** *vrid* **{ supervlan-if** *<vid>* **| vlan-if** *<vid>* **} reduced** *pri-value*	optional

28.2.7 Configure VRRP Ping Function

With the VRRP ping function enabled, the PC can ping the virtual IP.

Configure VRRP Ping function

operation	command	remark
Enter the global configuration mode	**configure terminal**	-
Enable the VRRP Ping function	**vrrp ping-enable**	-
Disable the VRRP Ping function	**no vrrp ping-enable**	

If the backup group is configured first, vrrp ping-enable cannot be enabled.

28.2.8 VRRP Display and Maintain

VRRP display and maintain

operation	command	remark
Display the virtual SWITCH information already configured on the current system	**show vrrp [vlan-interface** *vlan-id* [*vrid*] **]**	
Debug VRRP	**debug vrrp**	

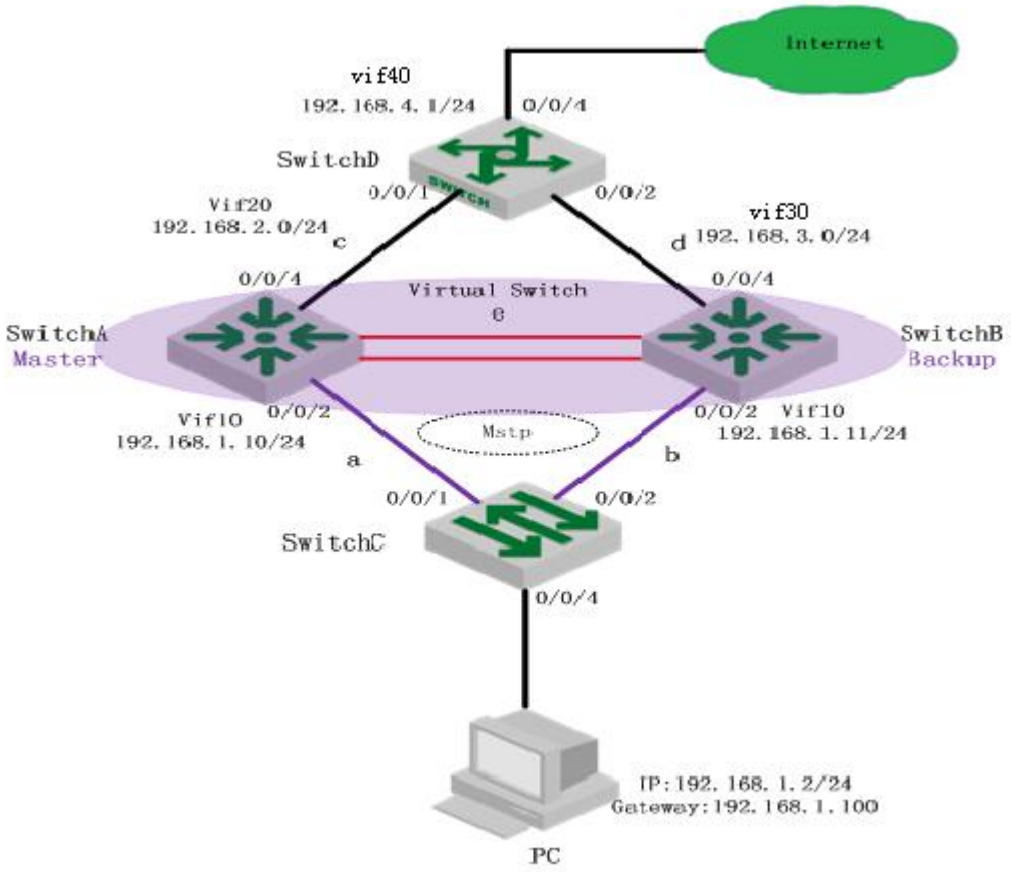
28.3 VRRP Configuration Example

Requirement and network

- VRRP technology is adopted to back up the gateway equipment, and improve the reliability of the gateway. When one gateway device fails, the host in the LAN can still access the external network through another gateway device.
- When the uplink of the gateway device fails, reduce the priority of the gateway device to prevent the gateway device from becoming the master device, causing the traffic forwarding to be interrupted.
- Redundant backup of the Layer 2 link in the LAN ensures that the traffic forwarding will not be interrupted when the downlink of the gateway is faulty. Use MSTP to avoid loops in Layer 2 networks.
- Add a heartbeat between Device A and Device B to provide redundancy for the downlink, and use MSTP to block the redundant links in the network to eliminate the Layer 2 loops.
- The gateway device is connected to the Internet through the egress device at the core layer.

  To implement the above requirements, user needs to adopt the MSTP + VRRP networking mode, as shown in below figure:

VRRP Configuration networking diagram



Switch A and switch B form a VRRP backup group. Switch A has higher priority. Normally, it is the master device of the VRRP backup group. Core device switch D connects with the Internet. Enable OSPF routing protocol between Switch A, Switch B and Switch D; The aggregation link e is added between Switch A and Switch B. MSTP is used between the switch C of the access layer to perform redundancy backup of Layer 2 links and eliminate Layer 2 loops. Use the VRRP track function on Switch A to reduce the priority of the gateway device when the uplink of the gateway device is faulty. The PC sets the virtual IP address as the gateway.

Configuration steps

1) Configure the switch interface, IP

SwitchA(config)#VLAN 10

SwitchA(config-if-vlan)#vlan 20

SwitchA(config-if-vlan)#exit

SwitchA(config)#interface ethernet 0/0/2

SwitchA(config-if-ethernet-0/0/2)#switchport default vlan 10

SwitchA(config-if-ethernet-0/0/2)#interface ethernet 0/0/4

SwitchA(config-if-ethernet-0/0/4)#switchport default vlan 20

SwitchA(config-if-vlan)#exit

SwitchA(config-if-ethernet-0/0/4)#interface vlan-interface 10

SwitchA(config-if-vlanInterface-10)#ip address 192.168.1.10 255.255.255.0

SwitchA(config)#interface vlan-interface 20
SwitchA(config-if-vlanInterface-20)#ip address 192.168.2.1 255.255.255.0

SwitchB(config)#vlan 10
SwitchB(config-if-vlan)#vlan 30
SwitchB(config-if-vlan)#exit
SwitchB(config)#interface ethernet 0/0/2
SwitchB(config-if-ethernet-0/0/2)#switchport default vlan 10
SwitchB(config-if-ethernet-0/0/2)#interface ethernet 0/0/4
SwitchB(config-if-ethernet-0/0/4)#switchport default vlan 30
SwitchB(config-if-ethernet-0/0/4)#exit
SwitchB(config)#interface vlan-interface 10
SwitchB(config-if-vlanInterface-10)#ip address 192.168.1.11 255.255.255.0
SwitchB(config-if-vlanInterface-10)#interface vlan-interface 30
SwitchB(config-if-vlanInterface-30)#ip address 192.168.3.1 255.255.255.0

SwitchD(config)#vlan 20
SwitchD(config-if-vlan)#vlan 30
SwitchD(config-if-vlan)#vlan 40
SwitchD(config-if-vlan)#exit
SwitchD(config)#interface ethernet 0/0/2
SwitchD(config-if-ethernet-0/0/2)#switchport default vlan 30
SwitchD(config-if-ethernet-0/0/2)#interface ethernet 0/0/1
SwitchD(config-if-ethernet-0/0/1)#switchport default vlan 20
SwitchD(config-if-ethernet-0/0/1)#interface ethernet 0/0/4
SwitchD(config-if-ethernet-0/0/4)#switchport default vlan 40
SwitchD(config-if-ethernet-0/0/4)#interface vlan-interface 20
SwitchD(config-if-vlanInterface-20)#ip address 192.168.2.2 255.255.255.0
SwitchD(config-if-vlanInterface-20)#interface vlan-interface 30
SwitchD(config-if-vlanInterface-30)#ip address 192.168.3.2 255.255.255.0
SwitchD(config-if-vlanInterface-30)#interface vlan-interface 40
SwitchD(config-if-vlanInterface-40)#ip address 192.168.4.1 255.255.255.0

2) Configure link aggregation between Switch A and Switch B
SwitchA(config)#vlan 10
SwitchA(config-if-vlan)#switchport ethernet 0/0/11
SwitchA(config-if-vlan)#switchport ethernet 0/0/12
SwitchA(config-if-vlan)#interface range ethernet 0/0/11 to ethernet 0/0/12
SwitchA(config-if-range)#channel-group 1 mode on

SwitchB(config)#vlan 10
SwitchB(config-if-vlan)#switchport ethernet 0/0/11 ethernet 0/0/12
SwitchB(config-if-vlan)#
SwitchB(config-if-vlan)#interface range ethernet 0/0/11 to ethernet 0/0/12

SwitchB(config-if-range)#channel-group 1 mode on

3) Configure OSPF routes between Switch A, Switch B, and Switch D
SwitchA(config)#router ospf
SwitchA(config-router-ospf)#network 192.168.1.10 0.0.0.255 area 0
SwitchA(config-router-ospf)#network 192.168.2.1 0.0.0.255 area 0

SwitchB(config)#router ospf
SwitchB(config-router-ospf)#network 192.168.1.11 0.0.0.255 area 0
SwitchB(config-router-ospf)#network 192.168.3.1 0.0.0.255 area 0

SwitchD(config)#
SwitchD(config)#router ospf
SwitchD(config-router-ospf)#network 192.168.2.2 0.0.0.255 area 0
SwitchD(config-router-ospf)#network 192.168.3.2 0.0.0.255 area 0
SwitchD(config-router-ospf)#network 192.168.4.1 0.0.0.255 area 0

4) Configure MSTP between Switch A, Switch B, and Switch C. Map VLAN 10 to instance 1. Configure the instance priority so that Switch A is the root bridge in instance 1, and Switch B is the backup root bridge in instance 1. The uplink port Switch D accessed Switch A and Switch B closes the spanning tree.
SwitchA(config)#spanning-tree
SwitchA(config)#spanning-tree mode mstp
SwitchA(config)#spanning-tree mst instance 1 vlan 10
SwitchA(config)#spanning-tree mst instance 1 priority 4096
SwitchA(config)#interface ethernet 0/0/4
SwitchA(config-if-ethernet-0/0/4)#no spanning-tree

SwitchB(config)#spanning-tree
SwitchB(config)#spanning-tree mode mstp
SwitchB(config)#spanning-tree mst name vrrp
SwitchB(config)#spanning-tree mst instance 1 vlan 10
SwitchB(config)#spanning-tree mst instance 1 priority 8192
SwitchB(config)#interface ethernet 0/0/4
SwitchB(config-if-ethernet-0/0/4)#no spanning-tree

SwitchC(config)#vlan 10
SwitchC(config-if-vlan)#switchport ethernet 0/0/1
SwitchC(config-if-vlan)#switchport ethernet 0/0/2
SwitchC(config-if-vlan)#interface ethernet 0/0/4
SwitchC(config-if-ethernet-0/0/4)#switchport default vlan 10
SwitchC(config)#spanning-tree
SwitchC(config)#spanning-tree mode mstp
SwitchC(config)#spanning-tree mst name vrrp

SwitchC(config)#spanning-tree mst instance 1 vlan 10
SwitchC(config)#

5) The Switch A and Switch B are configured as VRRP backup groups. The Switch A is the master device and the virtual IP of the backup group is 192.168.1.100.
SwitchA(config)#interface vlan-interface 10
SwitchA(config-if-vlanInterface-10)#ip vrrp 1 192.168.1.100
SwitchA(config-if-vlanInterface-10)#vrrp priority 1 110
SwitchA(config-if-vlanInterface-10)#vrrp track 1 vlan-if 20
SwitchA(config-if-vlanInterface-10)#vrrp track 1 vlan-if 20 reduced 20

SwitchB(config)#interface vlan-interface 10
SwitchB(config-if-vlanInterface-10)#ip vrrp 1 192.168.1.100

Configuration authentication

After the preceding configurations are complete, you can view the VRRP backup group information on the Switch A and Switch B. You can see that the Switch A is the master in the VRRP group and the Switch B is the backup in the backup group, so that the PC can communicate with external by the Switch A.
SwitchA(config)#show vrrp
VLAN-IF10 | Virtual Router 1
State : Master
Virtual IP : 192.168.1.100
Priority : 110
Preempt : YES
Delay Time (secs) : 0
Timer (secs) : 1
track interfaces:
VLAN-IF20, reduced priority: 20, status: up

SwitchB(config)#show vrrp
VLAN-IF10 | Virtual Router 1
State : Backup
Virtual IP : 192.168.1.100
Priority : 100
Preempt : YES
Delay Time (secs) : 0
Timer (secs) : 1
track interfaces:

Total entries:1

View the spanning tree status on Switch C. You can see that0/ 0/2 is the replacement port and the status is DIS, thus eliminating the loop.

SwitchC(config)#show spanning-tree mst instance brief 1
Current spanning tree protocol is MSTP
Spanning tree protocol is enable

MSTP Instance 1 vlans mapped:10
Bridge ID 32768-0001.7a00.0403
CIST root 32768-0000.0000.4302
Region root 4096-000a.5a15.1617
Bridge time HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External rpc: 0, Internal rpc: 20000

PortID	Role	Sts	ExternalCost	InternalCost	Prio.Nbr	Type
e0/0/1	Root	FWD	20000	20000	128.1	P2P
e0/0/2	Alte	DIS	20000	20000	128.2	P2P
e0/0/4	Design	FWD	20000	20000	128.4	P2P

The PC ping the IP 192.168.1.1 of the interface 30 of the Switch D. It can ping. By querying the ARP table, you can see that the PC communicates with the Switch D through the Switch A.

SwitchA(config)#show arp all
Informations of ARP
d - days, h - hours, m - minutes, s - seconds

IpAddress	Mac_Address	Vlan	Port	Type	ExpireTime	Status
192.168.1.2	c8:3a:35:d3:e3:99	10	e0/0/2	dynamic	19m58s	valid
192.168.1.11	00:00:00:00:43:02	10	e0/0/2	dynamic	03m31s	valid
192.168.2.2	00:0a:5a:21:93:fd	20	e0/0/4	dynamic	02m31s	valid

Total entries:3

SwitchB(config)#show arp all
Informations of ARP
d - days, h - hours, m - minutes, s - seconds

IpAddress	Mac_Address	Vlan	Port	Type	ExpireTime	Status
192.168.1.10	00:0a:5a:15:16:17	10	e0/0/2	dynamic	02m31s	valid
192.168.3.2	00:0a:5a:21:93:fd	30	e0/0/4	dynamic	01m03s	valid

Total entries:2

SwitchD Ping PC successfully.

SwitchD#ping 192.168.1.2
PING 192.168.1.2: with 32 bytes of data:

reply from 192.168.1.2: bytes=32 time<10ms TTL=127
reply from 192.168.1.2: bytes=32 time<10ms TTL=127
----192.168.1.2 PING Statistics----
2 packets transmitted, 2 packets received, 0% packet loss
round-trip (ms) min/avg/max = 0/0/0
Control-C

When link c or link a fails, check the VRRP backup group information. User can see that the priority of Switch A changes from 100 to 90, Switch A becomes the backup device in the backup group, and Switch B becomes the master in the VRRP group equipment. Check the state of the spanning tree of the Switch C. 0/0/2 becomes the root port and FWD status. At this point the PC can still communicate with Switch D through Switch B.

SwitchA(config)#interface ethernet 0/0/4
SwitchA(config-if-ethernet-0/0/4)#shutdown
SwitchA(config)#interface ethernet 0/0/2
SwitchA(config-if-ethernet-0/0/4)#shutdown
SwitchA(config-if-ethernet-0/0/4)#show ip interface vlan-interface 20
Show informations of interface
The mac-address of interface is 00:0a:5a:15:16:17
Interface name : VLAN-IF20
Primary ipaddress : 192.168.2.1/255.255.255.0
Secondary ipaddress : None
VLAN : 20
Address-range : NONE
Interface status : Down

Total entries: 1 interface.

SwitchA(config-if-ethernet-0/0/4)#show vrrp
VLAN-IF10 | Virtual Router 1
State : Backup
Virtual IP : 192.168.1.100
Priority : 90
Preempt : YES
Delay Time (secs) : 0
Timer (secs) : 1
track interfaces:
VLAN-IF20, reduced priority: 20, status: down

SwitchB(config)#show vrrp
VLAN-IF10 | Virtual Router 1
State : Master
Virtual IP : 192.168.1.100

Priority : 100
Preempt : YES
Delay Time (secs) : 0
Timer (secs) : 1
track interfaces:

Total entries:1

SwitchC(config)#show spanning-tree mst instance brief 1
Current spanning tree protocol is MSTP
Spanning tree protocol is enable

MSTP Instance 1 vlans mapped:10
Bridge ID 32768-0001.7a00.0403
CIST root 32768-0000.0000.4302
Region root 4096-000a.5a15.1617
Bridge time HelloTime 2,MaxAge 20,ForwardDelay 15,MaxHops 20
Cist Root time HelloTime 2,MaxAge 20,ForwardDelay 15,RemainingHops 19
External rpc: 0, Internal rpc: 40000

PortID	Role	Sts	ExternalCost	InternalCost	Prio.Nbr	Type
e0/0/2	Root	FWD	20000	20000	128.2	P2P
e0/0/4	Design	FWD	20000	20000	128.4	P2P

29.802.1X Configuration

29.1 802.1x Overview

IEEE 802.1X is the accessing management protocol standard based on interface accessing control passed in June, 2001. Traditional LAN does not provide accessing authentication. User can access the devices and resources in LAN when connecting to the LAN, which is a safety loophole. For application of mobile office and CPN, device provider hopes to control and configure user's connecting. There is also the need for accounting.

IEEE 802.1X is a network accessing control technology based on interface, which is the accessing devices authentication and control by physical accessing level of LAN devices. Physical accessing level here means the interface of LAN Switch devices. When authenticating, Switch is the in-between (agency) of client and authentication server. It obtains user's identity from client of accessing Switch and verifies the information through authentication server. If the authentication passes, this user is allowed to access LAN resources or it will be refused.

29.1.1 802.1x Authentication

802.1X operates in the typical client/server model and defines three entities: supplicant system, authentication system, and authentication server system:

- **Supplicant System**: It is required to access the LAN, and enjoy the services provided by the Switch equipment (such as PC), the client needs to support EAPOL agreement, and the client must run the IEEE 802.1X authentication client software.
- **Authentication System**: In the Ethernet system, the authentication Switch is mainly used to upload and deliver user authentication information and control whether the port is available according to the authentication result. As if between the client and the authentication server to act as a proxy role.
- **Authentication Server**: Normally refers to the RADIUS server. RADIUS checks the identity of the client (user name and password) to determine whether the user has the right to use the network system to provide network services. After the end of the authentication, results will be sent to the Switch.

Figure 1-1 shows the relationship between the three parts.

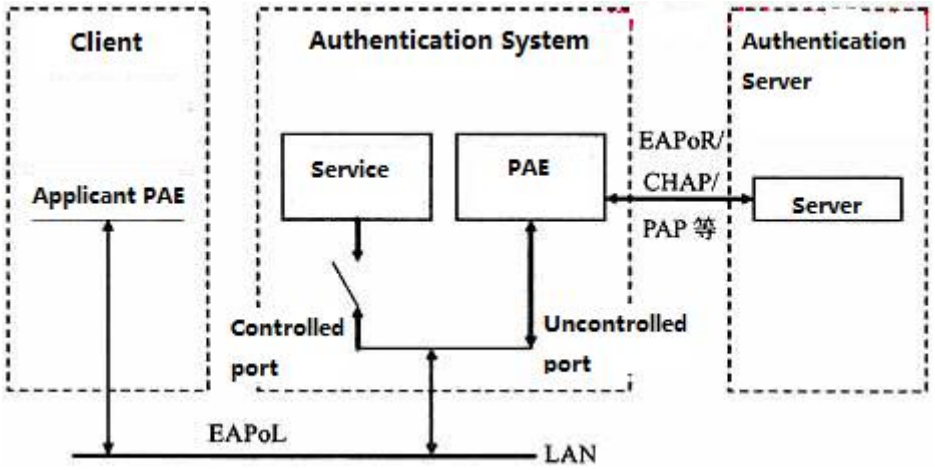


Figure 1-1 architecture of 802.1X

The above systems involve three basic concepts: PAE, controlled port, control direction:

1. PAE

Port Access Entity (PAE) refers to the entity that performs the 802.1x algorithm and protocol operations.

- PAE is the entity responsible for performing algorithms and protocol operations in the authentication mechanism. The PAE uses the authentication server to authenticate the clients that need to access the LAN, and controls the authorized / unauthorized status of the controlled ports accordingly according to the authentication result. The client PAE responds to the authentication request from the device and sends the user authentication information to the device. The client PAE can also send the authentication request and the offline request to the device.

2. Controlled port and uncontrolled port

An authenticator provides ports for supplicants to access the LAN. Each of the ports can be regarded as two logical ports: a controlled port and an uncontrolled port.

- The uncontrolled port is always enabled in both the ingress and egress directions to allow EAPOL protocol frames to pass, guaranteeing that the supplicant can always send and receive authentication frames.
- The controlled port is enabled to allow normal traffic to pass only when it is in the authorized state.
- The controlled port and uncontrolled port are two parts of the same port. Any frames arriving at the port are visible to both of them.

3. Control direction

In the non-authorized state, the controlled port is set to one-way controlled: the implementation of one-way controlled, prohibits the receiving frame from the client, but allows the client to send frames.

4． Port controlled manner

- Port-based authentication:

  As long as the first user authentication is successful under the physical ports, other access users without authentication can use the network source, when the first user is off line, other users will be refused to use network.

- MAC-address-based authentication:

  All the users on the physical port need to be authenticated separately. When userA goes offline, only the userA cannot use the network.

29.1.2 802.1x Authentication Process

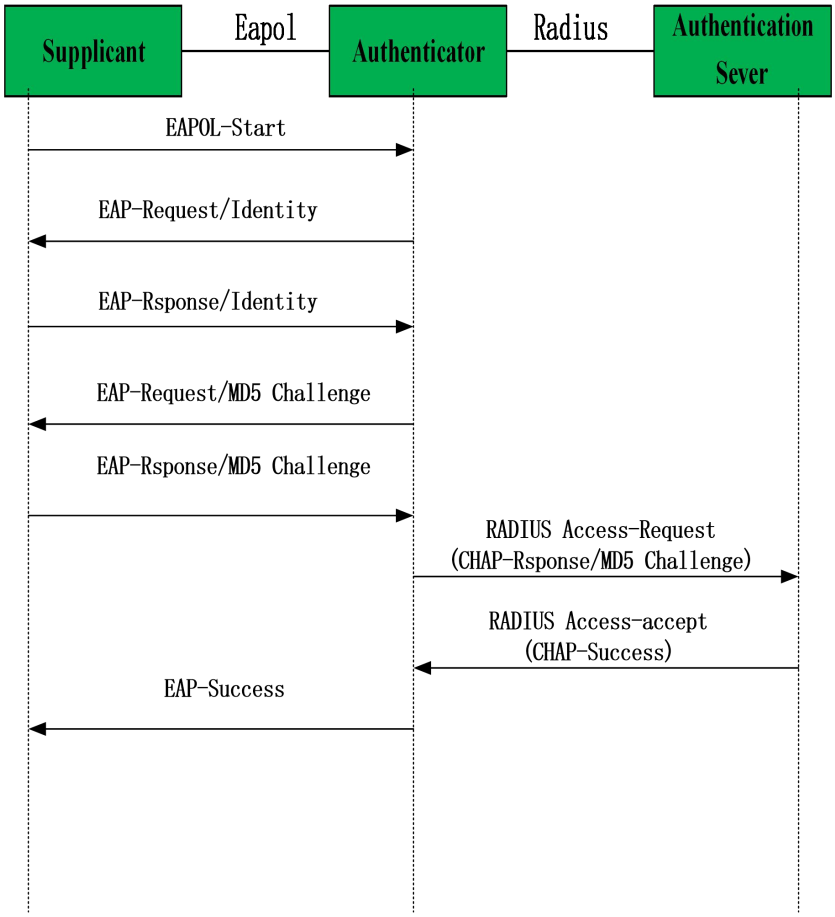
The 802.1x authentication system employs the Extensible Authentication Protocol (EAP) to exchange authentication information between the supplicant PAE, authenticator PAE, and authentication server.

At present, the EAP relay mode supports four authentication methods: EAP-MD5, EAP-TLS (Transport Layer Security), EAP-TTLS (Tunneled Transport Layer Security), and PEAP (Protected Extensible Authentication Protocol).

Switch supports EAP-Transfer mode and EAP-Finish mode to interactive with remote RADIUS server to finish the authentication.

1. EAP-Transfer

The following takes EAP-Transfer authentication process for an example to introduce the basic service procedure. As shown in the following:



EAP-Transfer authentication process

The authentication process is as follows:

(1) When the user needs to access the network, it will input the registered user name and password through the 802.1X client and initiate the connection request (EAPOL-Start packet). At this point, the client program will send the request message to the device, start an authentication process.;

(2) After receiving the requested data frame, the access device sends out a request frame (EAP-Request / Identity packet) to ask the user's client program for the user name;

(3) The client responds to the request from the device and sends the user name information to the device through the data frame (EAP-Response / Identity packet). The device encapsulates the RADIUS Access-Request packet and then sends it to the authentication server for processing after receiving the data frame packet from the client;

(4) After receiving the user name information from the device, the RADIUS server compares the information with the user name table in the database, finds the corresponding password information, and encrypts it with a randomly generated encryption key. And it sends

the encrypted keyword to the device through a RADIUS Access-Challenge packet. The message is then forwarded by the device to the client;

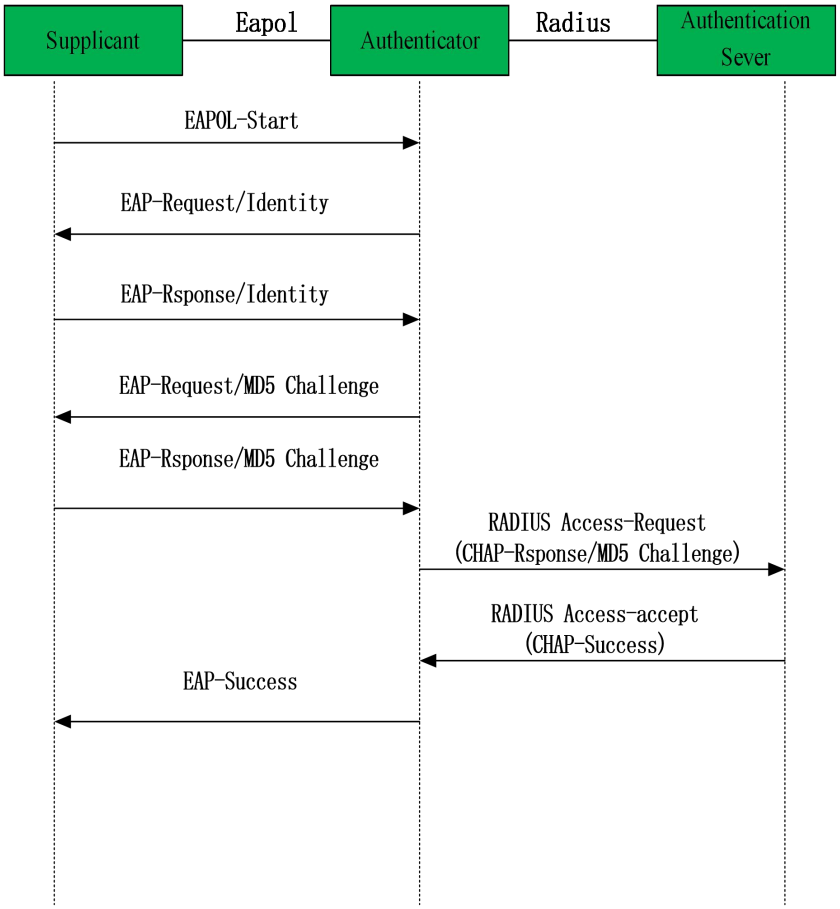
(5) After receiving the EAP-Request / MD5 Challenge packet, the client encrypts the encrypted part (this encryption algorithm is usually irreversible) and generates the EAP-Response / MD5 Challenge packets and pass the authentication packets to the authentication server.;

(6) The RADIUS server compares the received encrypted information (RADIUS Access-Request packet) with the local encrypted password information. If the password is the same, the RADIUS server considers the user to be a valid user and sends out the message -Accept and EAP-Success);

(7) After receiving the authentication message, the device changes the port to the authorized state, allowing the user to access the network through the port.

2. EAP-Finsh

In this way, EAP packets are terminated at the device end and are mapped to RADIUS packets. The RADIUS server uses the standard RADIUS protocol to complete authentication, authorization, and accounting. The PAP or CHAP authentication method can be adopted between the device and the RADIUS server. Our Switch defaults to this mode. The following takes the CHAP authentication method as an example to describe the basic service flow, as shown below:



EAP-Finish authentication process

The EAP termination mode differs from the authentication process of EAP relay mode in that a random encryption key for encrypting the user's password information is generated by the device, and then the device encrypts the user name, the random encryption key, and the encrypted password information of the client to the RADIUS server, and perform the related authentication process.

29.2 802.1X Configuration

29.2.1 Configure EAP

The 802.1x standard forwards the 802.1X authentication packets (Encapsulated with EAP frames) from the user to the RADIUS server without any processing. However, the traditional RADIUS server does not support the EAP feature. Therefore, the system supports the conversion of the authentication packets sent by the user to the data frames encapsulated by the standard

RADIUS protocol and then forwards the packets to the RADIUS server.

Configure EAP

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Set the protocol interaction mode between the system and the RADIUS server	**dot1x { eap-finish | eap-transfer }**	Optional eap-finish by default

29.2.2 Enable 802.1x

802.1x provides a user identity authentication scheme. However, 802.1x cannot implement the authentication scheme solely by itself. RADIUS or local authentication must be configured to work with 802.1x.

After enabling the 802.1X, the users who connected to the system can access to the LAN resources only after it had passed the authentication. When enabling the 802.1X, you should point out the whether the enabling way is based on interface authentication or MAC address authentication. The interface which does not participate in 802.1X authentication has no need to enable 802.1X authentication.

1) Interface configuration based on interface authentication: if one of the users under the port had passed the authentication, other users can use the network resources without authentication; However, if that user who had passed the authentication logoff, other users can not be able to use the network resources.

Interface configuration based on MAC address authentication: each user under the port should perform separate authentication. Only the user who had passed the authentication can he use the network resources. If a certain user logoff, it cannot affect other authenticated users to use the network resources.

Enable 802.1x

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable 802.1x	**dot1x method { macbased | portbased }**	required

	[*interface-list*]	

29.2.3 Configure 802.1x Parameters for a Port

After the interface enables the 802.1X authentication, this port needs to be authenticated by default while the uplink interface and the interface which connects to the server do not need, so you can configure the ports which do not need to be authenticated to be forceauthorized or disable their authentication functions. In addition, the interface which is banned to perform 802.1X authentication can be configured to be forceunauthorized.

Configure 802.1x Parameters for a Port

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure 802.1x parameters for a port	**dot1x port-control** { **auto** | **forceauthorized** | **forceunauthorized** } [*interface-list*]	optional

29.2.4 Re-authentication Configuration

In EAP-FINISH way, the port supports re-authentication. After the user is authenticated, the port can be configured to immediately re-certification, or periodic re-authentication.

re-authentication configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Immediately re-certification	**dot1x re-authenticate** [*interface-list*]	optional
Periodic re-authentication enabled on a port	**dot1x re-authentication** [*interface-list*]	optional
Periodic re-authentication time configuration port	**dot1x timeout re-authperiod** *time* [*interface-list*]	optional

29.2.5 Watch Feature Configuration

After enabling this function, a port sends a 1x watch message periodically when no user is present, triggering the following users to perform 802.1x authentication.

This triggering method is used to support clients that cannot send EAPOL-Start packets, such as 802.1X clients. Our device sends an EAP-Request / Identity packet to the client every N seconds to trigger authentication.

Watch Feature Configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable the watch function	**dot1x daemon** [*interface-list*]	optional
Configure the forwarding interval of watch packet	**dot1x daemon time** *time* [*interface-list*]	Optional 60S by default
Restore the default forwarding interval of watch packet	**no dot1x daemon time** [*interface-list*]	optional

29.2.6 Configure User Features

The operations mainly perform the operations, for example, the configurations for number of port users, delete users, heartbeat detection operations, etc.

Heartbeat detection: After this function is enabled, the device periodically forwards EAP-Request/Identity to the client ports, the normal online client responds with the EAP-Rsponse/Identity. If the four consecutive EAP-Request/Identity packets are not received the EAP-Rsponse/Identity packet from the client, the device considers the user to go offline, and then it will delete the session and change the port to an unauthorized state.

Quiesce function: After the user authentication fails, the device needs to quiesce for a period of time (The time can be configured through *dot1x quiet-period-value*. By default, no quiesced is required). During the quiesced period, the authenticator does not process the authentication request.

Configure User Features

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the maximum number of users that can	**dot1x max-user** *number*	optional

pass authentication		
Delete the specified online user	**dot1x user cut** { **username** *name* | **mac-address** *mac* }	optional
Enable heartbeat detection	**dot1x detect** [*interface-list*]	Optional 25s by default
Configure Heartbeat detection time	**dot1x detect interval** *time*	optional
Restore the default heartbeat detection time	**no dot1x detect interval**	optional
Configure the quiesce function	**dot1x quiet-period-value** *time*	Optional; 0 by default; No quiesce.
Restore the default quiet period value	**no dot1x quiet-period-value**	optional

29.2.7 Configure Host Mode Based on Port Authentication Mode

The host mode configuration only takes effect in port authentication method, please configure the port as port-based authentication; if the configuration of the host mode is the single-host, configure the port to be mac-based authentication, host mode will automatically become invalid.

（1） multi-hosts: Multi-hosts mode, when a user authentication is passed on the port, other users of the port can access network without authentication.

（2）single-host: Single-host mode, the user access network which the port allows only one authentication to pass and other users cannot access to the network, also can't go through authentication.

Configuration host-mode

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure host-mode based on port authentication mode	**dot1x portbased host-mode { multi-hosts | single-host }** [*interface-list*]	optional

29.2.8 Configure Guest VLAN

After enabling 1X authentication, the user can access only the network resources of the VLAN when the guest VLAN is configured on the port. Once the user authentication succeeds, the

port automatically reverts to the previously configured VLAN. If the authentication server delivers a valid VLAN, the port is automatically added to the assigned VLAN. After the user goes offline, the port reverts to the guest VLAN.

To ensure that all functions can be used normally, please assign different VLAN IDs for the Config VLAN, the radius distribution VLAN, and the Guest VLAN.

Guest VLAN configuration

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure Guest VLAN	**dot1x guest-vlan** *vlan-id* [*interface-list*]	optional

29.2.9 Configure Radius vlan

When 802.1X user pass the authentication via radius server, the server will transmit the authentication information to the device. If the device has enabled radius function and the server has configured to distribute VLAN (adopting Tunnel-Pvt-Group-ID (81) attribute), the authentication information will include the distributed VLAN information as a consequence, what is more, the device will add the user authentication online interface to radius distributed VLAN.

Configure Radius vlan

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter AAA configuration mode	**aaa**	
Enable radius vlan distribution function	**radius vlan enable**	Optional Disabled by default

Note:

Before using the radius vlan distribution function, you should create the corresponding VLAN and then add the user interface to the corresponding VLAN, so does Guest VLAN and Default-active-vlan；

Radius distributes VLAN, but it does not change the interface original VLAN configuration, so

does Guest VLAN and Default-active-vlan.

As to the interface-based authentication and the MAC-based authentication, radius vlan ，Guest VLAN and Default-active-vlan are effective.

29.2.10 Configure EAPOL Transmission

When a port disables 802.1x authentication, it requires to transmit user 802.1x EAPOL message. So the equipment will work as the relay, users can perform 802.1x authentication in the upper equipment. This function can only handle EAPOL packet forwarded to CPU. For packets that do not forward to CPU, the packets are processed by the hardware and are not subject to this configuration. You can configure EAPOL transparent transmission port and the corresponding uplink port only when the 802.1x authentication is disabled. That is, you can not configure transparent transmission function when the 802.1x authentication is enabled.

Configure EAPOL Transmission

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable port EAPOL message transmission function	**dot1x eapol-relay** [*interface-list*]	optional
Configure EAPOL message transmission uplink port	**dot1x eapol-relay uplink** [*interface-list*]	optional

29.2.11 Dot1x Display and Maintenance

Dot1x Display and Maintenance

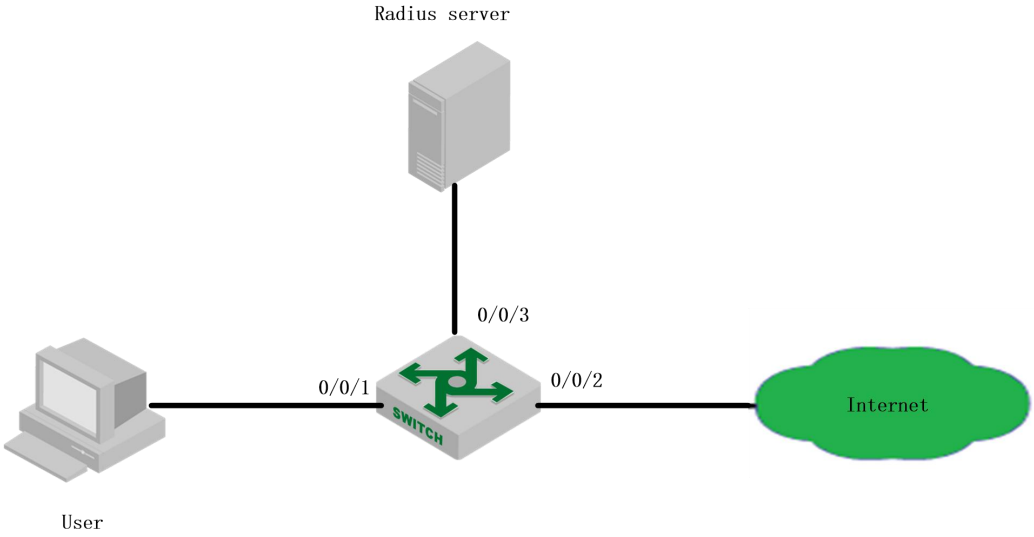
Operation	Command	Remarks
Display the status of 802.1X authentication function	**show dot1x**	
Display the configuration of 802.1x authentication interface watch function	**show dot1x daemon [interface** *interface-num* **]**	
Display interface configuration, such as	**show dot1x interface [** *interface-num* **]**	

the interface control mode, re-authentication state, the maximum number of users for the interface authentication.		
Display 802.1X session	**show dot1x session [{ interface** *interface-num* **} | { mac-address** *mac* **}]**	User online state (port number, VLAN ID, mac address, username, etc.)
Display EAPOL pass through configuration	**show dot1x eapol-relay [interface** *interface-num* **]**	
Display heartbeat detection configuration	**show dot1x detect [interface** *interface-num* **]**	
Display guest-vlan information	**show dot1x guest-vlan [interface** *interface-num* **]**	
Display whether the interface authentication is enabled or disabled	**show dot1x port-auth**	
Display quiet period	**show dot1x quiet-period-value**	
Debug DOT1X receive packet and transmit packet as well as module processing	**debug dot1x**	

29.3 Configuration Example

29.3.1 Networking Requirements

Local 802.1 x access user name is u1, and then password is 123. User can be able to access internet after login successfully. Network diagram are shown below:



network diagram of 802.1X configuration

29.3.2 Configuration steps

1) Enable the 802.1x authentication of Ethernet port 0/0/1

Switch(config)#dot1x method macbased interface ethernet 0/0/1

2) Configure the basic function of RADIUS server（create RADIUS 1，configure the master authentication server to be 1.1.1.1, primary accounting server to be 1.1.1.2, the authentication shared key and accounting shared key to be 123456. Please refer to 《Radius configuration 》for more RADIUS detailed configuration.)

Switch(config-aaa)#radius host 1

Switch(config-aaa-radius-1)#primary-auth-ip 1.1.1.1 1812

Switch(config-aaa-radius-1)#primary-acct-ip 1.1.1.2 1813

Switch(config-aaa-radius-1)#auth-secret-key 123456

Switch(config-aaa-radius-1)#acct-secret-key 123456

Switch (config-aaa)#domain abc.com

Switch (config-aaa-domain-abc.com)#radius host binding 1

Switch (config-aaa-domain-abc.com)#state active

Switch(config-aaa)#default domain-name enable abc.com

29.3.3 Result validation

User inputs the username and password on the 802.1X client to perform authentication. Through the command of “show dot1x session”, it shows the current user had passed the authentication and login successfully, that is to say, the user can be able to access the internet.

```
Switch(config)#show dot1x session
port       vid    mac                    username          login time
0/0/1    1       c8:3a:35:d3:e3:99 u1@abc.com    2000/01/01 05:13:42
```

30. RADIUS Configuration

30.1 Radius Overview

30.1.1 AAA Overview

AAA stands for *Authentication*, *Authorization* and *Accounting*.

AAA is actually a management of network security. Here, the network security mainly refers to the access control, including the users who can access the network server; what services are available to users with access rights; and how users are using network resources for billing.

AAA generally adopts the client / server structure: the client runs on the managed resource side, and the server stores the user information centrally. Therefore, the AAA framework has good scalability, and easy to achieve the centralized management of user information.

30.1.2 AAA Realization

AAA frame diagram is as shown in figure 1-1:

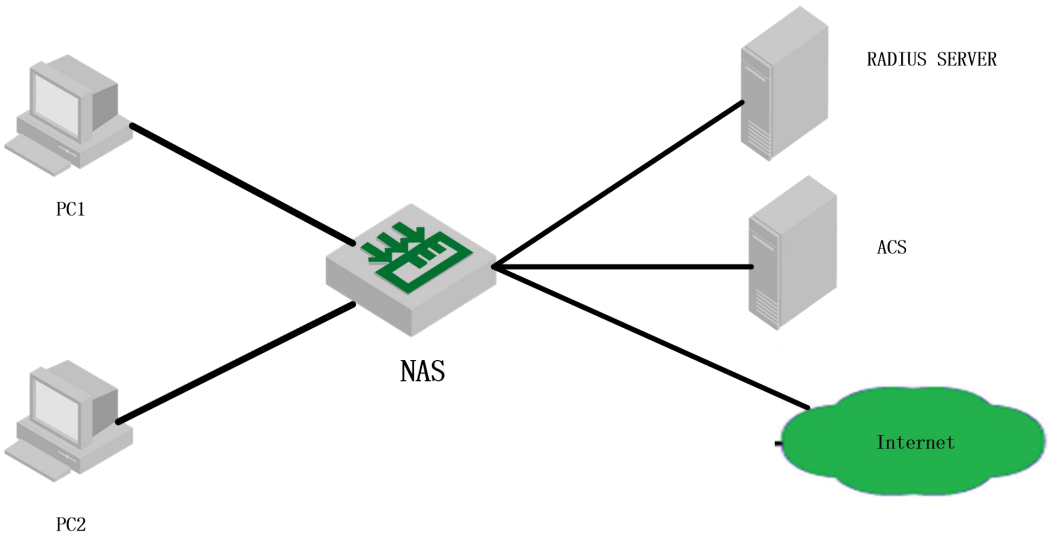


Figure 1-1 AAA frame diagram

There are two ways to realize AAA:

- via NAS;

- via R ADIUS, TACACS +, etc.

30.1.3 RADIUS Overview

RADIUS creates a unique user database, stores the user name and password of the user to authenticate, and stores the service type and corresponding configuration information that is passed to the user to complete the authorization. After the user is authorized, the RADIUS server performs the function of accounting for user accounts.

- RADIUS stands for Remote Authentication Dial in User Service.
- RADIUS is an AAA protocol for applications such as Network Access or IP Mobility.
- It works in both situations, Local and Mobile.
- It uses Password Authentication Protocol (PAP), Challenge Handshake Authentication Protocol (CHAP), or Extensible Authentication Protocol (EAP) protocols to authenticate users.
- It looks in text file, LDAP Servers, Database for authentication.
- After authentication services parameters passed back to NAS.
- It notifies when a session starts and stop. This data is used for Billing or Statistics purposes.
- SNMP is used for remote monitoring.
- It can be used as a proxy.

Here is a list of all the key features of Radius:

1. Client/Server Model

- NAS works as a client for the Radius server.
- Radius server is responsible for getting user connection requests, authenticating the user, and then returning all the configuration information necessary for the client to deliver service to the user.
- A Radius server can act as a proxy client to other Radius servers.

2. Network Security

- Transactions between a client and a server are authenticated through the use of a shared key. This key is never sent over the network.
- Password is encrypted before sending it over the network.

3. Flexible Authentication Mechanisms

- Point-to-Point Protocol - PPP
- Password Authentication Protocol - PAP
- Challenge Handshake Authentication Protocol - CHAP
- Simple UNIX Login

4. Extensible Protocol

Radius is extensible; most vendors of Radius hardware and software implement their own dialects.

30.2 RADIUS Configuration

30.2.1 RADIUS Server Configuration

RADIUS server saves valid user's identity. When authentication, system transfers user's identity to RADIUS server and transfers the validation to user. User accessing to system can access LAN resources only after authentication of RADIUS server.

Configure RADIUS server

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter AAA mode	**aaa**	-
Create and enter RAIDUS configuration schemes	**radius host** *name*	required
Configure primary RADIUS	**primary-auth-ip** *ipaddr port*	required
Configure second RADIUS	**second-auth-ip** *ipaddr port*	optional
Configure primary accounting server	**primary-acct-ip** *ipaddr port*	optional
Configure second accounting server	**second-acct-ip** *ipaddr port*	optional
Configure shared key of primary RADIUS	**auth-secret-key** *keystring*	required
Configure shared key of second RADIUS	**acct -secret-key** *keystring*	optional
Configure NAS-RAIDUS address	**nas-ipaddress** *ipaddr*	Optional If there is no

		configuration, the equipment IP address will also be OK.
Set whether the user name is to be carried with the domain name when the system passes the packet to the current RADIUS server	**username-format** { **with-domain** | **without-domain** }	optional
Configure the realtime accounting	**realtime-account**	optional
Configure the realtime accounting interval	**realtime-account interval** *time*	optional

30.2.2 Radius Master Server & Radius Slave Server Shift

RADIUS offers master/slave server redundancy function, that is: if both the master server and slave server can be able to perform the regular work, it can only perform the authentication via master server; if there is something wrong with the master server, the slave server will be enabled; if the master server recovers normal again, the slave server will be disabled, and then the master server will be enabled.

Realization Mechanisms:

When in radius authentication, if the master server cannot perform the regular work, just configure the master server as down, then the slave server will begin to work; if the master server is found had recovered the regular work, preemption timer will be enabled(time is configured as preemption-time). When the timer timeout, the master server will be configured as up, that is to say, you can perform the authentication operations via master server.

Radius Master Server & Radius Slave Server Shift

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter AAA configuration	**AAA**	-

mode		
Create and enter RAIDUS configuration schemes	**radius host** *name*	
Configure the preemption timer	**preemption-time** *Preemption-time*	Value range<0-1440>, the unit is minute; 0 by default, not preemption

30.2.3 Configure Local User

Client needs to configure local user name, password, etc.

Configure Local User

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter AAA mode	**AAA**	-
Configure local user	**local-user username** *name* **password** *pwd* **[vlan** *vid* **]**	optional

30.2.4 Configure Domain

Client needs to provide username and password during authentication. Username usually contains the corresponding user's ISP information, domain and ISP. The most important information of the domain is the RADIUS server authentication and accounting for the users in the domain.

Configure Domain

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter AAA mode	**aaa**	-
Configure the default domain- name	**default domain-name enable** *domain-name*	optional
Disable the default domain-name	**default domain-name disable**	
Create and enter a domain scenario	**domain** *name*	required
Configure to use radius server authentication	**scheme radius**	optional
Configure to use local user authentication	**scheme local**	
Configure to use local authentication after the radius authentication fails	**scheme radius loca**	
Select the RADIUS server for the current domain	**radius host binding** *radius-name*	optional
Enable the number limit of authentication users in the domain and set the number limit of allowed users	**access-limit enable** *number*	optional
Disable the number limit of authentication users in the domain	**access-limit disable**	
Activate the current domain	**state active**	required
Deactivate the current domain	**state block**	

30.2.5 Configure RADIUS Features

Configure RADIUS some compatible or special features as below:

Configure RADIUS features

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enter AAA mode	**aaa**	-
Configure accounting-on function	**accounting-on** { **enable** *sen-num* | **disable** }	optional
Configure H3C Cams compatibility	**h3c-cams** { **enable** | **disable** }	optional
Enable accounting function	**radius accounting**	optional
If the accounting packet does not respond, the user is shut down	**radius server-disconnect drop 1x**	optional
Configure RADIUS to distribute port priority	**radius 8021p enable**	optional
Configure RADIUS to distribute port PVID	**radius vlan enable**	optional
Configure RADIUS to distribute number limit of MAC address	**radius mac-address-number enable**	optional
Configure RADIUS to distribute bandwidth control	**radius bandwidth-limit enable**	optional

🕮 **Note:**

accounting-on: After the device reboots, it sends an Accounting-On packet to the RADIUS server to notify the RADIUS server to force the user of the device to go offline.

H3C Cams compatibility feature: In this feature, you can use the command of *radius attribute client-version* to forward the version information of the client to the RADIUS server. In this feature, you can use the command of *uprate-value / dnrate-value* to configure the attribute number of the upstream bandwidth / downstream bandwidth in the Vendor Specific.

RADIUS distributes port priority: After this function is enabled, if the user authenticates, the priority of the port where the user is located is modified. This function is carried out through the 77 attribute number in the Vendor Specific by default, which can be modified by using the *radius config-attribute.*

RADIUS distributes port PVID: After this function is enabled, if the user passes the authentication, the PVID of the port where the user is located will be modified. This function is carried out by using the *tunnel-Pvt-Group-ID*. The value of this attribute is a string. Use this string to find the VLAN name descriptor that matches the VLAN value.

RADIUS distributes number limit of MAC address: After this function is enabled, if the user passes the authentication, the MAC address learning limit of the port where the user resides is modified. This function is carried out through the 50 attribute number in the Vendor Specific by

default, which can be modified by using the *radius config-attribute.*

RADIUS distributes bandwidth control: After this function is enabled, if the user passes the authentication, the bandwidth control of the port where the user is located will be modified. The uplink bandwidth control is carried out through the 75 attribute number in the Vendor Specific by default, which can be modified by using the*radius config-attribute;* the downlink bandwidth control is carried out through the 76 attribute number in the Vendor Specific by default, which can be modified by using the *radius config-attribute.* The unit value defaults to **kbps** and can be modified through the *radius config-attribute access-bandwidth unit*.

RADIUS distributes ACL: This function has no control commands. It is enabled by default. Configure via 11 attributes of Filter-Id.

30.2.6 RADIUS Display and Maintenance

RADIUS Display and Maintenance

Operation	Command	Remarks
Display the radius attribute	**show radius attribute**	-
Display the radius attribute	**show radius config-attribute**	-
Display the radius service configuration information	**show radius host** *hostname*	
Enable the radius debugging function	**debug radius**	

30.3 RADIUS Configuration Example

30.3.1 Configure the networking and requirements

As shown below, user PC is connected to Switch 0/0/1 port, Switch 0/0/4 port is connected to radius server (radius server integrated with Windows 2003), and 802.1x authentication is enabled on 0/1.

Specific requirements are as follows:

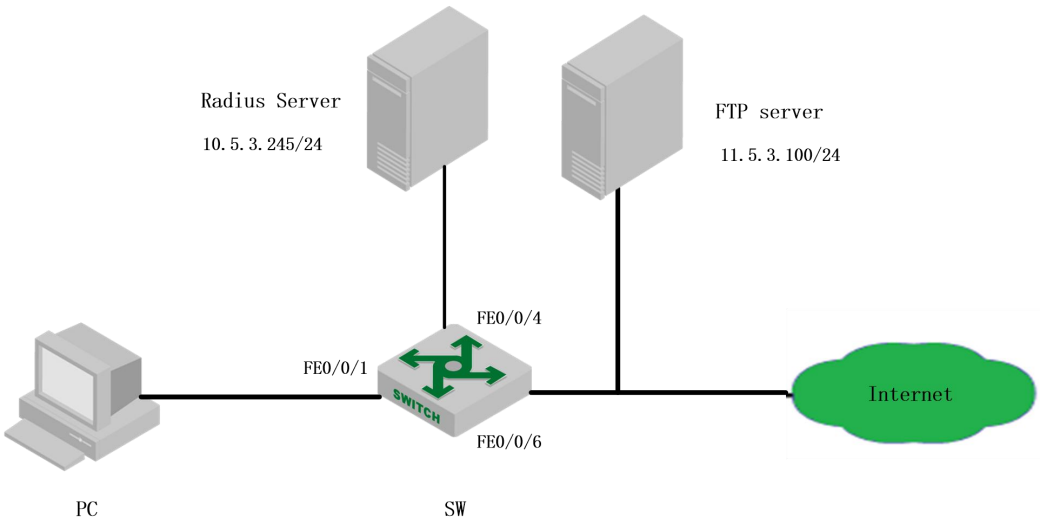
1. Use radius authentication；

2. The user PC must be authenticated before accessing the internet；

3. After the user passes the authentication, the ACL is distributed through the radius server.

In this case, the user can access the Internet but cannot access the FTP server;

4. After the user passes the authentication, distribute the bandwidth control via the RADIUS server to limit the uplink bandwidth to be 2M and the downstream bandwidth to be 1M.



networking diagram for radius configuration example

30.3.2 Configuration steps

一、initial preparation work:

1）. Install the 802.1X client on the PC, here adopts H3C Inode;

2）. Switch configuration user interface IP10.5.3.235 / 24 to ensure to PING radius server;

Switch(config-if-vlanInterface-1)#interface vlan-interface 1

Switch(config-if-vlanInterface-1)#ip address 10.5.3.235 255.255.255.0

This ipaddress will be the primary ipaddress of this interface.

Config ipaddress successfully!

Switch(config-if-vlanInterface-1)#

Switch(config-if-vlanInterface-1)#

Switch(config-if-vlanInterface-1)#exit

Switch(config)#ping 10.5.3.254

PING 10.5.3.254: with 32 bytes of data:

reply from 10.5.3.254: bytes=32 time<10ms TTL=128

reply from 10.5.3.254: bytes=32 time<10ms TTL=128

----10.5.3.254 PING Statistics----

2 packets transmitted, 2 packets received, 0% packet loss

round-trip (ms) min/avg/max = 0/0/0

Control-C

3）. radius server adds NAS IP, and the shared key is 123456;

4）.Configure the 802.1x client authentication username (test) and password (123456) on the radius server.

5）. The attribute value of the 75 attribute in the Vendor Specific on the radius server is set to 2048 Kbps, and the attribute value of the 76 attribute in the Vendor Specific is set to 1024 Kbps.

6）. The attribute value of the 11 attribute of the Filter-Id on the radius server is set to 100；

二、Access the switch 0/0/1 port to enable dot1x, configure the related service of RADIUS, and configure ACLs

Switch(config)#dot1x method portbased interface ethernet 0/0/1 // enable 802.1X

Switch(config)#aaa

Switch(config-aaa)#radius host ngn

Switch(config-aaa-radius-ngn)#primary-auth-ip 10.5.3.254 1812 // Configure accounting function, authentication IP, and port number

Switch(config-aaa-radius-ngn)#primary-acct-ip 10.5.3.254 1813

Switch(config-aaa-radius-ngn)#auth-secret-key 123456 // Configure to share the key

Switch(config-aaa-radius-ngn)#acct-secret-key 123456

Switch(config-aaa-radius-ngn)#exit

Switch(config-aaa)#radius bandwidth-limit enable // Enable the bandwidth sending function

Switch(config-aaa)#domain ngn.com

Switch(config-aaa-domain-ngn.com)#radius host binding ngn

Switch(config-aaa-domain-ngn.com)#state active

Switch(config-aaa-domain-ngn.com)#exit

Switch(config-aaa)#default domain-name enable ngn.com

Switch(config)#access-list 100 deny any 11.5.3.100 0.0.0.255 // Configure the ACL to deny access to the destination network segment

Switch(config)#access-list 100 permit any any

30.3.3 Result validation

Use the Inode client on the PC, and then enter the user name and password for authentication

After the authentication succeeds, the user can access the external network normally. The information of the online users can be found on the Switch. The command of *show dot1x radius-acl* displays the status of the *acl100* as enable, and the bandwidth of the ingress direction of the 0/ 0/1 port is limited to 2048 while the egress direction is limited to 1024.

```
Switch(config)#show dot1x session
port       vid    mac                    username          login time
e0/0/1     1      c8:3a:35:d3:e3:99      test@ngn.com      2000/12/11 15:07:00
Total [1] item(s).

Switch(config)#show dot1x radius-acl
The format of radius acl is string.
The prefix of radius acl is assignacl-.
Port       acl    Status
e0/0/1     100    enable
Total entries: 1.

Switch(config)#show bandwidth-control interface ethernet 0/0/1
port       Ingress bandwidth control    Egress bandwidth control
e0/0/1     2048 kbps                    1024 kbps
Total entries: 1.
```

31.Port Security Configuration

31.1 Port Security Overview

Port security is generally applied at the access layer. It can restrict host to access to the network through the device, and allow certain hosts to access the network, while other hosts cannot access the network.

The port security function binds the user's MAC address, IP address, VLAN ID, and PORT number flexibly, and prevents illegal users from accessing the network. This ensures the security of network data and the legal users can obtain sufficient bandwidth.

Users can restrict the hosts that can access the network through three rules: MAC rule, IP rule, and MAX rule. MAC rules are divided into three binding methods: MAC binding, MAC + IP binding, MAC + VID binding; the MAX rule defines the maximum number of MAC addresses that can be learned on a port. This address does not include the number of MAC rules and IP rules generated by the legitimate MAC address. In the MAX rule, there are sticky rules. If the deny rule is only configured on the port and the MAX rule is not configured, the other messages cannot be forwarded (Exception by allowing rule checking).

The MAC address of the Sticky rule can be learned automatically, and configured manually and saved in the running configuration file. If the configuration file is saved before the device reboots, the device does not need to be configured again after the device reboots, and these MAC addresses take effect automatically. When the sticky function is enabled on the port, the dynamic MAC address learned by the MAX rule is added to the sticky rule and saved to the running configuration file. In the case of the MAX rule is not full, it is allowed to continue learning the new MAC address and form the sticky rule until the number of sticky rules reaches the maximum configured by MAX.

MAC rules and IP rules can specify whether messages matching the corresponding rules are allowed to communicate. The user’s MAC address and VLAN, MAC address and IP address can be bound flexibly by the MAC rule. Because port security is software-based, the number of rules is not limited by hardware resources, make the configuration more flexible.

The rules of port security are triggered by the ARP messages of the terminal device. When the device receives an ARP message, port security extracts various messages information, and match with the three rules of the configuration. The order of match is MAC address, IP address and MAC rule. The Layer 2 forwarding table of the port is controlled by the matching result, in order to control the forwarding behavior of the port.

When the port security judgment message is illegal, messages are processed accordingly. There are three modes: protect, restrict and shutdown. Protect mode discards messages. The restrict mode discards messages and trap alarms (Receive an illegal message in two minutes of the alarm). Shutdown mode will shut down port in addition to restrict mode of action.

31.2 Port Security Configuration

Configure port security

operation	command	remark
Enter the port configuration mode	interface ethernet *port-number*	required
Enable/disable port security	port-security { enable | disable }	required
Configure MAC binding rule	[no] port-security { permit | deny } mac-address ***mac-address*** **{** **[** vlan-id ***vlan-id***] | ip-address ***ip-address*** }	optional
Configure IP rules	[no] port-security { permit | deny } ip-address ***start-ip*** [to ***end-ip***]	optional
Configure MAX rules	[no] port-security maximum ***value***	optional
Enable STICKY	[no] port-security permit mac-address sticky	optional
Configure MAC STICKY rules	[no] port-security permit mac-address sticky ***mac-address*** **[** vlan-id ***vlan-id***]	optional
Configure the address aging time	[no] port-security aging time ***value***	optional
Enable static address aging function	[no] port-security aging static	optional
Configure the policy for receiving invalid message	port-security violation { protect | restrict | shutdown }	optional
Enable shutdown automatic recovery	[no] port-security recovery	optional
Configure the automatic recovery time after shutdown	[no] port-security recovery time ***value***	optional
Delete the currently active MAC address	no port-security active-address {all | configured | learned }	optional
Delete all the port security related configurations	no port-security all	optional
Display the security configuration	show port-security [interface list]	optional
Display the MAC rule configuration	show port-security mac-address [interface ethernet *port-number*]	optional
Display the IP rule configuration	show port-security ip-address [interface ethernet *port-number*]	optional
Display the currently active MAC address	show port-security active-address [configured | learned | interface ethernet *port-number*]	optional
Display the configuration of automatic recovery after shutdown	show port-security recovery [interface ethernet ***port-number***]	optional

Note:

1. After the port-security function is enabled, deny all messages by default. Therefore, user must configure one of the mac\ip\max rules.

2. If the sticky function is effective, it is necessary that the port security is enabled, and the number of MAX rule isn't configured for 0. When this function is turned on, the dynamic addresses learned in the previous MAX rule are converted to STICKY rules and stored in the run

file. When the function is disabled, the learned STICKY rules are deleted. The number of STICKY rule entries of a port cannot exceed the configured number of MAX rules. If the configuration file is saved before the device reboots, the STICKY rule saved before the port reboots will take effect. When the port is shutdown, there are two ways to recovery: (1) configure the port for shutdown and no shutdown. (2) Automatic recovery after configuring shutdown.

4. When illegal message is received, trap alarms do not take effect immediately. Traps are generated within two minutes.

5. If a MAC address or ip address is denied, through the upper limit of MAX doesn't reach, the host can't communication.

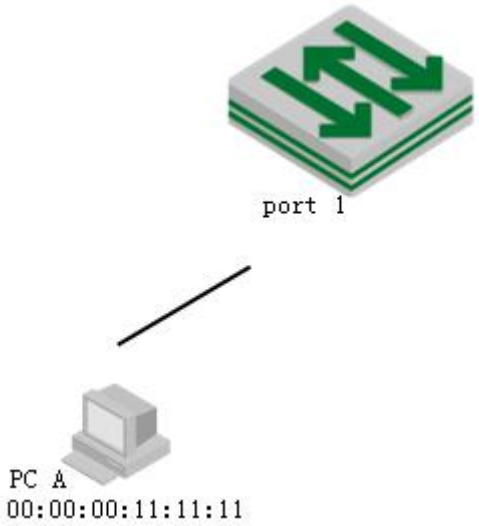
6. Port security cannot be enabled together with 802.1X or mac authentication.

7. Port security cannot be enabled together with anti-ARP flooding.

31.3 Port Security Configuration Example

I. Network requirements

Configure port 1 to allow only pc A communication;



Port security diagram

2. Configuration procedure

Configure port security

DUT(config)#interface ethernet 0/0/1

DUT(config-if-ethernet-0/0/1)#port-security enable

DUT(config-if-ethernet-0/0/1)#port-security permit mac-address 00:00:00:11:11:11

3. Verify results

(1) Using ixia emulation PC A, configure two network cards, all through DHCP to obtain IP, configure dhcp-snooping (configuration slightly) on the DUT, access to IP as follows:

DUT(config)#show dhcp-snooping clients

DHCP client information:

d - days, h - hours, m - minutes, s - seconds

IPAddress mac vlan port LeaseTime ExceedTime
192.168.1.100 00:00:00:11:11:11 1 e0/0/1 1d0h0m0s 23h51m21s
192.168.1.101 00:00:00:54:20:71 1 e0/0/1 1d0h0m0s 23h55m37s
Total entries: 2. Printed entries: 2.
2) Use the DUT to ping the two clients separately, obtain the ARP entry, and enable the DUT to establish the port security activation table.
DUT(config)#show dhcp-snooping clients
DHCP client information:
d - days, h - hours, m - minutes, s - seconds
IPAddress mac vlan port LeaseTime ExceedTime
192.168.1.100 00:00:00:11:11:11 1 e0/0/1 1d0h0m0s 23h51m21s
192.168.1.101 00:00:00:54:20:71 1 e0/0/1 1d0h0m0s 23h55m37s
Total entries: 2. Printed entries: 2.

Display the currently active MAC addresses. Only the permit mac rule entries are displayed
DUT(config)#show port-security active-address
Acitve mac-address:
Port MAC address VID IP Addr Derivation Action Age(min)
E1/0/1 00:00:00:11:11:11 1 192.168.1.100 MAC permit 1
Total entries: 1
DUT(config)#debug port-security
DUT(config)#logging monitor 0

(3) Try to communicate with the DUT using two PCs, respectively: The results are as follows
Use the ip = 192.168.1.100 (mac = 00: 00: 00: 11: 11: 11 match port-security rule) to ping the DUT. It can communicate, log is as follows:
00:29:48: DUT: %PORT-SECURITY-7-debug: port e0/0/1 recv packet mac[00:00:00:11:11:11] vlan [1] type[0x0806]
00:29:48: DUT: %PORT-SECURITY-7-debug: match with MAC RULE
00:29:48: DUT: %PORT-SECURITY-7-debug: action: PERMIT

Use the ip = 192.168.1.101 (mac = 00: 00: 00: 54: 20: 71 match port-security rule) to ping the DUT. It can communicate, log is as follows:
00:30:07: DUT: %PORT-SECURITY-7-debug: port e0/0/1 recv packet mac[00:00:00:54:20:71] vlan [1] type[0x0806]
00:30:07: DUT: %PORT-SECURITY-7-debug: match with MAX RULE
00:30:07: DUT: %PORT-SECURITY-7-debug: port e0/0/1 maxnum exceed
Maxnum rule by default is 0, so exceed, the message is discarded;

32.SNTP Client

32.1 SNTP Overview

Switch system time can be achieved in two ways, one is as sntp client, sntp server automatically synchronizes time; the other is the administrator own configuration.

The Simple Network Time Protocol (SNTP) is used for time synchronization between network devices. Normally, an SNTP server exists in the network and provides reference time for multiple SNTP clients. In this way, time synchronization is achieved among all network devices.

SNTP can work in four modes: unicast, broadcast, multicast, and anycast.

In the unicast mode, the client initiates a request to the server. After receiving the request, the server constructs a response message based on the local time and sends the response message back to the client.

In the broadcast and multicast mode, the server periodically sends broadcast or multicast messages to the client, and the client receives the messages from the server.

In the anycast mode, the client initiates a local broadcast address or a multicast address to send a request. In this case, the server in the network responds to the client. The client selects the server that receives the response message as the server, and discards the messages sent by the other server. After electing out of the server, the work pattern is same as unicast.

In all modes, the client receives a response message to parse the message to obtain the current standard time, and calculates the network transmission delay and local time compensation through a certain algorithm. The data is used to calibrate the current time.

32.2 Configure the SNTP Client

32.2.1 Enable/disable SNTP Client

Enable/disable SNTP client

operation	command	remark
Enter the global configuration mode	configure terminal	-
Enable/disable the SNTP client	[no]sntp client	Required, the default is off
Display the SNTP client configuration	show sntp client	Optional
View the system time	show clock	Optional

32.2.2 Configure the Work Mode of the SNTP Client

According to the network, administrators can use commands to modify the way of SNTP work - unicast, broadcast, multicast, or anycast.

Configure the Work Mode of the SNTP Client

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the Work Mode of the SNTP Client	sntp client mode { **broadcast** | **unicast** | **multicast** | **anycast [key *key-id*]** }	Optional, Default is broadcast mode
Display the SNTP client configuration	show sntp client	Optional

32.2.3 Configure the SNTP Server Address

When an SNTP client works in the unicast mode, user must configure the specified SNTP server.

Configure the SNTP Server Address

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the SNTP server address	[no]sntp server ***ip-address***	required
Configure the SNTP backup server	[no]sntp server backup ***ip-address***	optional
Display the SNTP client configuration	show sntp client	optional

32.2.4 Modify the Broadcast Transmission Delay

When the SNTP client works in the broadcast or multicast mode, it is necessary to use the broadcast transmission delay parameter. In the broadcast mode, the local system time of the SNTP client is equal to the time taken from the server plus the transmission delay. Administrators can modify the broadcast transmission delay based on the actual bandwidth of the network.

Configure the broadcast transmission delay

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the broadcast propagation delay	[no]sntp client broadcastdelay ***value***	Optional. The default value is 3ms
Display the SNTP client configuration	show sntp client	optional

32.2.5 Configure the Polling Interval

User needs to configure the polling interval when the SNTP client works in the unicast or anycast mode. The SNTP client initiates a request to the server every other polling interval to calibrate the local system time.

Configure the Polling Interval

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure the polling interval	[no]sntp client poll-interval ***value***	Optional. The

		default is 1000
Display the SNTP client configuration	show sntp client	optional

32.2.6 Configure Timeout Retransmission

Because the SNTP request message is a UDP message, it cannot guarantee that the request message can reach the destination. The timeout retransmission mechanism is adopted. The configured timeout interval is required when the SNTP client works in the unicast or anycast mode. When the client sends a request within a certain period of time without receiving a response, It will be re-sent the request until the number of retransmission exceeds the set value. The configured timeout retransmission mechanism takes effect only when the SNTP client works in the unicast or anycast mode.

Configure the timeout retransmission attempts and time interval

operation	command	remark
Enter the global configuration mode	configure terminal	
Configure the timeout retransmission interval	[no]sntp client retransmit-interval ***value***	Optional. Default 5s
Set the timeout retransmission attempts	[no]sntp client retransmit ***value***	Optional. Default 0
Display the SNTP client configuration	show sntp client	optional

32.2.7 Configure the Client Daylight Saving Time

Configure the Client Daylight Saving Time

operation	command	remark
Enter the global configuration mode	configure terminal	-
Configure daylight saving time	[no]sntp client summer-time dayly **{ *start-month start-day start-time end-month end-day end-time* }** [no]sntp client summer-time weekly { ***start-month start-week*** **[Fri |mon| sat | sun | thu | tue | wed**] ***start-time end-month end-week*** [**Fri | mon | sat | sun | thu | tue | wed**] ***end-time*** }	optional
Display the daylight saving time configuration	show sntp client summer-time	optional

32.2.8 Configure Legacy Server List

When an SNTP client works in broadcast or multicast mode, it will trust and receive the protocol messages from any SNTP server. If there is a malicious attack on the network server (which provides the wrong time), the local time cannot be synchronized to standard time.

After the list of valid servers is configured on the SNTP client, the client can only receive messages whose source addresses are in the legal server list, thus improve the security.

Configure legal server list

operation	command	remark
Enter the global configuration mode	configure terminal	-

Configure legal server list	sntp client valid-server ***ipaddress wildcard***	optional
Delate legal server list	no sntp client valid-server { **all** | ***ipaddress wildcard*** **}**	optional
Display the SNTP client configuration	show sntp client	optional

32.2.9 Configure Authentication

To further improve security, user can enable MD5 authentication between the SNTP server and the client. The SNTP client receives only authenticated messages. The authentication configuration is as follows:

Configure Authentication

operation	command	remark
Enter the global configuration mode	configure terminal	-
Switch certification	[no] sntp client authenticate	Optional, close by default
Configure the password for authentication	[no]sntp client authentication-key ***key-number*** md5 ***value***	optional
Configure a trusted password ID	[no]sntp trusted-key ***key-number***	Optional, For multicast and broadcast mode only, it must be equal to authentication-key
Configure the password ID used by the server	[no]sntp server key ***key-number***	Optional, it must be equal to authentication-key
Configure the password ID for anycast configuration	**sntp client mode anycast key** *key-number*	optional
Display the SNTP client configuration	show sntp client	optional

32.2.10 Manual Calibration of the System Clock

In addition to switch, which acts as sntp client automatically synchronize time from the sntp server, the other way is the administrator manual calibration system clock.

If the switch has built-in lithium battery, the switch power is off, the system clock runs normally; if there is no built-switch lithium battery, the switch power is off, the system clock stops running.

Configure system clock

operation	command	remark
Enter execution configuration mode	-	-
Configure the system clock	clock set ***HH:MM:SS YYYY/MM/DD***	required
Enter the global configuration mode	configure terminal	optional
Configure the system time zone	[no]clock timezone ***zone-name hours-offset minutes-offset***	optional
Configure daylight saving time	[no]sntp client summer-time dayly **{** ***start-month start-day start-time end-month end-day end-time*** } [no]sntp client summer-time weekly { ***start-month start-week*** **[Fri |mon| sat | sun | thu | tue | wed**] ***start-time end-month end-week*** [**Fri | mon | sat | sun | thu | tue | wed**] ***end-time*** }	optional

View the system time	show clock	optional

Example

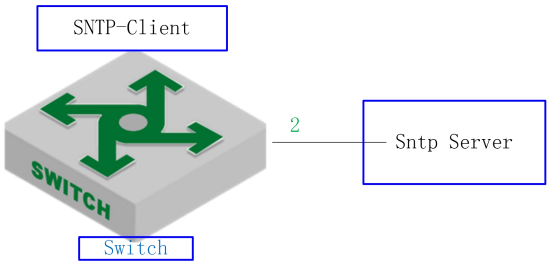
Configure the system clock
Switch#clock set 17:50:50 2015/11/25
Set clock successfully.
Clock will be reset to 2013/01/01 00:00:00 after system rebooting because there is no realtime clock chip.

Switch#show clock

Wed 2015/11/25 17:51:03 CCT 08:00

32.2.11 SNTP Client Configuration Example

1. Networking Requirements
The switch acts as the sntp client to synchronize time from the sntp server.
Make sure that the switch communicates properly with the sntp server.



Sntp client diagram

2. Configuration steps
#switch runs the broadcast, multicast, unicast, and anycast modes respectively.

Configuration of the authentication mode and broadcast mode
Enable the sntp client and configure it in broadcast mode (the broadcast mode is enabled by default).
Switch(config)#sntp client mode broadcast

Configure the trusted server (not configurable)
Switch(config)#sntp client valid-server 192.168.1.99 0.0.0.0
Switch(config)#sntp client

Configure the authentication key (make sure the configuration is consistent with that of the sntp server)

Switch(config)#sntp client authentication-key 1 md5 test
Configure the trusted key ID
Switch(config)#sntp trusted-key 1
Enable the authentication function
Switch(config)#sntp client authenticate

View the time synchronization result
Switch(config)#show sntp client
  Clock state   : synchronized   Current mode   : broadcast
  Use server   : 192.168.1.99   State   : idle
  Server state : synchronized   Server stratum   : 1
  Authenticate : enable   Bcast delay   : 3ms
  Last synchronized time: THU NOV 26 06:07:44 2015
  Summer-time is not set.
  Valid server list:
    Server address:192.168.1.99   wildcard:0.0.0.0

Multicast mode and authentication configuration
Enable the sntp client and configure it as the multicast mode
Switch(config)#sntp client mode multicast
Configure the trusted server list
Switch(config)#sntp client valid-server 192.168.1.99 0.0.0.0
Switch(config)#sntp client

Configure the authentication key (make sure the configuration is consistent with that of the sntp server)
Switch(config)#sntp client authentication-key 1 md5 test
Configure the trusted key ID
Switch(config)#sntp trusted-key 1
Enable the authentication function
Switch(config)#sntp client authenticate

View the switch time synchronization result
Switch(config)#show sntp client
  Clock state   : synchronized   Current mode   : multicast
  Use server   : 192.168.1.99   State   : idle
  Server state : synchronized   Server stratum   : 1
  Authenticate : enable   Bcast delay   : 3ms
  Last synchronized time: THU NOV 26 06:20:59 2015
  Summer-time is not set.
  Valid server list:
    Server address:192.168.1.99   wildcard:0.0.0.0

Configure unicast mode and authentication

Enable the sntp client and configure it as unicast
Switch(config)#sntp client
Switch(config)#sntp client mode unicast
Configure the sntp server
Switch(config)#sntp server 192.168.1.99

Configure the authentication key (make sure the configuration is consistent with that of the sntp server)
Switch(config)#sntp client authentication-key 1 md5 test
Configure the password ID for the server
Switch(config)#sntp server key 1
Enable the authentication function
Switch(config)#sntp client authenticate

View the time synchronization result
Switch(config)#show sntp client

Clock state : synchronized Current mode : unicast
Use server : 192.168.1.99 State : idle
Server state : synchronized Server stratum : 1
Retrans-times: 3 Retrans-interval: 30s
Authenticate : disable PrimaryServer: 192.168.1.99
Backup Server: 0.0.0.0 Poll interval : 1000s
Last synchronized time: THU NOV 26 09:05:29 2015
Last received packet's originateTime: TUE JAN 01 00:00:24 2013
Summer-time is not set.

Configure the authentication mode for anycast mode
Enable the sntp client and configure it to work in anycast mode
Switch(config)#sntp client mode anycast
Configure the sntp server (may not configure)
Switch(config)#sntp server 192.168.1.99
Switch(config)#sntp client

Configure the authentication key (make sure the configuration is consistent with that of the sntp server)
Switch(config)#sntp client authentication-key 1 md5 test
configure anycast mode and the key ID (if the authentication is not necessary, you don’t configure)
Switch(config)# sntp client mode anycast key 1

Enable the authentication function
Switch(config)#sntp client authenticate

View the time synchronization result
Switch(config)#show sntp client
  Clock state   : synchronized   Current mode   : anycast
  Use server   : 192.168.1.99   State   : idle
  Server state : synchronized   Server stratum   : 1
  Retrans-times: 3   Retrans-interval: 30s
  Authenticate : enable   Authentication-key: 1
  Poll interval   : 1000s
  Last synchronized time: THU NOV 26 09:22:25 2015
  Last received packet's originateTime: THU NOV 26 17:22:24 2015
  Summer-time is not set.

33.Link Backup Function

33.1 Flex Links

33.1.1 Flex links Overview

Flex links is a two-layer link backup protocol that provides options for STP. If you do not want STP on your network, you can select Flex links to implement link backup. If STP is enabled, you do not need flex links. The Flex links consist of a pair of ports, which can be physical or aggregated. When one port forwards data, the other port is in the standby state. When the primary link fails, the backup port starts to forward data. After the faulty port recovers, the status changes to backup. If the preemption mechanism is set, the status changes to forwarded data after 60 seconds. The STP must be disabled on the Flex links port. The Flex links ports can be configured with bandwidth and delay as the preemption mechanism, the condition precedence as the master port. There must be a trap alarm when the active / standby link fails.

Flex links have the following advantages:

1）When two links in a dual uplink network are normal, only one is in the connected state and the other is in the blocked state, thus preventing broadcast storms caused by the loop.

2） When the fault occurs on the primary link, the traffic will be switched to the standby link within the millisecond time to ensure the normal forwarding of data.

Configuration is simple, user-friendly operation.

- **Flex Links Group**

Flex Links group is also called flexible link group. A Flex Links group consists of two member ports, one of which is designated as the master port and the other as the slave port. A port can belong to only one Flex Links group. Normally, only one port (master port or slave port) is in the forwarded (ACTIVE) state, the other port is blocked (BLOCK), in standby (STANDBY) state. When a link fault (Currently, the fault of the link mainly refers to the state of the port being changed to DOWN, the Ethernet OAM link fault and so forth) occurs on a port which is in the forwarding state, the Flex Links group automatically blocks the port and switches the original blocked port which is in standby sate to the forwarding state.

1） Master Port

Master Port is a port role of the Flex Links group specified through the command line. The Master Port of the Flex Links group can be an Ethernet port (electrical or optical), or an aggregation group.

2） Slave Port

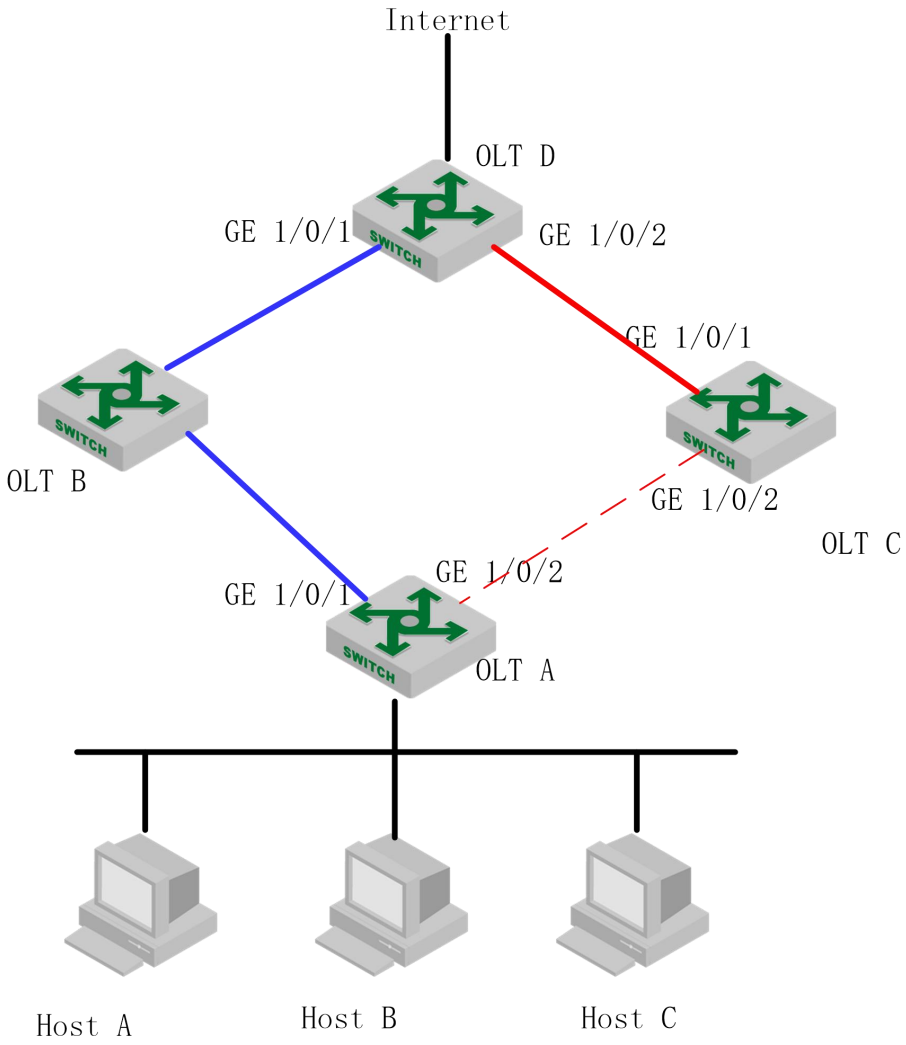
Slave Port is another port role of the Flex Links group specified through the command line. The Slave port of the Flex Links group can be an Ethernet port (electrical or optical), or an aggregation group. The link where the Slave port resides is also called the "backup link".

➢ **MMU Packet**

MMU is MAC address-table Move Update，MAC address drift；When the link is switched in the Flex Links group, the original MAC forwarding entries are no longer applicable to the new topology network. Therefore, the MAC address forwarding entries on the entire network need to be updated. Flex Links uses the MMU packet to notify other devices to refresh the MAC entries.

➢ **How Flex Links works**

Take the networking shown in Figure 1-1 as an example, describe the working mechanism of Flex Links according to the process of link normal-> link fault-> link recovery.



➢ **Link Normal Working Mechanism**

GE 1/0/1 of SwitchA is the master port and GE 1/0/2 is the slave port. When both uplinks are normal, the master port is in the forwarding state, and the link is the active link. The slave port is in the standby state and the link is the standby link. The data is transmitted

along the link represented by the blue lines. There is no loop in the network to avoid broadcast storms.

- **Link Fault Processing Mechanism**

When the active link on Switch A fails, GE 1/0/1 switches to the standby state and GE 1/0/2 switches to the forwarding state. In this case, the MAC address forwarding entries on the devices in the network may be incorrect. In this case, you need to provide a mechanism to update the MAC address so that the traffic can be quickly switched to avoid loss of traffic. Currently, this mechanism is used to prevent the traffic from being lost by sending MMU packets to notify the device to update the MAC entries.

This method is applicable to upstream devices (such as Switch B, Switch C, and Switch D in Figure 1-1) that support Flex Links and can identify MMU packets.

To implement fast link switching, you need to enable MMU packet transmission on Switch A, and enable the function of receiving and processing MMU packets on all the ports on the dual uplink network of the upstream device.

After the link switchover occurs on Switch A, the MMU packet is sent from the new active link. That is, the MMU packet is sent from GE1 / 0/2. When receiving an MMU packet, Switch C processes the MMU packet and sends the MAC address of the packet to GE 1/0/2 of Switch C.

After that, if Switch D receives the data packets destined for Host A, Host B, and Host C, Switch D forwards the packets through Layer 2 broadcast. After receiving the packet, Switch C looks up the MAC address table and forwards it to Switch A from GE 1/0/2, and finally forwarded by Switch A to Host A, Host B and Host C.

The mechanism of updating a device through MMU packets does not need to wait for the entries to be aged before they can be updated. This greatly reduces the time required for table entries to be updated. In general, the entire switching process of the link can be completed in milliseconds, with no traffic lost.

- **Link Recovery Processing Mechanism**

The Flex Links group supports three modes: role preemption mode, non-role preemption mode, and bandwidth preemption mode. The link recovery mechanism is different in different modes:

If the Flex Links group is configured as role preemption mode, after the active link fault is recovered, the master port will preempt the forwarding state and the slave port will enter the standby state. Only when the active link fails, the slave port will switch from the standby state to the forwarding state.

If the Flex Links group is configured as non-role preemption mode, after the active link fault is recovered, the slave port will continue to be in the forwarding state and the master port will remain in the standby state. This ensures that the traffic is stable.

If the Flex Links group is configured as the bandwidth preemption mode, if the bandwidth of

the slave port is higher after the active link fault is recovered, the slave port will remain in the forwarding state and the master port will remain in the standby state. If the bandwidth of the master port is higher, the master port preempts the forwarding state and the slave port enters the standby state.

As shown in Figure 1-1, after the link of GE 1/0/1 on Switch A is recovered, if the Flex Links group is configured as role preemption mode, GE1 / 0/2 will block and switch to the standby state, while GE1 / 0/1 will preempt the forwarding state. If the Flex Links group is configured as non-role preemption mode, GE1 / 0/1 will remain in the standby state and traffic will not be switched to maintain stable traffic. If the Flex Links group is configured as the bandwidth preemption mode, If the bandwidth of GE 1/0/2 is higher, GE 1/0/2 preempts the forwarding state and GE 1/0/1 continues to be in the standby state. If the bandwidth of GE1 / 0/1 is higher, GE 1/0/1 preempts the forwarding state and GE 1/0/2 is in the standby state.

33.1.2 Configure Flex Links Group

To configure the Flex Links group, you only need to specify the master port and the slave port. If the master port is an Ethernet port, you need to enter port mode to configure it. If the master port is an aggregation group, you can configure it in global mode.

Configure Flex Links Group

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Configure the Flex Links group	**channel-group** *cgroup ID1* **backup { interface** *port* **| channel-group** *cgroup ID2* **}**	*cgroup ID1* is the master port *cgroup ID2* is the slave port
Delete the Flex Links group	**no channel-group** *cgroup ID* **backup**	Optional
Enter the port configuration mode	**interface Ethernet** *port-num*	-
Configure the Flex Links group	**switchport backup { interface** *port2***| channel-group** *cgroup ID2* **}**	*Port 2/cgroup ID2* is the slave port
Delete the Flex Links group	**no switchport backup**	Optional

Note: The spanning tree of the master and slave ports must be down and not the ERRP port.

33.1.3 Configure Flex Links Preemption Mode

The Flex Links group supports three preemption modes: role preemption mode, non-role preemption mode, and bandwidth preemption mode. The default is non-role preemption mode.

Configure Flex Links Preemption Mode

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Configure the priority preemption mode for Flex Links	**channel-group** *cgroup ID* **backup { interface** *port2* **| channel-group** *cgroup ID2* **} preemption mode { Forced|Bandwidth|Off }**	*Cgroup ID1* 为主端口，*port_2/cgroup ID2* 为副端口
Enter the port configuration mode	**interface Ethernet** *port-num*	
Configure the priority preemption mode for Flex Links	**switchport backup { interface** *port2* **| channel-group** *cgroup***} preemption mode { Forced|Bandwidth|Off }**	*port2/cgroup ID2* 为副端口

33.1.4 Configure the Delay Time For Priority Preemption of Flex Links

After the configuration of the Flex priority preemption mode, the port will not immediately become the forwarding (ACTIVE) state. Instead, it takes a period of time to preempt the priority of the port to forward packets. The default time is 45 seconds.

Configure the Delay Time for Priority Preemption of Flex Links

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Configure the delay time for priority preemption of Flex Links	**channel-group** *cgroup ID1* **backup { interface** *port2* **| channel-group** *cgroup ID2* **} preemption delay** <1-60>	*Cgroup ID1* is the master port，*port_2/cgroup ID2* is the slave port
Enter the port configuration mode	**interface Ethernet** *port-num*	
Configure the priority preemption mode for Flex Links	**switchport backup { interface** *port2* **| channel-group** *cgroup ID2* **} preemption mode <Forced|Bandwidth|Off >**	*port2/cgroup ID2* is the slave port

33.1.5 Configure Flex Links MMU Function

When a link is switched in the Flex Links group, the original MAC forwarding entries are no longer applicable to the new topology network. The MAC address forwarding entries on the entire network need to be updated. Flex Links uses the MMU packet to notify other devices to refresh the MAC entries. The default is not to enable this function.

Configure Flex Links MMU Function

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Configure the MMU function for Flex Links	**mac-address-table move update { transmit | receive }**	*Cgroup ID1* is the master port，*port2/cgroup ID2* is the slave port

33.1.6 Flex Links Display and Maintenance

After completing the above configuration, you can run the following command to check the configuration.

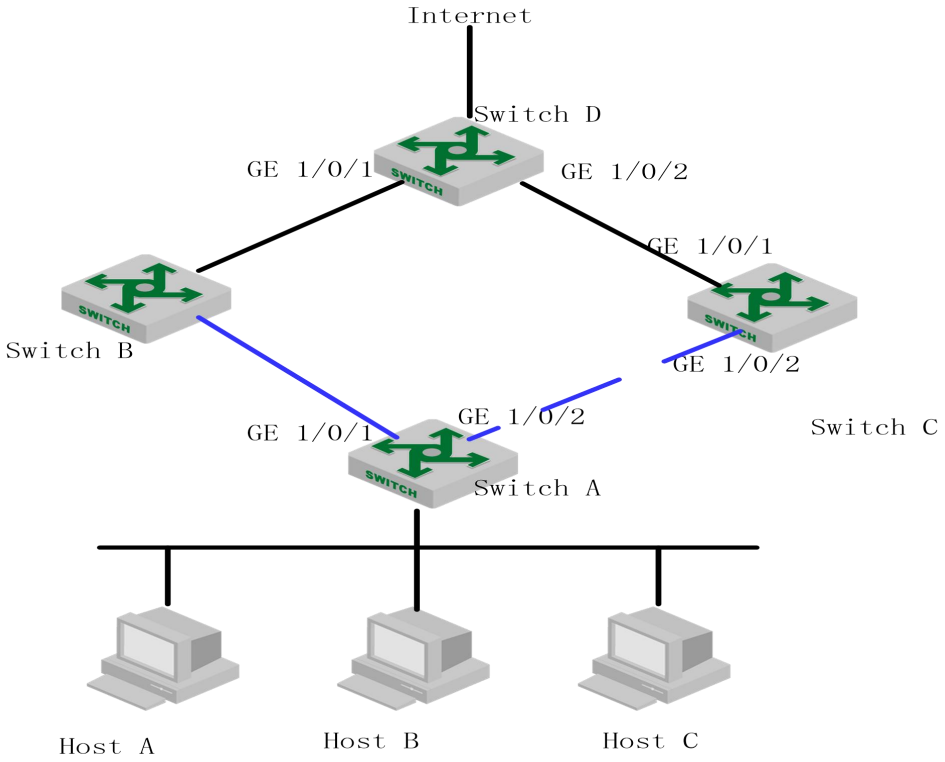
Flex Links Display and Maintenance

Operation	Command	Remarks
Display the configured Flex Links groups	**show interface switch backup**	All modes are executable
Display the MMU status of Flex Links and the number of sent and received packets	**sh mac-address-table move update**	

33.2 Monitor Link

33.2.1 Monitor Link Overview

➢ **Monitor Link Background**



As shown in Figure 2-1, Switch A is configured with the Flex Links function for link redundancy backup. GE 1/0/1 is the master port and GE 1/0/2 is the slave port.

When the active link on GE1 / 0/1 fails, traffic is switched to the standby link on GE1 / 0/2 in milliseconds. This ensures efficient and reliable link backup and fast convergence.

However, when the link on uplink port GE1 / 0/1 of Switch B fails, the link on GE1 / 0/1 of Switch A, which is a device in the Flex Links group, is not faulty. So there will be no link switching within the Flex Links group. In fact, traffic on Switch A can not be forwarded to Switch D through the link of GE 1/0/1, thus the traffic is interrupted. To solve this problem, Monitor Link technology come into being.

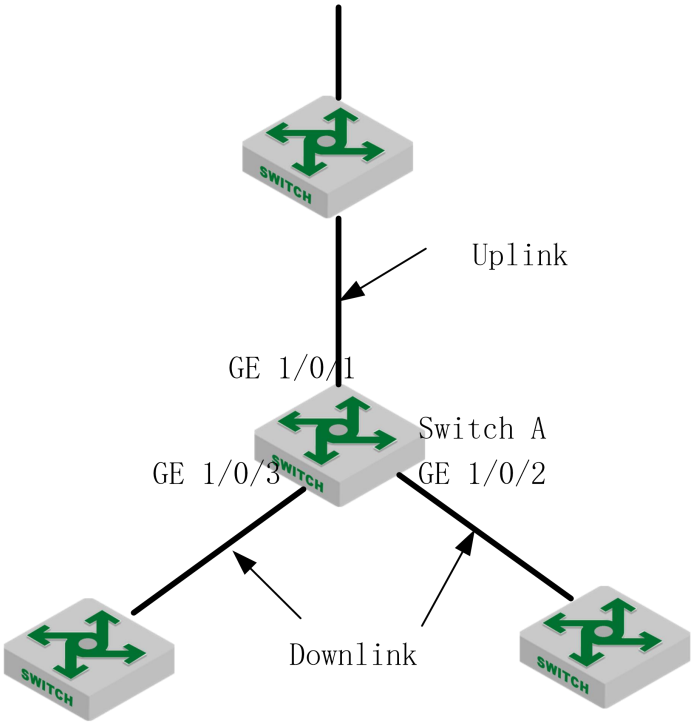
- **Technical Advantages**

Monitor Links is a powerful complement to Flex Links technology. It achieves a purpose of synchronizing the downlink with the uplink through monitoring the uplink.This status make the back-up role of Flex Links more perfect.

- **Monitor Link Basic Concept**

Monitor Link Group

Monitor Link Group consists of one or more uplink and downlink ports. The status of the downlink port varies with the status of the uplink port.



As shown in Figure 2-2, GE1 / 0/1, GE1 / 0/2, and GE1 / 0/3 of Switch A form a Monitor Link group.

Uplink Port

Uplink Port is a monitor object in the Monitor Link group and is a port role of the Monitor Link group specified through the command line. The uplink port of a Monitor Link group can be an Ethernet interface (electrical or optical) or an aggregation group interface.

As shown in Figure 2-2, GE 1/0/1 of Switch A is the uplink port of the monitor link group configured on the device.

If multiple ports are configured as the uplink ports of the monitor link group, the status of the monitor link group is UP as long as one of the ports is in the forwarding state; only when all the uplink ports fail, the status of the monitor link group is DOWN, and all downlink ports will be shut down. When the uplink port of the monitor link group is not specified, the uplink port is considered as faulty and all the downlink ports are shut down.

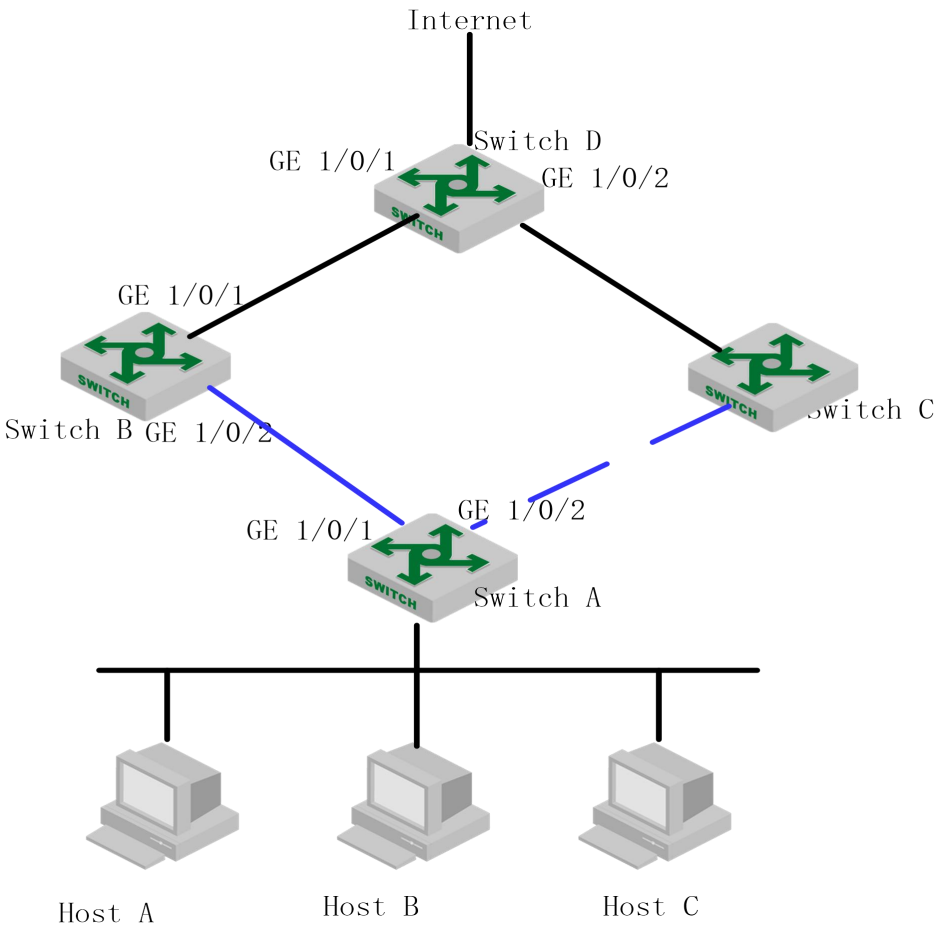
Downlink Port

Downlink port is the monitor in the Monitor Link group. It is another port role of the Monitor Link group specified through the command line. The downlink port of a Monitor Link group can be an Ethernet interface (electrical or optical) or an aggregation group interface.

As shown in Figure 2-2, GE1/0/2 and GE1/0/3 of Switch A are the two downlink ports of the monitor link group configured on the device.

When the uplink port of the monitor link group is restored to normal, the monitor link will only open the downlink port blocked by the uplink port failure, and can not enable the manually closed downlink port. And a downlink port failure has no effect on the uplink port and other downlink ports.

- **How Monitor Link Works**



As shown above, a flexible link group is configured on Switch A to ensure reliable access to the Internet. In the figure, GE 1/0/1 is the master port and is in the forwarding state. GE 1/0/2 is the slave port.

To prevent the traffic on Switch A from failing due to a link fault on GE 1/0/1 of Switch B, configure the monitor link group on Switch B and specify GE 1/0/1 as the uplink port, GE1 / 0/2 is the downlink port.

When the link on the uplink port GE 1/0/1 of Switch B fails, the monitor link group forcibly shuts down the downlink port GE 1/0/2 of the group, triggering the link switching of the Flex Link group on Switch A.

When the link failure occurred on the uplink port GE1/0/1 of Switch B recovers, the downlink port GE1/0/2 is also enabled. If the Flex link group on Switch A is configured with the role preemption mode, the link switching of the Flex Link group on Switch A will be triggered. Otherwise, the Switch will wait for the next link switching.

In this way, Monitor Link technology with Flex Links technology achieves a highly efficient and reliable link backup and fast convergence.

33.2.2 Configure Monitor Link Group

If the port is an Ethernet interface, you need to enter the port mode to configure it. If the port is an aggregation group interface, you can configure it in global mode.

Configure Monitor Link Group

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Configure the Monitor Link function of the aggregation group	**channel-group** *cgroup ID* **monitor-link-group** *group-ID* **{ uplink | downlink }**	-
Remove the aggregation group from the monitor link group	**no channel-group** *cgroup ID* **monitor-link-group** *group-ID* **{ uplink | downlink }**	optional
Enter the port configuration mode	**interface Ethernet** *port-num*	
Configure the monitor link function on a port	**switchport monitor-link-group** *group-ID* **{ uplink | downlink }**	optional
Remove the port from the Monitor Link group	**switchport monitor-link-group** *group-ID* **{ uplink | downlink }**	optional

33.2.3 Monitor Link Display and Maintenance

After you complete the above configuration, you can use the following command to view the configuration.

Monitor Link Display and Maintenance

Operation	Command	Remarks
Display the monitor link group configured	**show monitor-link-group**	All modes are executable

33.2.4 Configuration Example for Flex Links & Monitor Link

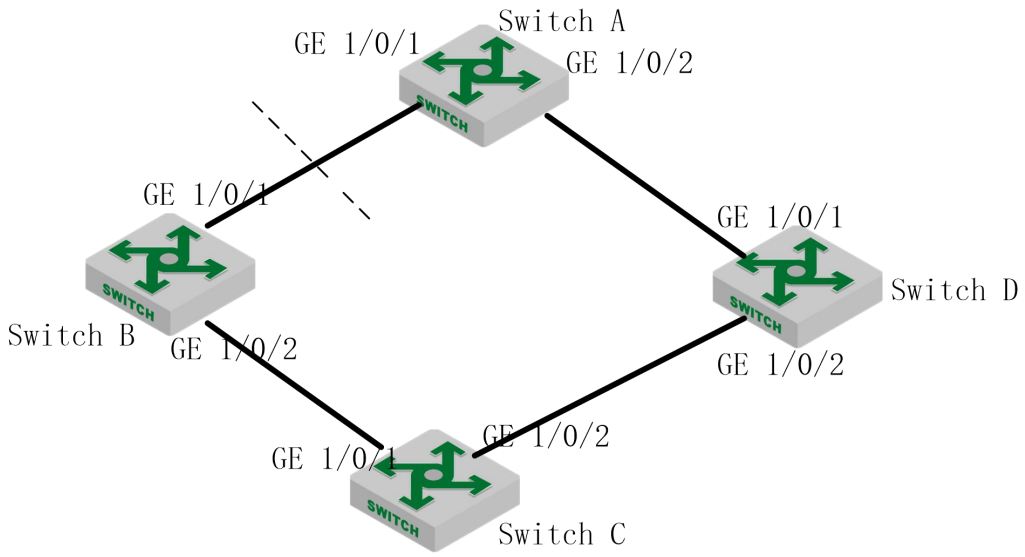


图 33-1

1. Network requirements

In the networking shown in Figure 2-4, Switch C is a Flex Links device, and Switch A, Switch B, and Switch D are related devices. Traffic on Switch C is dual-uplinked to Switch A through the Flex Links group. With this configuration, Switch C can flexibly back up dual uplinks, and when the link between Switch A and Switch B fails, Switch C can sense the fault and switch its uplink.

2. Configuration steps

Configure Switch C

Disable STP on GE1 / 0/1 and GE1 / 0/2, and configure the ports as trunk ports.

Switch-C(config)#interface range ethernet 1/0/1 ethernet 1/0/2

Switch-C(config-if-range)#no spanning-tree

Switch-C(config-if-range)#switchport mode trunk

Switch-C(config-if-range)#exit

Configure the Flex Links group with GE1 / 0/1 as the master port and GE1 / 0/2 as the slave port. and configure the Flex Links group as the role preemption mode. The delay time is 5s.

Switch-C(config-if-ethernet-1/0/1)#switchport backup interface ethernet 1/0/2

Switch-C(config-if-ethernet-1/0/1)#switchport backup interface ethernet 1/0/2 preemption mode forced

Switch-C(config-if-ethernet-1/0/1)#switchport backup interface ethernet 1/0/2 preemption delay

5
Switch-C(config-if-ethernet-1/0/1)#exit

Enable the function of sending MMU packets.
Switch-C(config)#mac-address-table move update transmit

Display Flex Links information
Switch-C(config)#show interface switchport backup
ActiveInterface BackupInterface State
---
e1/0/1 e1/0/2 active up /backup Standby
Preemption mode: Forced
Preemption Delay: 5 seconds

Total record 1.

Switch-C(config)#show mac-address-table move update
Dst mac-address: : 01:80:c2:00:00:10
Default/Current settings: : Rcv Off/Off,Xmt Off/On
Rcv Count: : 0
Xmt Count: : 0

Configure Switch A
Configure GE1 / 0/1 and GE1 / 0/2 as trunk ports and enable the function of receiving MMU packets.
Switch-A(config)#interface range ethernet 1/0/1 ethernet 1/0/2
Switch-A(config-if-range)#switchport mode trunk
Switch-A(config-if-range)#exit
Switch-A(config)#mac-address-table move update receive

Configure Switch B
Configure GE1 / 0/1 and GE1 / 0/2 as trunk ports and enable the function of receiving MMU packets.
Switch-B(config)#interface range ethernet 1/0/1 ethernet 1/0/2
Switch-B(config-if-range)#switchport mode trunk
Switch-B(config-if-range)#exit
Switch-B(config)#mac-address-table move update receive

Configure GE 1/0/1 as the uplink port of Monitor Link group 1 and GE 1/0/2 as the downlink port of Monitor Link group 1.
Switch-B(config)#interface ethernet 1/0/1
Switch-B(config-if-ethernet-1/0/1)#switchport monitor-link-group 1 uplink
Switch-B(config-if-ethernet-1/0/1)#exit
Switch-B(config)#interface ethernet 1/0/2

Switch-B(config-if-ethernet-1/0/2)#switchport monitor-link-group 1 downlink
Switch-B(config-if-ethernet-1/0/2)#exit

Display Monitor Link information
Switch-C(config)#show monitor-link-group
Monitor-link Group

Group 1:
UplinkID UplinkStatus
e1/0/1 UP
DownlinkID DownlinkStatus
e1/0/2 UP

Configure Switch D
Configure GE1 / 0/1 and GE1 / 0/2 as trunk ports and enable the function of receiving MMU packets.
Switch-D(config)#interface range ethernet 1/0/1 ethernet 1/0/2
Switch-D(config-if-range)#switchport mode trunk
Switch-D(config-if-range)#exit
Switch-D(config)#mac-address-table move update receive

Configure GE 1/0/1 as the uplink port of Monitor Link group 1 and GE 1/0/2 as the downlink port of Monitor Link group 1.
Switch-DB(config)#interface ethernet 1/0/1
Switch-D(config-if-ethernet-1/0/1)#switchport monitor-link-group 1 uplink
Switch-D(config-if-ethernet-1/0/1)#exit
Switch-D(config)#interface ethernet 1/0/2
Switch-D(config-if-ethernet-1/0/2)#switchport monitor-link-group 1 downlink
Switch-D(config-if-ethernet-1/0/2)#exit

Display Monitor Link information
Switch-D(config)#show monitor-link-group
Monitor-link Group

Group 1:
UplinkID UplinkStatus
e1/0/1 UP
DownlinkID DownlinkStatus
e1/0/2 UP

When a link fault occurs between Switch A and Switch B, display the information about Flex Links and Monitor Link on Switch B:
Switch-B(config)#show monitor-link-group
Monitor-link Group

Group 1:

UplinkID UplinkStatus

e1/0/1 DOWN

DownlinkID DownlinkStatus

e1/0/2 DOWN

Switch-B(config)#show mac-address-table move update

Dst mac-address: : 01:80:c2:00:00:10

Default/Current settings: : Rcv Off/On,Xmt Off/Off

Rcv Count: : 0

Xmt Count: : 0

Display the information about Flex Links and Monitor Link on SwitchC:

Switch-C(config)#show interface switchport backup

ActiveInterface BackupInterface State

---

e1/0/1 e1/0/2 active Standby /backup up

Preemption mode: Forced

Preemption Delay: 5 seconds

Total record 1.

Switch-C(config)#show mac-address-table move update

Dst mac-address: : 01:80:c2:00:00:10

Default/Current settings: : Rcv Off/Off,Xmt Off/On

Rcv Count: : 0

Xmt Count: : 1

Display the information about Flex Links and Monitor Link on Switch D:

Switch-D(config)#show monitor-link-group

Monitor-link Group

Group 1:

UplinkID UplinkStatus

e1/0/1 UP

DownlinkID DownlinkStatus

e1/0/2 UP

Switch-D(config)#show mac-address-table move update

Dst mac-address: : 01:80:c2:00:00:10

Default/Current settings: : Rcv Off/On,Xmt Off/Off

Rcv Count: : 1

Xmt Count: : 0

When a link is restored between Switch A and Switch B, the master port GE 1/0/1 of Switch C preempts the forwarding state after 5 seconds.

Display the information about Flex Links and Monitor Link on Switch B:

```
Switch-B(config)#show monitor-link-group
Monitor-link Group
-------------------------------------------------
Group 1:
UplinkID            UplinkStatus
e1/0/1                 UP
DownlinkID          DownlinkStatus
e1/0/2                 UP

Switch-B(config)#show mac-address-table move update
Dst mac-address:                  : 01:80:c2:00:00:10
Default/Current settings: : Rcv Off/On,Xmt Off/Off
Rcv Count:                        : 1
Xmt Count:                        : 0
```

Display the information about Flex Links and Monitor Link on Switch C:

```
Switch-C(config)#show interface switchport backup
ActiveInterface    BackupInterface      State
------------------------------------------------------------
e1/0/1                  e1/0/2                 active up /backup Standby
Preemption mode: Forced
Preemption Delay: 5 seconds

Total record 1.

Switch-C(config)#show mac-address-table move update
Dst mac-address:                  : 01:80:c2:00:00:10
Default/Current settings: : Rcv Off/Off,Xmt Off/On
Rcv Count:                        : 0
Xmt Count:                        : 2
```

Display the information about Flex Links and Monitor Link on Switch D:

```
Switch-D(config)#show monitor-link-group
Monitor-link Group
-------------------------------------------------
Group 1:
UplinkID            UplinkStatus
e1/0/1                 UP
DownlinkID          DownlinkStatus
```

e1/0/2 UP

Switch-D(config)#show mac-address-table move update
Dst mac-address: : 01:80:c2:00:00:10
Default/Current settings: : Rcv Off/On,Xmt Off/Off
Rcv Count: : 1
Xmt Count: : 0

33.3 VPRB Configuration

33.3.1 VPRB Overview

VPRB (vlan port redundancy backup) is a VLAN port backup function. If there are multiple ports in VLAN, you can specify one port as the master port and the other port as the backup port of the master port. If the master and backup ports are normal, the master port is in the forwarding state and the backup port is in the discarding state. All the packets of the vlan are sent from the master port. When the master port fails and packets cannot be forwarded, the backup port is immediately set to the forwarding state. The packets forwarded on the master port are switched to the backup port. If the master port returns to normal, the service packets are immediately switched back to the master port, so as to achieve the purpose of vlan port backup.

Reasonably planning the master and backup ports of different VLANs can achieve link load balancing.

To implement port backup and link load balancing for multiple VLANs, VPRB needs to be used with MSTP. First, add the VLAN to be backed up to the MSTP instance, and then configure the master port and backup port of the MSTP instance.

33.3.2 Configure Basic MSTP

Configure Basic MSTP

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the MSTP instance	[no] spanning-tree mstp instance ***in-id*** [vlan ***vlan-id*** | priority ***value***]	required
Run MSTP mode	spanning-tree mode mstp	required

Note:

33.3.3 Configure vprb

Configure vprb

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure vprb	[no] vprb major-port ethernet ***port-id*** bak-port ethernet ***port-id*** instance ***inst-id***	required
View the configuration information	show vprb	optional

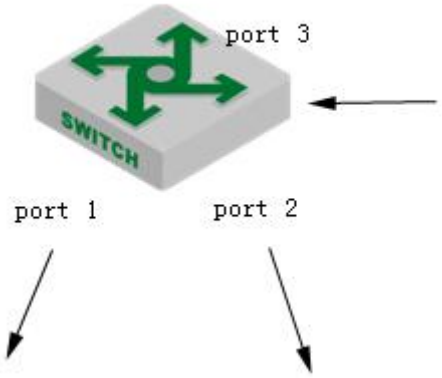
33.3.4 Configuration Example

1.Network requirements

The traffic of Vlan 1-200 enters from port 3.

Normally, the traffic of vlan 1-100 is forwarded from port1. When port 1 fails to become linkdown, the traffic of vlan 1-100 is forwarded from port 2; when port 1 recovers linkup, it is forwarded from port 1.

Normally, the traffic of vlan 101-200 is forwarded from port2. When port 2 fails to become linkdown, the traffic of vlan 101-200 is forwarded from port 1; when port 2 recovers linkup, it is forwarded from port 2.



sketch map of VPRB

2.Configuration steps

Configure the vlan

SW(config)#vlan 1-200

SW(config-if-vlan)#switchport range ethernet 0/0/1 to ethernet 0/0/3
SW(config-if-vlan)#exit

Configure the mstp instance for vprb
SW(config)#spanning-tree mst instance 1 vlan 1-100
SW(config)#spanning-tree mst instance 2 vlan 101-200

Configure stgp to run mstp mode
SW(config)# spanning-tree mode mstp

Configure the vprb function
SW(config)#vprb major-port ethernet 0/0/1 bak-port ethernet 0/0/2 instance 1
SW(config)#vprb major-port ethernet 0/0/2 bak-port ethernet 0/0/1 instance 2

3.Result validation
View the vprb information
SW(config)#s vprb

major-port	bak-port	instance
e0/0/1	e0/0/2	1
e0/0/2	e0/0/1	2

Total entries: 2 .

34. PPPoE Plus

34.1 Overview for PPPoE Plus

Point-to-Point Protocol over Ethernet (PPPoE) is a network tunneling protocol that encapsulates Point-to-Point Protocol (PPP) in an Ethernet frame. Because PPP is integrated in the protocol, it can implements the functions of authentication, encryption and compression that cannot be provided by traditional Ethernet. It can also be used for protocol architecture that provides access services for users by Ethernet protocol, such as cable modem and digital subscriber line (DSL) and so forth.

The PPPoE Plus function means that the physical information of the user side (the connected port, the VLAN where it resides, the MAC address of the local Switch and so forth) is added to the Sub-tag field in the PPPoE protocol packet by the Switch directly connected to the end-user. In this way, the authentication server can read the information to know the location of the user in the network in order to manage, maintain and service users.

Note that this function requires a server that supports PPPoE Plus to work with.

34.2 PPPoE Plus Configuration

34.2.1 Enable/disable PPPoE Plus

By default, the pppoe plus function is disabled. When you need to use the function, you must configure it as follows:

a. Enable the pppoe plus function on the port;

b. The port connecting to the pppoe-server is configured as a trusted port (all ports are untrusted by default).

Enable/disable PPPoE Plus

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Enable/disable the function	[no] pppoeplus	Required The default is off
Configure the uplink port as a trusted port	[no] pppoeplus trust	Required The default is untrust
View the configuration	show pppoeplus interface ethernet ***port-number***	optional

information		

34.2.2 Configure the Option Processing Strategy

If the PPPOE downlink port is directly connected to a PC or an Switch with PPPoE disabled, the DUT receives the PPPOE packet with no options. The DUT processes the packet according to the standard. If the downlink port is connected to an Switch with PPPoE enabled, the received PPPOE packet may already contain the option contents. In this case, the administrator needs to specify how the DUT handles the option contents. Allow three processing strategies: drop, keep and replace, the default use the replace strategy.

Configure the Option Processing Strategy

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure the option processing strategy	[no] pppoeplus strategy { **drop** | **keep** | **replace** }	Optional The no command recovers the default strategy :replace
View the configuration information	show pppoeplus interface ethernet ***port-number***	optional

34.2.3 Discard padi/pado Packet

In some specific cases, you may not want the port to process the received PPPOE padi / pado packets. The DUT provides the drop function and is disabled by default.

Configure to Discard padi / pado Packets

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-number***	-
Configure the option processing strategy	[no] pppoeplus drop { **padi** | **pado** }	Optional The default is processing
View the configuration information	show pppoeplus interface ethernet ***port-number***	optional

Note:

Most devices do not have this configuration command, indicating that they do not support this function. Please refer to the actual configuration.

34.2.4 Configure the Packet Type

The PPPOE packet needs to be added with the option contents before the packet is forwarded. the option contents can be determined in a number of ways.

You can specify the option contents in port mode.

If no content is specified, the configuration is performed according to the configuration rules, and the type is configured using the pppoe plus type. When the type is self-defined, you need to determine whether the format is in binary format or text format. The text format also needs to determine the connection symbol format.

The Switch provides three modes for PPPoE Plus packets：

a. Standard mode: User-side information includes the connected port, VLAN, and local Switch MAC. Encoding is as follows:

“0 0/0/0:4096.VID Switch MAC/0/0/slot/subslot/port”

b. HuaWei mode: Support connecting with HuaWei BRAS. User-side information includes the connected port, VLAN, the hostname of local Switch, and local Switch MAC. Encoding is as follows:

“0 0/0/0:4096.VID Switch MAC/hostname/0/slot/sub-slot/port”

c. Self-defined mode: Support user-defined message format.

Configure the Packet Type

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure the packet type	[no]pppoeplus type **{ huawei | standard |** self-defined **{** circuit-id **{ [** ***circuit-string*** **] [vlan] [port] [switch-mac] [hostname] [client-mac] } |** remote-id **{ [** ***remote-string*** **] [switch-mac] [hostname] [client-mac] } }}**	Optional The no command recovers the default type:standard
Configure the format	[no]pppoeplus format { **binary** | **ascii** }	Optional The no command recovers the default format: binary
Configure the delimiter	[no]pppoeplus delimiter { **colon** | **dot** | **slash** | **space** }	Optional The no command recovers the default

		configutation: space
Enter the port configuration mode	interface ethernet ***port-number***	-
Configure the circuit-id	[no] pppoeplus circuit-id ***circuit-string***	Optional The default is no configuration
View the configuration information	show pppoeplus interface ethernet ***port-number***	Optional

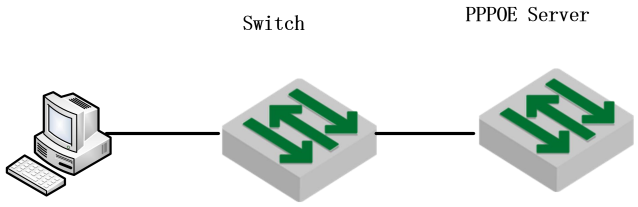
Note:

If both global and port are configured with circuit-id, the port takes precedence.

34.2.5 Configuration Example

1.Network requirements

The pppoe+ function is enabled on the Switch and is configured with self-defined type. Configure the circuit-id and remote-id contents, and verify that the PC pppoe dial-up is successful and that the options in the pppoei packet are correct.



sketch map of pppoe+

2.Configuration steps

Enable the pppoe plus function on the port connected to the client PC；

switch(config)#interface ethernet 0/0/1

switch(config-if-ethernet-0/0/1)#pppoeplus

The port connected to the server is configured as a pppoe plus trusted port；

switch(config-if-ethernet-0/0/1)#interface ethernet 0/0/3

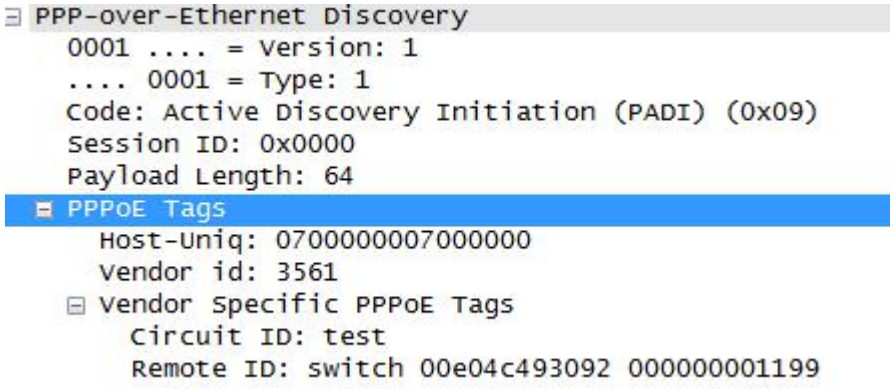
switch(config-if-ethernet-0/0/3)#pppoeplus trust

Configure self-defined type and configure the circuit-id and remote-id contents

switch(config-if-ethernet-0/0/3)#exit

switch(config)#pppoeplus type self-defined circuit-id test
switch(config)#pppoeplus type self-defined remote-id hostname client-mac switch-mac
switch(config)#pppoeplus format ascii

3.Result validation
(1) PC through pppoe dialing success；
(2) The circuit-id and remote-id fields of the mirrored trust port and padi packets are consistent with the configuration:



35.File Upload and File Download

35.1 Overview for File Download

File download is to download files from the external to the DUT's flash, such as the upgrade file (host file, bootrom file), the configuration file, and the ssh key file.

Host file name suffix must be .arj; bootrom file name suffix must be .bin; the configuration file name suffix must be .txt; the ssh key file name suffix must be .txt.

Download tools include xmodem, tftp, ftp.

When using the xmodem tool, after entering the command,, select "Send" -> "Send File" in the HyperTerminal menu. In the "Send File" dialog box, enter the full path and file name of the file in the File Name field. Select Xmodem from the Protocol drop-down list, and then click Send.

When an external file is downloaded to the DUT, it is saved in the flash memory and does not take effect immediately. You need to use the related configuration commands. After upgrading the host and bootrom, you need to restart the DUT. When you download the configuration file, it will overwrite the original configuration file in flash. You need to use the downloaded configuration file in the privilege mode: "copy startup-config running-config". Refer to the ssh module user manual for key usage.

35.1.1 Configure file download

Configure file download

Operation	Command	Remarks
privilege configuration mode	-	-
Upgrade the master host file	xmodem: load application xmodem	optional
	tftp: load application tftp inet[6] ***server-ip xxx.arj***	required
	ftp: load application ftp inet[6] ***server-ip xxx.arj*** grn 123	optional
Upgrade the backup host file	ftp: load application ftp inet[6] ***server-ip xxx.arj*** grn 123	required
Upgrade the bootrom file	xmodem: load whole-bootrom xmodem	optional
	tftp: load whole-bootrom tftp inet[6] *server-ip xxx.bin*	required
	ftp: load whole-bootrom ftp inet[6] ***server-ip xxx.bin*** grn 123	optional
Download the configuration file	xmodem: load configuration xmodem	optional
	tftp: load configuration tftp inet[6] ***server-ip xxx.txt***	required
	ftp: load configuration ftp inet[6] ***server-ip xxx.txt*** grn 123	optional
Download the ssh	tftp:	required

key	load keyfile private tftp[6] ***server-ip xxx.txt*** load keyfile public tftp[6] ***server-ip xxx.txt***	
	ftp: load keyfile private ftp[6] ***server-ip xxx.txt*** grn 123 load keyfile public ftp[6] ***server-ip xxx.txt*** grn 123	optional
Use the backup host program at boot time	startup secondary application	optional
Use the host program at boot time	no startup secondary application	optional

Note:

Run the main host application by default.

35.1.2 Configuration Example for File Download

1.Network requirements

The DUT connects to the file server to ensure proper communication；

2.Configuration steps

Check that the DUT and the file server are communicating properly；

Switch#ping 192.168.1.99

PING 192.168.1.99: with 32 bytes of data:

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

----192.168.1.99 PING Statistics----

4 packets transmitted, 4 packets received, 0% packet loss

round-trip (ms) min/avg/max = 0/0/0

Control-C

Upgrade the host file

Switch#load application tftp 192.168.1.99 host.arj

Downloading application via TFTP...

Download application via TFTP successfully.

EPON(onu-0/1/1)#onu-bandwidth unknown-ucast downstream 300000

Upgrade the bootrom file

Switch#load whole-bootrom tftp 192.168.1.99 bootrom_rom.bin

Reboot and use the downloaded host and bootrom files

Switch#reboot

Download the configuration file
Switch#load configuration tftp 192.168.1.99 test.txt
Startup config will be updated, are you sure(y/n)? [n]y
Downloading config file via TFTP...
Download config file via TFTP successfully.
Use the downloaded configuration file
Switch#copy startup-config running-config
Running config will be updated, are you sure(y/n)? [n]y
Start to load startup-config, please wait for a while ...
Load successfully

35.2 Overview for File Upload

File uploading refers to uploading files in DUT flash to external file servers, such as host files, configuration files, ssh key files, and log files in the upgrade file for analysis, backup, or migration to other compatible devices.

It is recommended that the uploaded file name is the same suffix as the file download: Host file name suffix is .arj; bootrom file name suffix is .bin; configuration file name suffix is .txt; ssh key file name suffix is .txt.

Support upload tools including tftp, ftp.

35.2.1 Configure File Upload

Configure File Upload

Operation	Command	Remarks
Enter privilege mode	-	-
Upload the host file	tftp: upload application tftp inet[6] ***server-ip xxx.arj***	required
	ftp: upload application ftp inet[6] ***server-ip xxx.arj*** grn 123	optional
Upload the log file	tftp: upload logging tftp inet[6] ***server-ip xxx.arj***	required
	ftp: upload logging ftp inet[6] ***server-ip xxx.arj*** grn 123	optional
Save the current configuration to flash	copy running-config startup-config	required
Upload the configuration file	tftp: upload configuration tftp inet[6] ***server-ip xxx.arj***	required
	ftp: upload configuration ftp inet[6] ***server-ip xxx.arj*** grn 123	optional
Automatically upload the configuration	tftp: upload automatically configuration tftp inet[6] ***server-ip xxx.txt*** per hours ***hours*** minutes ***minutes***	required
	ftp: upload automatically configuration ftp inet[6] ***server-ip xxx.txt*** grn 123	optional

file	per hours ***hours*** minutes ***minutes***	
Upload the ssh key	tftp: upload keyfile private tftp[6] ***server-ip xxx.txt*** upload keyfile public tftp[6] ***server-ip xxx.txt***	required
	ftp: upload keyfile private ftp[6] ***server-ip xxx.txt*** grn 123 upload keyfile public ftp[6] ***server-ip xxx.txt*** grn 123	optional

35.2.2 Configuration Example for File Upload

1.Network requirements

The DUT connects to the file server to ensure proper communication.

2.Configuration steps

Check that the DUT and the file server are communicating properly

Switch#ping 192.168.1.99

PING 192.168.1.99: with 32 bytes of data:

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

reply from 192.168.1.99: bytes=32 time<10ms TTL=64

----192.168.1.99 PING Statistics----

4 packets transmitted, 4 packets received, 0% packet loss

round-trip (ms) min/avg/max = 0/0/0

Control-C

Upload the host file

Switch#upload application tftp 192.168.1.99 host.arj

Uploading APP file via TFTP...

Upload APP file via TFTP successfully.

Save the current configuration to flash

Switch#copy running-config startup-config

Startup config in flash will be updated, are you sure(y/n)? [n]y

Building, please wait...

Update startup config successfully.

Upload a configuration file to an external server

Switch#upload configuration tftp 192.168.1.99 text.txt

Uploading config file via TFTP...

Upload config file via TFTP successfully.

Upload the files in the current flash to the file server

Switch#upload logging tftp 192.168.1.99 logg.txt
Uploading syslog file via TFTP...
Upload syslog file via TFTP successfully.

36.Decompilation Configuration

36.1 Overview for Decompilation Configuration

Device configuration can be divided into two sources: the first is called the default configuration, that does not require user configuration. After the DUT is powered on for the first time, or after the startup configuration is cleared, the existing configurations, such as the admin user, ensure that the DUT satisfies the simple usage environment. The second configuration is to increase or modify the configuration, such as creating vlan 2, modifying pvid = 2.

Device configuration can be divided into three types by saving: The first one is called temporary cache configuration or the current running configuration, such as creating vlan 2. This configuration does not exist after the DUT restarts. The second configuration is called the startup configuration, which can be loaded (either automatically or manually) after the DUT is restarted. The first configuration can be saved to the startup configuration with the command. The third configuration is saved in the flash. In the configuration, a small number of particularly important configuration will be saved directly to the flash: such as stacking configuration, user name configuration; stacking configuration will not enter the decompilation, that is, "show running" will not show, it can only be displayed by the show command in the module. User name configuration will enter the decompilation, that is, "show running" will show, it can also be displayed by the show command in the module. The configuration in Flash is permanent and does not need to be saved with commands. If you want to delete the flash configuration, you can only delete it through the corresponding no command in the module.

36.2 Basic Commands for Decompilation

Configure Decompilation

Operation	Command	Remarks
View the decompilation of the current configuration	show running-config [***module*** | interface ethernet ***port-num***]	required
View the startup configuration	show startup-config [***module***]	required
Save the current configuration to the startup configuration	copy running-config startup-config	required
Load the startup configuration at reboot	During the restart process,the default is to load the configuration automatically after 6s.Press "enter" according to the prompt message to load immediately	required
Do not load the startup configuration at reboot	During the restart process, Press "ctrl + c" according to the prompt message	optional
Load the boot configuration	copy startup-config running-config	required

at the command line		
Clear the startup configuration	clear startup-config	required

36.3 Configure the Switchover of File Execution Mode

You can change the execution mode of the configuration file through the command line interface. The system-saved configuration file can be executed in both interruptible and non-interruptible modes. When an error is encountered while executing the configuration file, execution in the interruptible mode stops immediately and echoes the error. In non-interruptible mode, execution is not stopped, the error is echoed, and the configuration file continues execution. The default is non-interruptible mode.

Configure the Switchover of File Execution Mode

Operation	Command	Remarks
Set the execution mode to interruptible	buildrun mode stop	Optional. Execute in privileged mode
Set the execution mode to non-interruptible	buildrun mode continue	Optional. Execute in privileged mode

36.3.1 Configuration Example for Decompilation

1. Configuration Example
View the decompilation of the current configuration
Switch#show running-config
!LanSwitch BuildRun
enable
configure terminal
![DEVICE]
interface ethernet 0/1
exit
interface ethernet 0/2
exit
interface ethernet 0/3
exit
interface ethernet 0/4
exit
interface ethernet 0/5
exit
interface ethernet 0/6

exit
interface ethernet 0/7
exit
interface ethernet 0/8
exit
……

Save the current configuration to the startup configuration:
Switch#copy running-config startup-config
Startup config in flash will be updated, are you sure(y/n)? [n]y
Building, please wait...
Update startup config successfully.

Use the startup configuration
Switch#copy startup-config running-config
Running config will be updated, are you sure(y/n)? [n]y
Start to load startup-config, please wait for a while ...
Load successfully.

37.Utilization Alarm

37.1 Overview for Utilization Alarm

The device utilization alarm function is used to monitor the bandwidth of the device, CPU resource consumption, and generate alarm notification in the event of congestion, so that the administrator can keep abreast of the network and equipment running.

The port utilization alarm function can set two trigger alarm thresholds. The detailed description is as follows:

- exceed：When the port bandwidth utilization equals or exceeds the "exceed" value, a congestion alarm is triggered.
- normal： When the port bandwidth utilization falls below the "normal" value, the recovered alarm is triggered.

CPU utilization alarm function can also set two trigger alarm thresholds, described in detail as follows:

- busy： When the CPU utilization equals or exceeds the "busy" value, an alarm is triggered, indicating that the CPU is busy.
- unbusy： When the CPU utilization is equal to or lower than the "unbusy" value, an alarm is triggered, indicating that the CPU is idle.

Note:

The alarm information is output to Syslog by default. If you want to output to the terminal, you need to enable the command.

37.2 Utilization Alarm Configuration

37.2.1 Configure the Port Utilization Alarm

Configure the Port Utilization Alarm

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enable/disable alarms on all ports	[no] alarm all-packets	required

Enter the port configuration mode	interface ethernet ***port-num***	required
Enable/disable port alarm	[no] alarm all-packets	required
Configure threshold information	alarm all-packets threshold exceed ***value*** normal ***value***	optional
View the alarm information	show alarm all-packets interface [ethernet ***port-num***]	optional

37.2.2 Configure the CPU Utilization Alarm

Configure the CPU Utilization Alarm

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enable/disable CPU alarm	[no] alarm cpu	required
Configure threshold information	alarm cpu threshold {[busy ***value***] | [unbusy ***value***]}	optional
View the alarm information	show alarm cpu	optional

37.2.3 Configuration Example for Utilization Alarm

1.Network requirements

When the utilization of the port 0/2 is up to 50M, alarm information is sent.

When the CPU utilization reaches 90%, alarm information is sent.

2.Configuration steps

Configure the port alarm function

Switch(config)#alarm all-packets

Switch(config)#interface ethernet 0/0/2

Switch(config-if-ethernet-0/0/2)#alarm all-packets

Switch(config-if-ethernet-0/0/2)#alarm all-packets threshold exceed 50 normal 40

Configure the cpu alarm

Switch(config)#alarm cpu

Switch(config)#alarm cpu threshold busy 90 unbusy 85

3.Result validation

Enable the serial output log

Switch(config)# logging monitor 0

When the CPU reaches the configured threshold, the following alarm information is printed
02:44:02: Switch: %OAM-5-CPU_BUSY: cpu is busy.
Switch(config)#show alarm cpu
CPU status alarm : enable
CPU busy threshold(%) : 90
CPU unbusy threshold(%) : 85
CPU status : busy

When the cpu utilization returned to normal, print as follows:
02:47:05: Switch: %OAM-5-CPU_UNBUSY: cpu is not busy.
Switch(config)#show alarm cpu
CPU status alarm : enable
CPU busy threshold(%) : 90
CPU unbusy threshold(%) : 85
CPU status : unbusy

38.Mail Alarm

38.1 Overview for Mail Alarm Function

The DUT sends the system log to the specified mailbox by mail.

38.2 Configure the Alarm

Configure Mail Alarm

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enable/disable mail alarm	[no] mailalarm	required
Configure the smtp server	[no] mailalarm server ***ip-address***	required
Configure smtp identity	mailalarm smtp authentication username ***name*** passwd ***password***	required
Configure the mail sender	[no] mailalarm sender ***sender@co.com***	required
Configure the mail recipient	[no] mailalarm receiver ***receiver@abc.com***	required
Configure the mail cc	[no] mailalarm ccaddr ***ccaddr @abc.com***	optional
Configure the alarm log level	[no] mailalarm logging level ***value***	optional
View the configuration	show mailalarm	optional

Note:

1.Mailalarm is disabled by default。

2 The default log level uses 0, and when configured to n, logs of level 0-n are sent to the email specified by mailalarm.

38.3 Configuration Example for Mail Alarm

1. Network description

Let the DUT send 0-3 level logs to receive@126.com.The sender is test/test and mailbox is test@126.com, mail server is 183.222.100.112.

2.Configuration steps

Configuration

Switch(config)#mailalarm

Switch(config)#mailalarm server 183.222.100.112

Switch(config)#mailalarm smtp authentication username test passwd test

Switch(config)#mailalarm sender test@126.com

Switch(config)#mailalarm receiver receive@126.com

Switch(config)#mailalarm ccaddr cc@126.com

Switch(config)#mailalarm logging level 3

View the information

Switch(config)#show mailalarm

mailalarm state : on

smtp authentication : on

smtp server address : 11.1.1.1

mailalarm logging level : 3

sender e-mail address : test@126.com

receiver e-mail address : receive@126.com

ccaddr : cc@126.com

39.System Log

39.1 System Log Overview

Syslog is the system information center, to complete the unified processing and output of information.

The other modules in the system will send the output information to Syslog. Syslog determines the output format of the information according to the configuration of the user and outputs the information to the specified display device according to the information switching and filtering rules of each output direction configured by the user.

With the Syslog information producer, which is each module of output information, you do not need to export information to the console, Telnet terminal, or log host (Syslog server). You only need to output the information to Syslog. By configuring the appropriate filtering rules, information consumers, which are console, Telnet terminal, history buffer, log host and SNMP agent, can choose whatever they want to receive the information they need and discard unwanted information.

39.2 System Log Configuration

39.2.1 Enable/disable Syslog

Enable/disable Syslog

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Enable/disable log function	[no] logging	optional
View the configuration information	Show logging	optional

Note:

The logging function is enabled by default and stored in the buffer.

39.2.2 Configure the Log Serial Number

Configure the Log Serial Number

Operation	Command	Remarks
Enable/disable the log serial number	[no] logging sequence-numbers	required

Note:

The logging function is enabled by default.

39.2.3 Configure the Timestamp

Configure Syslog Timestamp

Operation	Command	Remarks
Configure the timestamp type	logging timestamps { notime | uptime | datetime }	optional
Recover the default configuration	no logging timestamps	optional

Note:

There is no separate timestamp switch. There are three timestamp types: notime: do not show the time; uptime: show the boot time; Datetime: show the date and time. The default is uptime.

39.2.4 Output to the Terminal

In privileged mode, configure the log to output to the terminal switch. In global mode, you can configure information display and filtering rules. By default, the device logs are not output to the terminal, but output to the buffer. There is a slight difference between the command for the serial terminal and the Telnet or ssh terminal.

Output to the Terminal

Operation	Command	Remarks
User configuration mode	end	required
Enable/disable	[no] terminal monitor	optional

output to the terminal		
Global configuration mode	configure terminal	required
Enable/disable log display	[no] logging monitor { all | ***monitor-num*** }	optional
Configure the filtering rules	logging monitor { all | ***monitor-num*** } { ***level-value*** | none | level-list { ***start-leve***l to ***end-level*** | ***value*** } } [module ***module-name***]	optional
View the filtering rules	show logging filter monitor ***monitor-num***	optional
Delete the filtering rule	no logging monitor { all | ***monitor-num*** } filter	optional

Note:

1. Log output to the terminal: In the serial console, the default configuration is terminal monitor; in other terminal console, the default is no terminal monitor;
2. Log information display: In the non-console terminal configuration, only affect this landing of the current terminal, the other terminals, the next landing of the current terminal is invalid;
3.monitor-num is 0 for the console, and 1 to 5 for Telnet and ssh terminals.
4. Output log default rule: all modules, log level 0-5,7. Deleting the filtering rule restores the default rule.

39.2.5 Output to Buffer

Output Syslog to Buffer

Operation	Command	Remarks
Enter global configuration mode	configure terminal	optional
Enable/disable output to buffer	[no] logging buffered	optional
Configure the filtering rules	logging buffer { ***level-value*** | none | level-list { ***start-leve***l to ***end-level*** | ***value*** } } [module ***module-name***]	optional
View the filtering rules	show logging filter buffered	optional
Delete the filtering rule	no logging buffered filter	optional
View the log information in the buffer	Show logging buffered [***level-value*** | level-list { ***start-leve***l to ***end-level*** | ***value*** }] [module ***module-name***]	optional

Note:

The default is log output to buffer. The default rule is to output all modules and logs at the level 0-6. Deleting the filtering rule restores the default rule.

39.2.6 Output to Flash

In global mode, you can configure Syslog to save to Flash, which is not saved in flash memory by default.

Output Syslog to Flash

Operation	Command	Remarks
Enter global configuration mode	configure terminal	optional
Enable/disable output to flash	[no] logging flash	required
Configure the filtering rules	logging flash { ***level-value*** | none | level-list { ***start-leve***l to ***end-level*** | ***value*** } } [module ***module-name***]	optional
View the filtering rules	show logging filter flash	optional
Delete the filtering rule	no logging flash filter	optional
Configure the save period	[no] logging flash interval ***value***	optional
Configure the log size to be saved each time	[no] logg flash msg-number ***value***	optional
Check the log information in the flash	Show logging flash [***level-value*** | level-list { ***start-leve***l to ***end-level*** | ***value*** }] [module ***module-name***]	optional

Note:

1. When the log is output to flash, the default rule is to output all modules and the log level is 0-5. Deleting the filtering rule restores the default rules.

2. When the log output to flash, the default cycle is 30M. By default, 100 logs are saved at one time.

39.2.7 Output to External Server

Configure the specified server address for log output, information output switch, filtering rule, and logging tool and source address in global mode.

Output Syslog to External Server

Operation	Command	Remarks
Enter global	configure terminal	必先

configuration mode		
Configure the log server	[no] logging ***ip-address***	required
Enable/disable the log server	[no] logging host { all | ***ip-address*** }	required
Configure the filtering rules	logging host { all | ***ip-address*** } { ***level-value*** | none | level-list { ***start-leve***l to ***end-level*** | ***value*** } } [module ***module-name***]	optional
Recover the default filtering rules	no logging host { all | ip-address } filter	optional
Configure the logging tool name	[no] logging facility { clock1 | clock2 | ftp | kernel | lineprinter | localuse0 | localuse1 | localuse2 | localuse3 | localuse4 | localuse5 | localuse6 | localuse6 | localuse7 | logalert | logaudit | mail | networkknews | ntp | security1 | security2 | syslogd | system | userlevel | uucp }	optional
Configure the sip for log packet	[no] logging source { ***ip-address*** | loopback-interface ***if-id*** }	optional

Note:

1. The sip of log messages must be the interface of the device. The Layer 3 device uses the IP address of the corresponding interface of the log server by default. The Layer 2 device automatically uses the system IP and does not need to be configured.

2. The default logging tool name uses localuse7.

39.2.8 Output to the SNMP Agent

Configure Syslog output to the SNMP agent in global mode. To send Syslog messages to SNMP Workstation in trap messages, you must also configure the Trap host address. Refer to the SNMP configuration instructions.

By default, logs are not output to the SNMP agent.

Output Syslog to SNMP Agent

Operation	Command	Remarks
Enable/disable log output to the snmp agent	[no] logging snmp-agent	required
Configure the filtering rules	logging snmp-agent { ***level-value*** | none | level-list { ***start-leve***l to ***end-level*** | ***value*** } } [module ***module-name***]	optional
View the filtering rules	show logging filter snmp-agent	optional
Recover the default	no logging snmp-agent filter	optional

filtering rules		

Note:

1. Log output to the snmp agent, the default rule: output all modules, the log level is 0-5.

39.2.9 Debugging Function

In the global mode, you can configure the debugging function to print the debugging information of the corresponding module. By default, debugging information of all modules is disabled.

Configure the Debugging Function

Operation	Command	Remarks
Enable/disable debugging function	[no] debug { all | ***module-name*** }	required
View the configuration information	show debug	optional

39.2.10 Syslog Configuration Example

1.Network requirements

Output the logs of the stp module and the device module at levels 0-4 to the console terminal; turns on the serial number display; timestamp use datetime; the log is output to the flash memory; the log information of level 3 and 4 is output to buffer; output logs to external server: 192.168.1.3; open arp debugging information.

2.Configuration steps

Enable the terminal output function
Switch#terminal monitor
Switch#configure terminal

Enable the logging function
Switch(config)#logging

Enable the terminal display function
Switch(config)#logging monitor all

Output the logs of the stp module and the device module at levels 0-4 to the console terminal
Switch(config)#logging monitor all level-list 0 to 4 module stp device

Enable the serial number display
Switch(config)#logging sequence-numbers

Configure the timestamp to use datetime
Switch(config)#logging timestamps datetime

Enable log output to flash
Switch(config)#logging flash

Configure the buffer log filtering rule
Switch(config)#logging buffered level-list 3 4

Configure the log server
Switch(config)#logging 192.168.1.3

Enable the log server
Switch(config)#logging host 192.168.1.3

Configure the facility
Switch(config)#logging facility ftp

Enable the debugging function of the arp module
Switch(config)#debug ARP

View the configuration information
Switch(config)#show logging
state: on;
logging sequence-numbers: on;
logging timestamps: datetime;
logging language: english
logging monitor:
Console: state: on; display: off; 96 logged; 0 lost; 0 overflow.
logging buffered: state: on; 249 logged; 0 lost; 0 overflow.
logging flash: state: on; 37 logged; 0 lost; 0 overflow.
logging loghost:
logging facility: ftp;logging source: off
192.168.1.3: state: on; 23 logged; 0 lost; 0 overflow.
logging SNMP Agent: state: off; 0 logged; 0 lost; 0 overflow.

40.System Maintenance

40.1 View the System Status

This section describes some of the show commands for the system.

View the System Status

Operation	Command	Remarks
Display version information	show version	-optional
Display system information	**show system**	optional
Display memory information	**show memory**	optional
Display CPU utilization	**show cpu-utilization**	optional
Display the statistics of CPU packets	**show cpu-statistics** *ethernet port-num*	Optional. According to port statistics
Display the statistics of CPU packet types.	**show cpu-classification** *interface ethernet port-num*	optional
Display the information of the administrator who can log in to the system	**show username**	optional
Display the information of the administrator who has logged in to the system	**show users**	optional
Display the system clock	**show clock**	optional
Display all fdb tables	**show ip fdb**	optional
Display the fdb table of the specified ip	**show ip fdb *ip***	optional
Displays the fdb table of the specified address segment	**show ip fdb *ip mask***	optional

Note：Fdb table is arp table which is issued to the three-tier switching chip, that is hardware arp table. The Layer 2 Switch does not have this entry.

40.2 Set the Switch Host Name

Run the hostname command and set the system command line interface prompt in the global configuration mode.

Set the Switch Host Name

Operation	Command	Remarks
Enter privilege configuration mode	**enable**	

Enter global configuration mode	**configure terminal**	-
Set the system command line interface prompt	**hostname** *host-name*	optional
Cancel the system command line interface prompt	no **hostname**	optional

40.3 Set the System Clock

The device has a clock that can be calibrated by command. Set the system clock by operating the clock set command in privileged user mode.

Set the System Clock

Operation	Command	Remarks
Enter privilege configuration mode	**enable**	
Set the clock	**clock set** *HH:MM:SS YYYY/MM/DD*	**required**
Enter the global mode	**configure terminal**	**-**
Set the time zone	**clock timezone** *name hour minute*	**optional**
View the clock of the device	**show clock**	**optional**

Configuration example:
// Set the clock for the device
switch#clock set 13:12:33 2014/08/10
Set clock successfully.

// Set the time zone for the device to the Beijing time zone
switch(config)#clock timezone beijing 8 0
Set timezone successfully.

// View the clock of the device
switch(config)#show clock

Sun 2014/08/10 13:19:15 beijing 08:00

40.4 Network Connection Test Command

Ping is used to check whether the network connection and the host is reachable. Operate the following configurations in privileged user mode or normal user mode.

Ping Test Command

Operation	Command	Remarks

Run the ping command	**ping** {- **i** *ttl*|-l *packet length*|**-n** *count* |**-s** *sourceip* | **-t** *timeout*} host	optional
Run the ping6 command	ping6 {**-a** *ipv6 source address* | **-c** *count* |**-h** *hop_limit*| **-s** *packet length* | **-t** |**-w** *time_out}*	optional

IpV4 ping Command Parameter Description:

- i ttl: TTL value to be sent

-l packetlength：Length of the sent packet, in bytes

-n count：Number of sent packets

-s sourceip: Source IP address of the sent packet

-t timeout：The timeout to wait for the response after sending the packet，in seconds.

IPV6 ping Parameter Description:

-a source address：The source IPV6 address to be sent;

-c count：Number of sent packets

-h hop limit：Hop limit

-s packet length：Length of the sent packet, in bytes

-w time out：the timeout to wait for the response after sending the packet.

40.5 Route Tracking Command

Tracert is mainly used for route tracking and checking network connections. Operate the following configurations in privileged user mode or normal user mode.

Route Tracking Command

Operation	Command	Remarks
Perform IPV4 route tracking	**tracert** { **-u** | **-c** } { **-p** *udpport* | **-f** *first_ttl* | **-h** *maximum_hops* | **-w** *time_out* } target_name	optional
Perform ipv6 route tracking	**tracert6** { **-c** |**-h** *maximum_hops* | **-w** *time_out* } **ipv6_host_address**	optional

Parameter Description:

-u means to send udp packets, -c means to send icmp echo packets, the default is -c mode;

udpport：Destination port address for sending udp packets. It ranges from 1 to 65535, the default port is 62929;

first_ttl：Initial ttl value of the sent packets, in the range of 1 to 255. The default is 1.

maximum_hops：Maximum ttl value of the sent packets, in the range of 1 to 255. The default is 30.

time_out：The timeout to wait for the response after sending the packe, in the range of 10 to 60, in seconds. The default value is 10 seconds.

target_name：Destination host or router address

40.6 Banner

After setting the banner, the manufacturer information will appear when the device is logged in

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	-
Enable the banner function	**banner**	-
Disable the banner function	**no banner**	
Customize the first line of the banner	**banner line1** *string*	
Customize the second line of the banner	**banner line2** *string*	
Customize the third line of the banner	**banner line3** *string*	
Customize the fourth line of the banner	**banner line4** *string*	

The device has a default banner, but the banner function is turned off by default

Example:

Enable the banner function, log in again after exiting, it will display the banner.

Switch(config)#banner

witch(config)#exit

Switch#quit

Username:admin

Password:*****

****

* It's XX owner of this equipment, please remember for legal

* liability.

****

40.7 The Number of Lines Displayed When Viewing Information

You can display up to 25 lines of information at one time by default when viewing device information.

Operation	Command	Remarks

Enter global configuration mode	**configure terminal**	-
Enable the banner function	**screen-rows per-page** *inter*	-

The range of per-page is 0 to 256, and 0 means that all information is displayed. There is no limit; setting per-page to 0 can collect information efficiently.

40.8 Restart Switch

40.8.1 Command to Restart Switch Immediately

You can use the following commands to restart the Switch immediately.

Restart the Switch Immediately

Operation	Command	Remarks
Restart the Switch immediately	reboot	Optional. Run the command in privileged mode

40.8.2 Restart the Switch Periodically

The device allows the customer to set the restart time.

Restart the Switch periodically

Operation	Command	Remarks
Restart the Switch periodically	**auto-reboot** { **in** { hours *hour* | minutes *min* } | **at**{ *hh:mm:ss YYYY/MM/DD* | *hh:mm:ss* daily | *hh:mm:ss weekday* weekly } }	Optional. Run the command in global mode
Cancel the timing reboot configuration	**no auto-reboot**	optional
View the timing reboot configuration	**show auto-reboot**	optional

Note：In auto-reboot at mode, configure the system clock to be used together.

Configuration Example:

// Configure the Switch to reboot after 3 minutes；

Switch(config)#auto-reboot in hours 0 minutes 3

Enable auto-reboot successfully.

// View the timing reboot configuration；
Switch(config)#show auto-reboot
Auto-reboot setting
  Type: one-off/in
  Time: 2014-01-02 00:47:35
  Auto-reboot in 0 hours, 2 minutes and 24 seconds. // The configuration starts with a countdown；

// Restart the switch
Switch(config)#
It's time to reload, ready to reboot.....
0
    Count down to auto-boot...
     0
    boot default application from flash.....

Loading image...OK
Unarj image...OK

41.sFlow Configuration

41.1 sFlow Overview

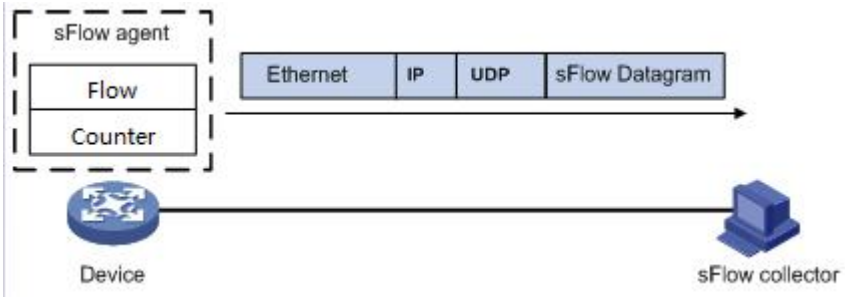
sFlow is a network traffic monitoring technology based on packet sampling, which is mainly used for statistical analysis of network traffic.

As shown in the figure, the sFlow system consists of the sFlow Agent embedded in the device and the remote sFlow Collector. The sFlow agent obtains the statistics and packet information of the interface through the sampling mechanism, encapsulates the information into sFlow packets. When the sFlow packet buffer is full or the sFlow packet sending timer (Timer interval is fixed to 1 second) expires, the sFlow packet is encapsulated in UDP packet and sent to the specified sFlow Collector. The sFlow Collector analyzes the sFlow packet and displays the analysis result. An sFlow Collector can monitor multiple sFlow Agents.

SFlow uses the following two sampling mechanisms:

· Flow Sampling: Packet-based stream sampling is used to obtain information about the contents of the packet.

· Counter Sampling：Time-based interface statistics sampling is used to obtain the interface statistics.



41.2 sFlow Configuration

41.2.1 Configure sflow agent IP

The IP address of the sFlow agent is the source IP address that the Switch communicates with the remote sFlow Collector. The IP address must be the IP address of the Switch itself. You can configure only one IP address for the sFlow agent on the device. The newly configured IP address overwrites the existing configuration.

Operation	Command	Remarks
Enter the global mode	**configure terminal**	
Configure the sflow agent IP	**sflow agent ip** *A.B.C.D*	
Delete the sflow agent IP	**no sflow agent ip**	

【For example】

For example:

！Configure the IP address of the sflow agent to 1.1.1.1

switch(config)#sflow agent ip 1.1.1.1

41.2.2 Configure sFlow Collector

sflow collector is used to monitor the traffic of the Switch device. The Switch must be configured with sflow collector ip and port number.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	

Configure the sflow collector ip and port number	**sflow collector** *id* **ip** *ip-address* **[port** *port-number* **]**	
Delete the sflow collector	**no sflow agent ip**	

【Parameter Description】

id: collector number (in the range of 1-10)

ip-address: the IP address of the collector

port-number: The port where the collector listens for sflow packets (6343 by default)

【For example】

For example:

！Configure the collector with the number 2, IP address 1.1.1.2 and port number 6345

switch (config)#sflow collector 2 ip 1.1.1.2 port 6345

41.2.3 Configure sflow sampling-rate

This command is used to configure the packet sampling rate for Flow sampling, that is, a packet is sampled in rate packets and the Flow sampling function is enabled. Flow Sampling uses the random sampling mode.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enter port configuration mode	**interface ethernet** device/slot/port	
Configure the sampling rate of the Flow packet	**sflow sampling-rate** *rate*	
Delete the sampling of the Flow packet	**no sflow sampling-rate**	

【For example】

For example:

！Configure the Flow sampling rate of port 2 to be one per 3000 packets

switch (config-if-ethernet-0/0/2)#sflow sampling-rate 3000

41.2.4 Configure sflow flow max-header

This command is used to configure the maximum number of bytes that can be copied from the head of the original packet when Flow sampling performs packet content copying. By default, the maximum number of bytes that can be copied is 128 bytes.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enter port configuration mode	**interface ethernet** device/slot/port	
Configure the content copy length of the Flow packets	**sflow flow max-header** *length*	
Restore the default	**no sflow flow max-header**	

【Parameter Description】

length: Maximum number of bytes allowed to be copied (in the range 18-512)

【For example】

For example:

！Configure the maximum number of bytes that can be copied to 200 for Flow sampling on

port 2

switch (config-if-ethernet-0/0/2)#sflow flow max-header 200

41.2.5 Configure sflow flow collector

Flow sampling and sFlow Collector are bound by the collector number

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enter port configuration mode	**interface ethernet** device/slot/port	
Configuration	**sflow flow collector** *id*	
Restore the default	**no sflow flow collector**	

41.2.6 Configure sflow counter interval

The Switch can also be sampled at regular intervals.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enter port configuration mode	**interface ethernet** device/slot/port	
Configure the counter sampling interval	sflow counter interval time	
Restore the default	**no sflow counter interval**	

41.2.7 Configure sflow counter collector

This command is used to configure the collector number of the counter sampling. The no form of this command cancels this setting.

Operation	Command	Remarks
Enter global configuration mode	**configure terminal**	
Enter port configuration mode	**interface ethernet** device/slot/port	
Configure the Counter sampling number	**sflow counter collector** *id*	
Delete the Counter sample number	**no sflow counter collector**	

【For example】

For example:

！Set the collector number of the port counter sampling to 1

switch (config-if-ethernet-0/0/2)#sflow counter collector 1

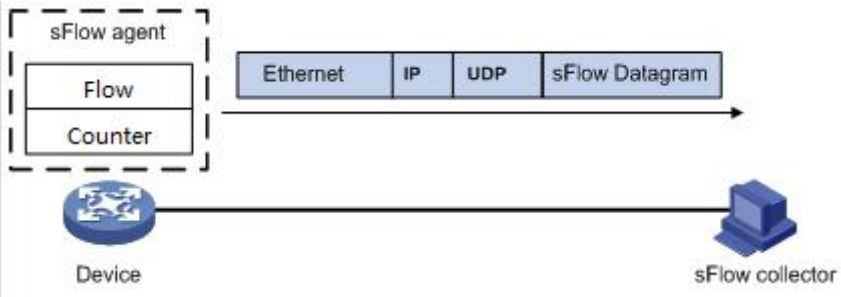
41.2.8 The Command of show sflow

This command is used to display the sflow configuration

Operation	Command	Remarks
Any mode	show sflow	

41.3 Example

Device ip = 192.168.2.1; PC as sflow collector, IP = 192.168.2.100



Configure the device ip
switch(config)#interface vlan-interface 2
Create vlan-interface successfully!

switch(config-if-vlanInterface-2)#ip address 192.168.2.1 255.255.255.0
This ipaddress will be the primary ipaddress of this interface.
Config ipaddress successfully

Set the sflow agent ip
switch(config)#sflow agent ip 192.168.2.1

Configure sflow collector ip
switch(config)#sflow collector 1 ip 192.168.2.100 port 6000
Port configuration
switch(config-if-ethernet-0/0/3)#sflow counter collector 1

42.CFM Configuration

42.1 CFM Overview

CFM (Connectivity Fault Management Protocol), defined by the IEEE 802.1ag standard, is a VLAN-based side-to-side OAM mechanism on Layer 2 link for fault management of carrier Ethernet.

42.1.1 CFM Concept

CFM Concept

Concept	Description
Maintenance Domain	The maintenance domain indicates the network covered by connectivity fault detection, whose boundaries are defined by a series of maintenance end points configured on the port. The maintenance domain is identified by the name of the maintenance domain. According to the network planning, the maintenance domain can be classified into eight levels. Different maintenance domains can be adjacent to each other or nested, but can not be crossed. Nesting can only be performed from high-level maintenance domain to low-level maintenance domain. That is, a low-level maintenance domain must be included in a high-level maintenance domain.
Maintenance association	You can configure multiple maintenance associations as required in the maintenance domain. Each maintenance association is a collection of maintenance points in the maintenance domain. The maintenance association is identified by the "Maintenance Domain Name + Maintenance Association Name". The maintenance association serves a VLAN. The packets sent by the maintenance point in the maintenance association are tagged with the VLAN. The maintenance point in the maintenance association can receive the packets sent from other maintenance points in the maintenance association.
Maintenance point	The maintenance point is configured on a port and belongs to a maintenance association. It can be classified into two types: maintenance end points and maintenance intermediate points. A maintenance end point is identified by a MEP ID, which determines the scope and boundaries of the maintenance domain. The maintenance end points are directional, and are classified into UP MEP and DOWN MEP. The direction of the maintenance end point indicates the location of the maintenance domain relative to the port. The DOWN MEP sends a packet to the port on which it resides. Instead of sending a packet to its port, the UP MEP sends a packet to the other port of the device. The maintenance intermediate point is located inside the maintenance domain, It can not actively send CFM protocol packets but can process and respond to CFM protocol

	packets.

42.1.2 CFM Main Functions

The effective application of connectivity fault detection is based on reasonable network deployment and configuration. Its function is implemented between the configured maintenance points. The main functions are as follows:

CFM Main Functions

Function	Description
Continuity detection	It is an active OAM function used to detect the connectivity between maintenance end points. Failure to connect may be caused by a device failure or a configuration error.
Loopback function	It is an on-demand OAM function used to verify the connection status between the local device and the remote device.
Link tracing function	It is an on-demand OAM function that determines the path between the local device and the remote device to locate the link fault.

42.1.3 Configure CFM

Before configuring the CFM function, make the following planning for the network:

a. The maintenance domain of the entire network is classified to determine the boundaries of the maintenance domains at each level.

b. Identify the names of the various maintenance domains. The names of the same maintenance domain are the same on different devices.

c. Determine the maintenance association in each maintenance domain according to the VLAN to be monitored.

d. Determine the name of each maintenance association. The same maintenance association in the same maintenance domain has the same name on different devices.

e. Determine the list of maintenance end points for the same maintenance association in the same maintenance domain, which should be the same on different devices.

f. The maintenance end points can be planned on the boundary ports of the maintenance domain and the maintenance association. The maintenance intermediate points can be planned on the non-border devices or ports.

After you have completed the network planning, perform the following configuration.

42.2 CFM Configuration

42.2.1 Configure the MD(Maintenance Domain)

Configure the MD

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	
Create a maintenance domain and enter maintenance domain configuration mode	**cfm md** *md-index*	required

42.2.2 Configure the Maintenance Domain Name and Level

In order to distinguish the various maintenance domains, you can specify different domain names for each maintenance domain. The domain name consists of two parts: name format and name content. The domain name is preferred to be unique throughout the network. To indicate nested relationships among maintenance domains, you must also specify maintenance domain level. Only a maintenance domain with a high level can nest a maintenance domain with a small level.

Configure the Maintenance Domain Name and Level

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	
Enter the maintenance domain configuration mode	**cfm md** *md-index*	When a domain does not exist, the domain is created;
Delete the MD	**no cfm md** *md-index*	
Configure a maintenance domain with no name and specify only the level of the maintenance domain	**cfm md format** none **level** *md-level*	The two must be one
Configure the name of the maintenance domain, and specify the name and level of the maintenance domain	**cfm md format** { dns-name | mac-uint | string } **name** *md-name* **level** *md-level*	

42.2.3 Configure the Maintenance Association

Configure the Maintenance Association

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Create a maintenance association and enter the	**cfm ma** *ma-index*	required

maintenance association configuration mode		
Delete the maintenance association configuration	**no cfm ma** *ma-index*	-

42.2.4 Configure the Maintenance Association Name and Associated VLAN

In order to distinguish the maintenance association in each maintenance domain, you can specify different instance names for each maintenance association. The instance name consists of two parts: name format and name content. The domain name and the instance name of the maintenance domain where the maintenance association is located must be unique to the entire network.

Configure the Maintenance Association Name and Associated VLAN

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Configure the name of the MA and the primary VLAN	**cfm ma format** { primary-vid | string | uint16 | vpn-id } **name** *ma-name* **primary-vlan** *vlan-id*	required

42.2.5 Configure the MEP (Maintenance End Points)

The CFM function is mainly used for various operations on the maintenance end points. You can configure the maintenance end points on the network edge ports according to the network planning.

Configure the MEP

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Create a MEP and specify its associated port	**cfm mep** *mep-id* **direction** { up | down } [**primary-vlan** *vlan-id*] **interface ethernet** *port-id*	required
Enable the MEP management state	**cfm mep** *mep-id* **state** enable	Required. By default, it is off
Close the management state of the MEP	**cfm mep** *mep-id* **state** disable	-
Configure the priority that the MEP sends to CCM and LTM	**cfm mep** *mep-id* **priority** *priority-id*	Optional By default, the priority

		is 0

42.2.6 Configure the Remote MEP

The remote MEP is relative to the local MEP. In the entire maintenance association, all MEPs other than the MEPs of the local should be configured as remote MEPs on the local.

Configure the Remote MEP

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Create the remote MEP and specify the relative local MEP	**cfm rmep** *rmep-id* **mep** *mep-id*	required

42.2.7 Configure the MIP

The MIPs are used to respond to various CFM test packets. You can configure the MIPs on non-border devices or ports based on network planning.

Configure the MIP

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Create a maintenance intermediate point and specify its associated port	**cfm mip** *mip-id* **interface ethernet** *port-id*	optional

42.2.8 Configure the Continuity Check Function

By configuring the continuity check function, you can enable the MEPs to send CCM packets between them to detect the connectivity between these MEPs, and thus to manage the link connectivity.

Configure the Continuity Check Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Configure the interval at which	**cfm cc interval** { 1 | 10 | 60 | 600 }	Optional.

CCMs are sent by the MEP		The default is 1s.
Enable the ccm sending function on the MEP	**cfm mep** *mep-id* **cc** enable	Required. By default, it is off
Cancel the ccm sending function of the MEP	**cfm mep** *mep-id* **cc** disable	Optional

Note:

The time interval for sending CCMs must be the same on the maintenance end points in the same maintenance domain and the maintenance association on different devices.

42.2.9 Configure the Loopback Function

By configuring the loopback function, you can check the link status between the source and destination MEPs or the MIPs to verify link connectivity.

Configure the Loopback Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Enable the loopback function	**cfm loopback mep** *mep-id* { **dst-mac** *mac-address* | **dst-mep** *rmep-id* } [**priority** *pri-id* | **count** *pkt-num* | **length** *data-len* | **data** *pkt-data*]	optional

42.2.10 Configure the Link Tracking Function

By configuring the link tracking function, you can locate the path between the source and destination MEPs or the MIPs, and locate the link faults.

Configure the Link Tracking Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Enable link tracking	**cfm linktrace mep** *mep-id* { **dst-mac** *mac-address* | **dst-mep** *rmep-id* } [**timeout** *pkt-time* | **ttl** *pkt-ttl* | **flag** { use-mpdb | unuse-mpdb }]	optional

42.2.11 Y.1731 Frame Loss Rate Detection Function

Y.1731 Frame Loss Rate Detection Function

Operation	Command	Remarks
Enter the global configuration	**configure terminal**	-

mode		
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Perform the frame loss rate detection function	**cfm eth-slm mep** *mep-id* { **dst-mac** *mac-address* | **dst-mep** *rmep-id* } [**timeout** *pkt-time* | **priority** *priority-identifier* | **interval** *second* **/count** *packet-num]*	optional

42.2.12 Y.1731 Frame Delay Measurement Function

Y.1731 Frame Delay Measurement Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the maintenance domain configuration mode	**cfm md** *md-index*	-
Enter the maintenance association configuration mode	**cfm ma** ma-index	-
Perform frame delay measurement	**cfm eth-2dm mep** *mep-id* { **dst-mac** *mac-address* | **dst-mep** *rmep-id* } [**timeout** *pkt-time* | **priority** *priority-identifier* | **interval** *second* **/count** *packet-num]*	optional

42.2.13 CFM Display and Maintenance

After completing the above configuration, you can use the following command to display the CFM configuration.

CFM Display and Maintenance

Operation	Command	Remarks
Clear the CCM statistics	**clear cfm cc**	Enter global configuration mode
Clear the CCM database information	**clear cfm cc database**	
Display the maintenance domain information	**show cfm md** [*md-index*]	Any mode can be viewed
Display the maintenance association information	**show cfm ma**	
Display the local maintenance points information	**show cfm mp local**	
Display the remote maintenance point information	**show cfm mp remote**	
Display CCM statistics	**show cfm cc**	
Display CCM database information	**show cfm cc database**	
Display CFM alarm information	**show cfm errors**	

42.3 Configuration Examples

Switch(config)#cfm md 1 // Create a maintenance domain

Switch (config-cfm-md-1)#cfm md format none level 1 // Enter the maintenance domain configuration mode, configure a maintenance domain with no name and the maintenance domain level is 1.

Switch (config)#cfm md 1 // Enter the maintenance domain configuration mode

Switch (config-cfm-md-1)#cfm ma 1 // Enter the maintenance association configuration mode and set the name of the maintenance association to 1

Switch (config-cfm-md-1-ma-1)#cfm ma format primary-vid name 1 primary-vlan 2

// The associated VLAN is VLAN 2

Switch (config-cfm-md-1-ma-1)#cfm mep 1 direction up primary-vlan 2 interface

ethernet 0/0/2 // Create MEP 1 and specify the associated port as VLAN 2

Switch (config-cfm-md-1-ma-1)#cfm mep 1 state enable

Switch (config-cfm-md-1-ma-1)#cfm mep 1 priority 1 // Set the priority of sending CCMs and LTMs to 1 by the MEP

43.EFM Configuration

43.1 EFM Overview

EFM (Ethernet of First Mile), known as the first mile Ethernet, is defined by the IEEE 802.3ah standard for the management and maintenance of point-to-point Ethernet links between two devices.

43.1.1 EFM Main Function

EFM can effectively improve the management and maintenance of Ethernet capacity to ensure the stable operation of the network, its main functions include:

EFM Main Function

Function	Explanation
EFM auto-discovery function	The EFM function is established on the basis of the EFM connection. The EFM connection establishment process is realized by the EFM automatic discovery function. Between the connected EFM entities, the information about the EFM configuration and the EFM capability supported by the local EFM are notified through the information OAMPDUs. After the EFM entity receives the configuration parameters of the opposite side, it determines whether to establish the EFM connection. There are two EFM modes: active mode and passive mode. An EFM connection can only be initiated by an EFM entity in active mode. An EFM entity in passive mode can only wait for a connection request from an opposite EFM entity. EFM connections can not be established between two EFM entities in passive mode.
Remote fault indication function	When the device detects an emergency link event, the faulty EFM entity reports the fault information (that is, the emergency link event type) to the remote EFM entity through the Flag field in the Information OAMPDU. In this manner, the administrator can dynamically learn the link status by observing the log information and process the corresponding errors in time. The emergency link event types include Link Fault, Dying Gasp, and Critical Event.
Link monitoring function	The link monitoring function is used to detect and discover the link layer faults in various environments. The EFM monitors the links by exchanging Event Notification OAMPDUs：When an EFM entity detects a general link event, it sends Event Notification OAMPDU for notification, administrators can monitor the status of the network dynamically by observing the log information. The general link event types include errored-symbol-period, errored-frame, errored-frame-period and errored-frame-seconds.
Remote loopback	Remote loopback means that an EFM entity in active mode sends all the packets except the OAMPDUs to the remote side. After receiving the packet, the remote device does not forward the packet according to its destination address. Instead, it sends the packet back to the Local.

	The remote loopback function controls the remote side to perform the remote loopback function or cancel the remote loopback operation through the loopback control OAMPDU. This function can be used to detect the link quality and locate the link fault.
Remote MIB variable acquisition	The EFM entity can obtain the MIB variable value of the remote entity by exchanging the Variable Request / Response OAMPDU. The MIB variables contain all the performance parameters and error statistics on the Ethernet link. It provides the local EFM entity with a common detection mechanism for the performance and error of the remote entity.

Note:

The EFM-enabled port is called an EFM entity.

43.1.2 EFM Protocol Packet

EFM works at the data link layer, and its protocol packet is called OAMPDU (OAM Protocol Data Units). EFM reports the link status by periodically exchanging OAMPDUs between devices so that the network administrator can manage the network effectively.

EFM Protocol Packet Types and Effects

Packet type	Function
Information OAMPDU	It is used to send the status information of the EFM entity (including local information, remote information, and custom information) to the remote EFM entity to keep the EFM connection.
Event Notification OAMPDU	Generally, it is used for link monitoring and alarming of faults on the link connecting the local and remote EFM entities.
Loopback Control OAMPDU	It is mainly used for remote loopback control. It is used to control the EFM loopback status of a remote device. This packet contains information about enabling or disabling the loopback function. The remote loopback function is enabled or disabled based on this information. .
Variable Request/ Response OAMPDU	It is used to obtain the MIB variable value of the remote device to monitor the remote status.

43.2 Configure EFM

43.2.1 EFM Basic Configuration

EFM works in active mode and passive mode. When EFM is enabled, the Ethernet port starts using the default working mode to establish EFM connections with its opposite ports.

EFM Basic Configuration

Operation	Command	Remarks
Enter the global	**configure terminal**	

configuration mode		
Enter the port configuration mode	Interface ethernet *interface-num*	Required
Start EFM	**efm**	Required. By default, EFM is disabled
Close EFM	**No efm**	Port mode
Configure the EFM working mode	**efm mode** { **passive** | **active** }	Optional By default, EFM works in active mode

43.2.2 Configure EFM Timer Parameter

After an EFM connection is established, EFM entities at both sides send information OAMPDUs at intervals of a certain interval to check whether the connection is normal. The interval is called handshake packet sending interval. If an EFM entity does not receive an Information OAMPDU from the remote EFM entity within the connection timeout period, the EFM connection is considered interrupted.

By adjusting the EFM handshake packet sending interval and connection timeout time, you can change the detection precision of the EFM connection. Configure the timeout time for responses of remote devices to OAMPDU request packets, if the response times out, the OAMPDU response packets received are discarded.

EFM Timer Parameter Configuration

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	interface ethernet *interface-num*	If Switch products, only ethernet type port; if PON products, there are Ethernet and pon type port
Set the interval for sending EFM handshake packets	**efm pdu-timeout** *time*	Optional The default is 1s
Set the timeout time for EFM connections	**efm link-timeout** *time*	Optional The default is 5s
Set the response timeout time	**efm remote-response-timeout** *time*	Optional The default is 2s
Recovery Response Timeout Time	**No efm remote-response-timeout**	Optional, restore the default

Note:

After the EFM connection times out, the local EFM entity ages the connection with the opposite EFM entity and disables the EFM connection. Therefore, the connection timeout time must be greater than the handshake packet sending interval (recommended to be 3 times or more). Otherwise, the EFM connection will become unstable.

43.2.3 Configure the Remote Fault Detection Function

Configure the Remote Fault Detection Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	interface ethernet *interface-num*	-
Enable the remote fault detection function	**efm remote-failure** { **link-fault** | **dying-gasp** | **critical-event** }	Optional By default, the remote fault detection function is enabled
Disable the remote fault detection function	**no efm remote-failure** { **link-fault** | **dying-gasp** | **critical-event** }	Optional

Note:

The remote fault detection function requires the device to support the single-link function to notify the emergency link event of the local to the remote. On the device that does not support the single-link function, after the local detects an emergency link event, it only reports the alarm to the local and can not notify the remote.

43.2.4 Configure the Link Monitoring Function

Configure the Link Monitoring Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	**interface ethernet** device/slot/port	-
Enable the link monitoring function	**efm link-monitor** { **errored-symbol-period** | **errored-frame** | **errored-frame-period** | **errored-frame-seconds** }y	Optional. By default, the link monitoring function is enabled
Disable the link monitoring function	**no efm link-monitor** { **errored-symbol-period** | **errored-frame** | **errored-frame-period** | **errored-frame-seconds** }	optional
Configure the errored-symbol-period event detection interval	**efm link-monitor errored-symbol-period window high** *win-value1* **low** *win-value2*	optional
Configure the errored-symbol-period event detection threshold	**efm link-monitor errored-symbol-period threshold high** *th-value1* **low** *th-value2*	optional
Configure the errored-frame event detection interval	**efm link-monitor errored-frame window** *win-value*	optional
Configure the errored-frame event detection threshold	**efm link-monitor errored-frame threshold** *th-value*	optional
Configure the errored-frame-period event	**efm link-monitor errored-frame-period window** *win-value*	optional

detection interval		
Configure the errored-frame-period event detection threshold	**efm link-monitor errored-frame-period threshold** *th-value*	optional
Configure the errored-frame-seconds detection interval	**efm link-monitor errored-frame-seconds window** *win-value*	optional
Configure the errored-frame-seconds event detection threshold	**efm link-monitor errored-frame-seconds threshold** *th-value*	optional

Note:

The detection period and threshold of the errored-symbol-period event is a 64-bit integer value. The parameter values after high and low represent the upper 32 bits and the lower 32 bits of this value, respectively, i.e., the integer value = (high * (2 ^ 32)) + low.

43.2.5 Enable Remote Loopback

By default, the remote loopback function is disabled. You can enable remote loopback only on devices that support remote loopback.

Enable Remote Loopback

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	interface ethernet *interface-num*	optional
Enable the remote loopback function	**efm remote-loopback**	optional
Disable the remote loopback function	**no efm remote-loopback**	optional

43.2.6 Reject Remote Loopback Request From Remote Side

To avoid the problem that normal services is affected by the remote loopback function, you can use the configuration to prevent the local port from being controlled by the Loopback Control OAMPDU from the opposite side, thus rejecting the EFM remote loopback request initiated by the opposite side.

Reject Remote Loopback Request From Remote Side

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	
Enter the port configuration mode	interface ethernet *interface-num*	
Reject remote loopback request from remote side	**efm remote-loopback ignore**	Optional By default, the remote loopback request from the remote side is denied
Process the remote loopback request	**efm remote-loopback process**	optional

from the remote side		

43.2.7 Enable the Remote MIB Variable Acquisition Function

Enable the Remote MIB Variable Acquisition Function

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	interface ethernet *interface-num*	optional
Enable the remote MIB variable acquisition function	**efm variable-retrieval**	Optional. By default, the remote MIB variable acquisition function is enabled

43.2.8 Initiate the Remote MIB Variable Acquisition Request

Initiate the Remote MIB Variable Acquisition Request

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the port configuration mode	interface ethernet *interface-num*	optional
Get the port MIB variable value of the remote device	**show efm port** *port-id-list* **remote-mib** { **phyadminstate** | **autonegadminstate** }	optional
Get the global MIB variable value of the remote device	**show efm remote-mib** { **fecability** | **fecmode** }	optional

Note:

Only when the EFM connection on the port is established, the EFM working mode is in active mode and the remote port supports the remote MIB variable acquisition function, in this case, the remote MIB variable acquisition request can be initiated on the port.

Currently, only the FEC capability, FEC mode, port enabled state, and port auto-negotiation enabled state can be queried. Other MIB variables can be supplemented according to requirements.

43.2.9 EFM Display and Maintenance

After completing the above configuration, you can use the following command to display the EFM configuration.

EFM Display and Maintenance

Operation	Command	Remarks
Display the running status of the EFM protocol	**show efm status interface** [*interface-name*]	Any mode is executable
Display EFM summary information	**show efm summary**	

Display EFM discovery information	**show efm discovery interface** [*interface-name*]	
Display EFM protocol packet statistics	**show efm statistics interface** [*interface-name*]	
Clear EFM protocol packet statistics	**clear efm statistics interface** [*interface-name*]	Global configuration mode

43.3 Configuration Example

Switch(config)#i e 0/0/2
Switch(config-if-ethernet-0/0/2)#efm // Start EFM
Switch(config-if-ethernet-0/0/2)#efm mode passive // Set the EFM working mode to passive
Switch(config-if-ethernet-0/0/2)#efm pdu-timeout 1 // Set the interval for sending EFM handshake packets to 1s
Switch(config-if-ethernet-0/0/2)#efm link-timeout 5 // The connection timeout period is 5s
Switch(config-if-ethernet-0/0/2)#efm remote-response-timeout 2 // The response timeout period is 2s
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-symbol-period // Enable the link monitoring function
Switch (config-if-ethernet-0/0/2)#efm link-monitor errored-frame-period
Switch (config-if-ethernet-0/0/2)#efm link-monitor errored-frame-seconds
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-symbol-period window high 1 low 3 // Set the errored-symbol-period detection period to 1 to 3
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-symbol-period threshold high 1 low 3 // The detection threshold is from 1 to 3
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-frame window 20 // The detection period of the errored-frame event is 20
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-frame threshold 2 // The detection threshold is 2
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-frame-period window 2 // The detection period of the errored-frame-period event is 2
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-frame-period threshold 2 // The detection threshold is 2
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-frame-seconds window 200 // Set the errored-frame-second event detection interval to 200
Switch(config-if-ethernet-0/0/2)#efm link-monitor errored-frame-seconds threshold 2

44.BFD

44.1 BFD Function Overview

BFD is a simple "Hello" protocol, which in many respects resembles the neighbor detection part of the well-known routing protocols. A pair of systems periodically sends test packets over the channels on which they establish sessions, if a system does not receive the test packet from the opposite side within a sufficient period of time, it considers that a fault occurs in some part of the bi-directional channel to the adjacent system. In some cases, in order to reduce the load, the transmission and reception rates between the systems need to be negotiated.

BFD needs to set up peer-to-peer sessions at both ends of the channel before detection. After the session is established, the BFD control packets are sent to each other at the negotiated rate to implement fault detection.

The session establishment process is a three-way handshake process. After this process, the sessions at both ends become Up, and the corresponding parameters are negotiated at the same time. The later state change is based on the detection result of the defect, and the corresponding processing is done.

The BFD protocol describes the mechanism of bidirectional detection. It can be divided into two modes: asynchronous mode and query mode. There is also an auxiliary function: echo function, which can be used in combination with these two modes. The essential difference between asynchronous mode and query mode is that the position of detection is different. In asynchronous mode, the local side sends BFD control packets at a certain transmission interval and needs to detect the BFD control packets sent from the local system at the remote side. In query mode, BFD control packets sent from the local side are detected at the local system.

In asynchronous mode, the system periodically sends BFD control packets to each other. If a system does not receive the BFD control packet from the opposite side within the detection time, it declares the session to be Down.

In query mode, it is assumed that each system has a separate method to confirm that it is connected to other systems. So that once a BFD session is established, the system stops sending BFD control packets unless a system needs to explicitly verify connectivity. In case the connectivity needs to be verified explicitly, the system sends a short series of BFD control packets. If no response packet is received within the detection interval, the session is declared Down. If the response packet is received from the opposite side, the protocol remains silent again.

The local system sends a series of BFD echo packets, and the remote system loops them back through its forwarding channel. If the local system does not receive several echo packets in succession, the session is declared Down. Echo function can be used with the above two detection modes, you can use the echo function instead of the BFD control packet detection task, this reduces the transmission period of the control packet(in asynchronous mode) or completely disables the BFD control packet (in query mode).

Note:

Currently, BFD is used for OSPF and stacking. This document is only used for OSPF.

44.2 BFD Configuration

44.2.1 Enable/disable BFD

Enable/disable BFD

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	-
Enable the BFD function	bfd enable	required
Disable the BFD function	bfd disable	required

44.2.2 Apply to OSPF

Apply to OSPF

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	-
Enter interface configuration mode	interface vlan-interface ***if-num***	required
Configure Bfd to apply to OSPF	ip ospf bfd	required
Disable Bfd to apply to OSPF	no ip ospf bfd	required

44.2.3 Configure the Session Mode

BFD session establishment is divided into active and passive two modes. If a device is in active mode, BFD control packets are actively sent regardless of whether a BFD control packet is received from the opposite side when the session is established. If a device is in passive mode, BFD control packets will not be actively sent when the session is established, and will not be sent until the BFD control packet from the opposite side is received. At least one of the two ends of the BFD session needs to be in active mode. Active mode is used by default.

Configure the Session Mode

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	-
Enter interface configuration mode	interface vlan-interface ***if-num***	required
Configure the session mode	bfd session init-mode { active | passive }	optional

44.2.4 Configure the Query Mode

Configure the Query Mode

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	-
Enter interface configuration mode	interface vlan-interface ***if-num***	required
Enable the query mode	bfd demand on	optional
Disable the query mode	bfd demand off	optional

44.2.5 Configure Time Parameters

Configure Time Parameters

Operation	Command	Remarks
Enter the global configuration mode	configure terminal	-
Enter interface configuration mode	interface vlan-interface ***if-num***	required
Configure the minimum receiving interval	[no] bfd min-receive-interval ***value***	optional
Configure the minimum transmission interval	[no] bfd min-transmit-interval ***value***	optional
Configure the detection time	[no] bfd detect-multiplier ***value***	optional

44.2.6 Information Display and Maintenance

Information Display and Maintenance

Operation	Command	Remarks
View session	show bfd session [verbose]	optional

information		
Display information about interface configuration	show bfd interface [verbose]	optional

44.3 BFD Configuration Example

1.Network requirements

SW1 and SW2 are connected through interface vlan 1 and run OSPF protocol. In order to immediately detect the reachability of ospf neighbors, the bfd function is enabled.(Ospf configuration is not listed here, please refer to the OSPF user manual part. In addition to bfd run the necessary configuration, all other use the default configuration.)

sketch map of BFD

2.Configuration steps

SW1 CONFIGURATION

Enable the global bfd function and enable ospf bfd on interface 1；

SW1(config)#bfd enable

SW1(config-if-vlanInterface-1)#ip ospf bfd

Display the bfd information of the interface

SW1(config-if-vlanInterface-1)#s bfd interface

Global bfd state: enable

Interface: Vlan-interface1, Session Num: 1

Min Trans Inter: 400ms, Min Recv Inter: 400ms

DetectMult: 5, Min Echo Recv Inter: 0ms

Auth mode: NULL

Interface protocol: OSPF

#SW2 CONFIGURATION:

SW2(config)#bfd enable

SW2(config-if-vlanInterface-1)#ip ospf bfd

3.Result validation

Create a bfd session and display the following information:

SW1(config)#show bfd session

Total Session Num: 1

Init Mode: Active

Session Working Under Asynch Mode

LD	SourceAddr	DestAddr	State	Holdtime	Interface
0x06bb1f14	192.168.4.24	192.168.4.52	UP	1620ms	Vlan1

45.LLDP

45.1 LLDP Overview

LLDP (Link Layer Discovery Protocol) is a Layer 2 discovery protocol defined in IEEE 802.1AB. Through the use of LLDP technology, when the network scale is expanded rapidly, the network management system can quickly understand the Layer 2 network topology information and topology change information.

The basic principle of LLDP is as follows: A device in a network sends a notification of its status information to its neighboring devices, and each port of each device stores its own information. If a local device has a state change, it can also send updated information to the neighbor device directly connected to it. The neighbor device stores the information in the standard SNMP MIB. The network management system can inquire the connection situation of the current second layer from the SNMP MIB. Note that LLDP is a remote device state information discovery protocol. It can not perform the functions of network device configuration and port control..

45.2 LLDP Configuration

45.2.1 Enable/disable the LLDP

By default, LLdp is disabled. LLDP is enabled in global mode. After the configuration, the function takes effect.

Enable/disable the LLDP

Operation	Command	Remarks
Enter global configuration mode	configure terminal	optional
Enable/disable the LLDP	[no] lldp	required

45.2.2 Configure the Working Mode

LLDP working mode:

TxRx: Send and receive LLDP packets, the port works in this mode by default.

Tx: Only send LLDP packets, but do not receive.

Rx: Only receive LLDP packets, but do not send.

Disable: LLDP packets are neither sent nor received.

The port will initialize the protocol state machine if the LLDP working mode changes.

Sending mechanism:

When the port works in TxRx or Tx mode, the device periodically sends LLDP packets to the neighbor. If the local configuration of the device changes, LLDP packets are sent immediately to notify the neighboring devices of the changes in the local information. However, to prevent large amount of LLDP packets from being sent due to frequent changes of local information, after sending an LLDP packet, it needs to be delayed for a period of time before next sending.

Receiving mechanism:

When the port works in TxRx or Rx mode, the device checks the validity of received LLDP packets and TLVs carried in the packets. After checking, the neighbor information is saved to the local device and the aging time of the neighbor information on the local device is set according to the TTL (Time To Live) value in the TLV. If the value is zero, the neighbor information is aged out immediately.

Configure the Working Mode

Operation	Command	Remarks
Port configuration mode	interface ethernet ***port-num***	required
Configure the working mode	lldp [rxtx | tx | rx]	optional
Disable the LLDP	no lldp	optional

Note:

There is no separate command to turn off the working mode.

45.2.3 Configure Time Parameters

Configure Time Parameters

Operation	Command	Remarks
Enter global configuration mode	configure terminal	required
Configure Hello-time	[no] lldp hello-time ***value***	optional
Configure Hold-time	[no] lldp hold-time ***value***	optional

45.2.4 Configure the Management Address

The Layer 2 device does not support Management Address TLV.

By default, the Layer 3 device uses the IP address of the PVID interface. If there is no

corresponding interface IP for the corresponding vlan, the Management Address TLV is not sent in the LLDPDU. Use the following command to modify. The loopback interface is currently not supported.

Configure the Management Address

Operation	Command	Remarks
Port configuration mode	interface ethernet ***port-num***	required
Configure the management address	[no] lldp management-address **{** supervlan-interface ***value*** **|** vlan-interface ***value*** **}**	optional

45.2.5 Information Display and Maintenance

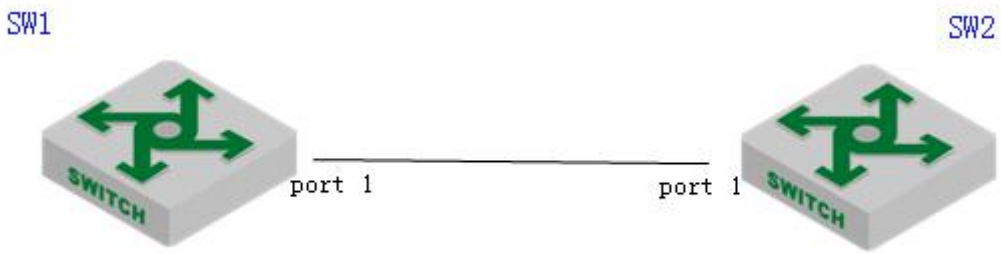
Information Display

Operation	Command	Remarks
Enter the global configuration	configure terminal	required
Display information	show lldp interface [ethernet ***port-num***]	optional

45.3 Configuration Example

1. Network requirements

SW1 and SW2 are directly connected through port 1, and then open the lldp function.



sketch map of networking

2.Configuration steps

#SW1 CONFIGURATION

SW1(config)#lldp

#SW2 CONFIGURATION
SW2(config)#lldp

3.Result validation

View the configuration and lldp neighbor information
SW1(config)#show lldp interface ethernet 0/0/1
System LLDP: enable
LLDP hello-time: 30(s) LLDP hold-time: 4 LLDP TTL: 120(s)
Interface Ethernet 0/0/1
Port LLDP: rxtx Pkt Tx: 158 Pkt Rx: 160
Total neighbor count: 1
Neighbor (1):
TTL: 106(s)
Chassis ID: 00:0a:5a:20:4d:ad
Port ID: port e0/0/1
System Name: SW2
System Description: New GreenNet Switch
Port Description: NULL
Management Address: 192.168.4.52
Port Vlan ID: 4
Port SetSpeed: auto
Port ActualSpeed: FULL-1000
Port Link Aggregation: support ,not in aggregation

46.UDLD

46.1 UDLD Overview

The actual network sometimes has the following situation: fiber cross connect; an optical fiber is not connected; one line of a fiber or the twisted pair is disconnected. In those cases, one of the ports at both ends of the link can receive the link-layer packet from the opposite side, but the opposite side cannot receive the packet from the local side. This link is a unidirectional link. In a unidirectional link, because the physical layer is in the connected state and can work normally, the detection mechanism of the physical layer (such as auto-negotiation mechanism) cannot find the communication problems between the devices, resulting in incorrect forwarding of traffic.

The role of UDLD (UniDirectional Link Detection) is to detect the existence of unidirectional links and take corresponding measures. It is responsible for monitoring the link status of physical lines on devices connected by fiber or copper twisted pair. When a unidirectional link is found, the device sends alarm information to the user. And according to the user configuration, the device automatically shuts down or notifies the user manually close the corresponding port to prevent network problems.

UDLD Protocol Status:

Status	Description
Initial	The initialization state of disabled UDLD protocol
Inactive	The UDLD protocol is enabled, but the link is Down
Active	The UDLD protocol is enabled and the link is Up, or the neighbor entry is cleared
Undetermined	The new neighbor is discovered in the normal working mode. The neighbor enters the undetermined state when no echo packet is received before the echo wait timer expires (5 seconds). When all the neighbors are in the undetermined state, the port enters the undetermined state
Unidirectional	The neighbor enters the unidirectional state due to the presence of a cross link or a unidirectional link. When all the neighbors enter the unidirectional state, the port will enter the unidirectional state. Output log and trace information. Send the flush packet. Manually or automatically shut down the local port according to the UDLD Down mode configured by the user. The device can only receive and send UDLD packets and delete the neighbor entry.
Bidirectional	Neighbors can correctly answer. When all neighbors are bidirectional, the port is bidirectional.
DelayDown	When the port is linkdown, the neighbor is not deleted immediately. Instead, it enters the temporary DelayDown state. In

	this case, the port can only respond to the linkup event. In this state, the UDLD neighbor information is retained and the DelayDown timer is started.

UDLD Working Process:

1) Neighbor Discovery: When a port is active, it sends its own information through the probe packet to request neighbor information. After receiving the probe packet, the opposite port implements the neighbor discovery based on the content information of the probe packet. When the port receives a probe packet, it determines whether the sending port is already in the neighbor table. If it is not, indicating that it is a new neighbor, it replies with the echo packet and adds it to the neighbor table and marks it as undetermined. If the sending port is already in the neighbor table but the RSY flag is set in the packet, an echo packet is sent to the port. If the neighbor is a bidirectional neighbor, the neighbor information is updated.

2) Neighbor aging: After the bidirectional neighbor is added to the neighbor table, the port sets an aging time T1f according to the contents of the probe packet, the hello packet and the echo packet. When time T1f is over, the port does not receive any new hello packets from the neighbor. The neighbor is aged out and is deleted from the neighbor table. Normally, the value of the aging time T1f is greater than the keepalive packet transmission interval Tmsg(T1f=3*Tmsg) of the neighbor. The value of Tmsg is obtained from the TLV of the hello packet.

3) Unidirectional probing: A port will perform a unidirectional probing only when there is a change in the neighbor list. The probe initiator first initiates a probe packet(RSY=1) with a synchronization request. After receiving the packet, the opposite side responds with an echo packet and adds its own neighbor table information to the echo-tlv field of the echo packet. If the initiator receives the echo packet and checks that the echo-tlv field is correct, it considers the port to be bidirectional. If the content of the received echo packet is not correct, the port is considered to be unidirectional. If no echo packet is received, there are different processing methods depending on the UDLD check mode. The processing method will be described in the following sections.

4) Unidirectional processing: After the port state is determined to be unidirectional, the neighbor list of the port is cleared, and then the FLUSH packet is sent to notify the neighbor of deleting the port information, the port is shut down. According to the port shutdown mode, recovery methods are inconsistent.

5) Keepalive mechanism： After the port is in stable state, the port periodically sends hello packets to notify other ports of its status. The opposite side will use this packet to refresh the neighbor state. If no hello packet is received within the keepalive period, the port will be deleted from the neighbor table. Hello packets carry all the neighbor information of the port. The sending period Tmsg of hello packets can be set using the global command.

The received UDLD packets are processed as follows:

Packet Type	Processing	
Hello	Remove the	If there is no neighbor entry on the local device, the packet is discarded

<table>
<tr><td>packet</td><td>neighbor information （device-id and port-id）</td><td colspan="4">If the neighbor entry exists on the local device and echo-TLV is correct, the entry aging timer is refreshed if the neighbor is bidirectional. Otherwise, the entry is discarded. If the neighbor entry exists and the echo-TLV is incorrect, it indicates that a cross-link occurs and the neighbor is marked as a cross-link neighbor. If the neighbor is bidirectional, it is set to unidirectional. If all the neighbors are unidirectional, the port enters the unidirectional state and sends the Flush packet. The port will be shut down manually or automatically according to the configured UDLD Down mode, and the neighbor entry will be deleted.</td></tr>
<tr><td>Flush packet</td><td colspan="5">Delete the neighbor entry on the local device. If the neighbor entry is empty, the port enters a unidirectional state. Otherwise, the port status is determined according to the status of all the neighbors.</td></tr>
<tr><td rowspan="2">Probe packet</td><td rowspan="2">Send the Echo packet containing neighbor information and its own information to the opposite side.</td><td colspan="4">If the neighbor does not exist on the local device, the neighbor entry is created</td></tr>
<tr><td colspan="4">If the neighbor entry exists on the local device, the entry aging timer is refreshed.</td></tr>
<tr><td rowspan="3">Echo packet</td><td rowspan="3">Check whether the neighbor information carried in the packet is the same as that of the local device</td><td>different</td><td colspan="3">A cross-link occurs, if the machine has a related neighbor, the neighbor is marked as a cross-link neighbor, if the neighbor is bidirectional, it is set to unidirectional. If all the neighbors are unidirectional, the port enters the unidirectional state and sends the Flush packet. The port will be shut down manually or automatically according to the configured UDLD down mode, and the neighbor entry will be deleted.</td></tr>
<tr><td rowspan="2">same</td><td rowspan="2">The device detects whether a related neighbor exists based on the device-id and port-id in the packet</td><td>existent</td><td>If the neighbor is in the undetermined or unidirectional state, the neighbor is set to bidirectional. If all the neighbors are bidirectional, the port enters the bidirectional state. Refresh neighbor entries.</td></tr>
<tr><td>nonexistent</td><td>The packet is discarded</td></tr>
</table>

Note:

S5900-24S does not support unidirectional, if both ends are S5900-24S, you can not detect unidirectional link. So if UDLD is to be used, S5900-24S needs to be connected with other vendor devices that support unidirectional..

46.2 UDLD

46.2.1 Enable/disable UDLD

UDLD has global switch and port switch, you must enable both global switch and port switch simultaneously to work normally, the default configuration are turned off. DUT connected ports must be enabled the function to work normally.

Enable/disable UDLD

Operation	Command	Remarks
Enter the global configuration	configure terminal	required
Enable/disable UDLD	[no] udld	required
Port configuration mode	interface ethernet ***port-num***	optional
Enable/disable UDLD	[no] udld	required
Display udld information	show udld interface [ethernet ***port-num***]	optional

46.2.2 Reset

When UDLD detects a unidirectional link, and the port is Down, the port state can be reset through reset command, and then re-perform the UDLD.

Reset

Operation	Command	Remarks
Enter the global configuration	configure terminal	optional
Configure reset	udld reset	optional
Port configuration mode	interface ethernet ***port-num***	optional
Configure reset	udld reset	optional

46.2.3 Configure Time Parameters

Configure Time Parameters

Operation	Command	Remarks
Enter the global configuration	configure terminal	optional
Configure hello send interval	udld message-interval ***value***	optional
Configure the down delay	udld delaydown-time ***value***	optional

46.2.4 Configure the Working Mode

UDLD works in two modes: Normal and Aggressive. The default is Normal.

In normal mode, if a port does not receive a packet from the opposite side, the neighbor is aged out. And the port is in an undetermined state; If the port receives a probe keepalive packet or an echo packet without the information of the local port, which causes all neighbors of the local port to be aged out, the port is considered to be in a unidirectional state. Normal mode is used to check for the unidirectional link due to cross-connections. S5900-24S hardware temporarily does not support the unidirectional caused by cross-connections.

In aggressive mode, a port does not receive a packet from the opposite side or receives packet without the local port information, as a result, all its neighbors are aged out, and the port is considered to be in a unidirectional state. Aggressive mode is used to check for the unidirectional connection caused by fiber-optic cross-connections or disconnections..

The following table compares the two modes:

Working mode	Whether to actively detect the presence of a neighbor when aging the neighbor table	Whether to enable the Entry aging timer when aging the neighbor table	Whether to enable the aggressive timer when the Entry aging timer expires
Normal mode	Active detection is not performed	Yes（After the Entry aging timer expires, it directly ages the neighbor entry）	No
Aggressive mode	Active detection is performed	Yes（After the Entry aging timer expires, the aggressive timer is enabled ）	Yes（After the aggressive timer expires, the neighbor state is set to unidirectional, and the neighbor entry is aged out）

Configure the Working Mode

Operation	Command	Remarks
Port configuration mode	interface ethernet ***port-num***	optional
Configure the working mode	udld work-mode { normal | aggressive }	optional

46.2.5 Configure Unidirectional Processing

When a port is detected as a unidirectional port, a different processing is adopted depending on the working mode. Send out the alarm information (need to turn on the logging function). If it is Auto mode, it will automatically shut down / restore the port; if it is Manual, the administrator needs to shut down / restore port at the command line.

Configure Unidirectional Processing

Operation	Command	Remarks
Enter port mode	interface ethernet ***port-num***	optional
Configure the working mode	udld unidirectional-shutdown { Auto | Manual }	optional
Manually close in unidirectional state	udld port shutdown	optional
Restore closed port	no udld port shutdown	optional

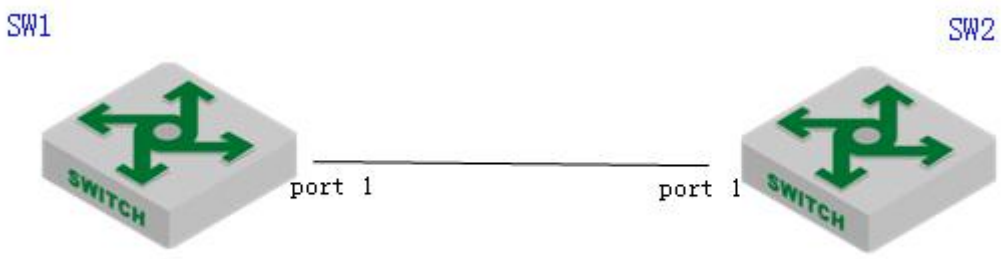
Note:

"[no]udld port shutdown "Only in the Manual mode need to configure；

46.3 Configuration Example

1. Network description

SW1 and SW2 are directly connected through port 1, and then turn on UDLD function.



sketch map of networking

2.Configuration steps

#SW1 CONFIGURATION:

SW1(config)#udld

SW1ch(config)#interface ethernet 0/0/1

SW1(config-if-ethernet-0/0/1)#udld

SW1(config-if-ethernet-0/0/1)#udld work-mode aggressive

#SW2 CONFIGURATION:

SW2(config)#udld

SW2(config)#interface ethernet 0/0/1

SW2(config-if-ethernet-0/0/1)#udld

SW2(config-if-ethernet-0/0/1)#udld work-mode aggressive

3.Result validation

When the unidirectional state is detected, the following information is displayed:

SW1(config)#show udld interface e 0/0/1

Udld state : ON

Message interval : 15

TimeOut interval : 5

DelayDown time : 1

port udld information:

Port ID : e0/0/1

NOTICE:this port'ingress was shutdown by manual!

Port udld state : ON

Port status : Unidirectional

Port link status : LinkUp

Port udld work mode : Aggressive

Port shutdown mode : Auto

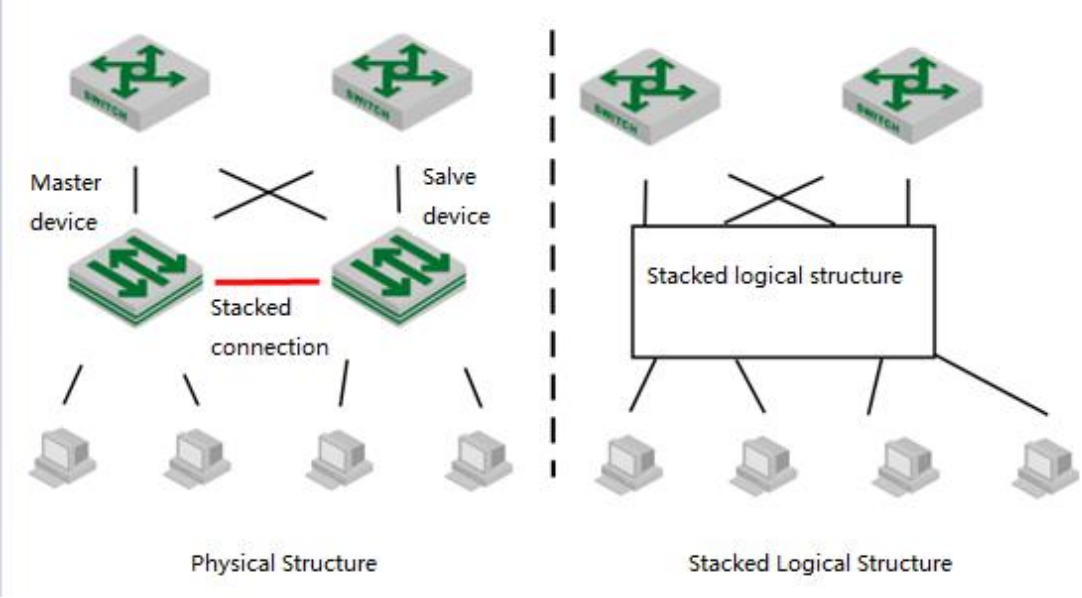
neighbor information:

Total Neighbor Record: 0

47.Stack

47.1 Stack Overview

Stack means multiple devices are connected via stack port to form a fictitious logic device, and users perform the management via performing the management on fictitious logic device.



schematic diagram of stack networking applications

(1) Technical advantages of stack

Stack possesses the following advantages:

- Network scalability. At the early stage of the network construction, it uses less devices to build the network and it will expand the port number and bandwidth via adding the stack devices in middle and later periods of the network construction.
- Reliability. Stack system is made up by a master device and multiple slave devices. Master device takes the responsibility to finish the administration and maintenance of the stack system whereas slave devices take part in service data processing. If there is something wrong with master device, system will select a new master device to perform the backup job. In addition, physical ports between devices support aggregation function to help to finish the port backup job.
- Available management. Any port of any device in stack system can be able to login stack system to perform management configuration, and there is no need to perform separate management configuration on each member device.
- Low cost on operations and maintenance. Network upgrading needn't to replace existing devices, just add the new devices will be OK. Multiple devices form a logic

device can effectively reduce maintenance cost

(2) Stack basic conception

- master-slave devices

There are two types devices in stack system:

Master device is the management device of the stack system.

Slave device. The backup device of the master device and it will be selected as the master device if there is something wrong with the master device.

Master-slave equipment is produced automatically by the system. There is only one master device and multiple slave devices in a stack system.

- Device ID

Each device in the system must be manually specified a non-repeatable to uniquely identify the ID number of this equipment, and the device port is shown as “device ID/ slot number/ port number”, for example, “0/0/1” means that device ID is “0”, and its slot number is 0 and port number is 1.

All devices in one stack system should not exist the same ID.

The smaller the device ID is, the higher priority it will be.

- Stack port

Devices connect with each other via stack port in stack system. Stack port can use dedicated stack interface or normal device port to perform devices connection.

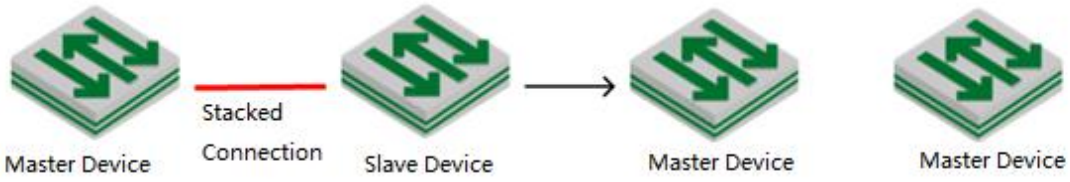
All device stack port in one stack system should be the same or it cannot form the stack. For example, if SW1 uses port{1,2,3,4} as stack port, SW2 must use port{1,2,3,4} as stack port as a consequence. Or these two DUT cannot form stack.

- service port

Other ports are called as service port except stack port.

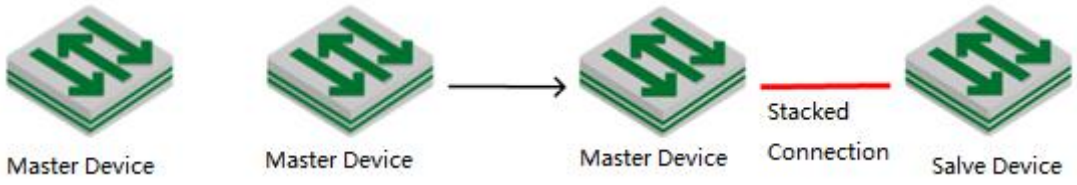
- Split and merge

In stack system, it will split two independent stack systems if the device port lose connection.



schematic diagram of stack split

In reverse, two independent stack systems will be merged in one stack system because of the adding connection of the stack ports.



(3) Stack Topology

Currently, stack can only use the devices with the same capacities to help to form loop topology structure and rectilinear topology structure.

“same capacities” refers to the same software version number and the same hardware type.

(4) left-right virtual port

For example, stack is just like few people hand in hand to form a loop or rectilinear figure.

Left-right port is geared to virtual port, forming by one or multiple physical stack port or no physical stack port.

In stack system, stack port must be left-right virtual port connection, that is, the virtual port left-left virtual port connection and right-right virtual port connection are not allowed.

(5) Master-Slave Selection

In stack system, there will be a master device of automatic select according to stack protocol, and the other devices will act as slave devices.

To understand the master-slave selection mechanism, you need to understand different stack scenarios:

- Power on the devices at the same time: devices connect with each other via stack port before start up, and then power on at the same time；
- Add the devices which are booting to existing stack system: devices connect to an existing stack system through the stack port, then starts to power on.
- Add the devices which have finished booting to existing stack system: it is geared to stack merge. Devices have finished booting and formed its own stack, and then this stack system will connect with the other system via stack port. Stack merger occurs at this time.
- Re-select slave device: if there is something wrong with the master device or the stack splitting, device needs to re-select a master device.

Selection mechanism is as following:

- Running priority. This rule applies to the situation when the stack is added during boot process. If the device discovers the existing master device during boot process, it will add the existing master device to the stack system as a slave device.
- Number of members. This rule applies to the stack merging situation. In the process of stack merging, the master devices with large number of members will stay to be master device while the master devices with fewer number of members will turn into slave devices.
- Device ID. The device with small device ID will be selected to be master device in the following situation: multiple devices establish stack system at the same time; master device re-selection situation because of master device losing; members of the same number when stack merging.

(6) Device mode

stack equipment can work in two different working modes:

- Standalone Mode. It is the same as ordinary switch, that is, it does not offer stack function.
- Stack Profile. This mode can enable stack function as well as to form a stack system with other device.

(7)MAD Multi-Active Detection

If there is something wrong with a certain stack link, it will disconnect with Master device while it can still connect with multiple Slave devices and select one of them to be the new Master device. So there will be two or multiple master devices with the same configuration in stack system, and it is the so called Multi-Active.

As to the networking devices in the external stack system, logical devices corresponding to stack system are divided into two or multiple logical devices with the same configuration. Therefore, it will appear network configuration conflict and then lead to transmission confusion between uplink and downlink. For this reason, multi-active arises at the historic moment. Multi-Active can be able to detect stack line fault in the shortest time and then perform duly handle.

Therefore, it asks to ensure the following networking function when stack system splitting:

A. Multi-Active Detection

It judges whether there exists multiple logical devices from one stack system and they are now in **active** state via LACP and BFD.

B. Conflict resolution

After stack system splitting, it will be detected that there are multiple logical devices are in active state via multi-active detection mechanism. Conflict resolution will make the logical device with the highest ActivePriority continue the regular work (remain *Active* state)while the other logical devices will be moved to *Recovery* state (Disable state) via a certain election algorithm. When the logical devices are in the *Recovery* state, it will perform Recovery Action: disabled VLAN interface (except the saving-port) and all service port which are in the Recovery state to ensure the logical devices completely disconnected with the network.

C. Failure recovery

Stack system reminder user to restore the stack link by log record. After restoring the stack link, conflicting equipment will reboot and then restore the stack system. At the same time, the closed port will also restore the service transmission.

Note:

In MAD，use device-id as activePriority; save the small id number as active state, and the other as recovery state.

47.2 Stack

47.2.1 Stand-Alone Mode Configuration

Stand-Alone Mode Configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure stack	[no] stack	required
Configure device ID	stack device-id ***value***	required
Configure stack port	[no] stack { left-port | right-port } ***port-num***	required
Configure device priority	[no] stack priority ***value***	optional
Display stack information	show stack	optional

Note:

1. Stack function will go into effect after rebooting. Except for debug configuration, all stack configuration will be written into Flash and will not be shown in show running；

2. Stack configuration cannot be deleted via the command of *clear startup-config*, and it should be use the command of *no command* to delete one by one.

47.2.2 Stack Mode Configuration

Stack Mode Configuration

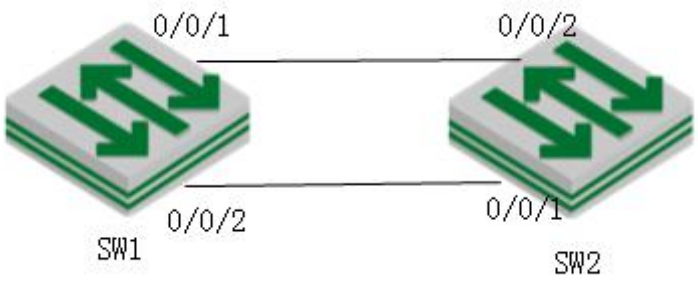
Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure device ID	stack device-id ***value***	optional
Configure stack port	[no] stack { left-port | right-port } ***port-num***	optional
Configure device priority	[no] stack priority ***value***	optional
Display stack information	show stack	optional
Configure stack debug function	[no] stack debug { all | default |error |event |packet | packet-detail | smd | smd-msg | state }	optional
Display stack debug function configuration	show stack debug	optional
Display stack member information	show stack members	optional
Display stack neighbor information	show stack neighbors	optional

Display statistical information of stack packet	show stack statistic	optional
Configure keeping active time	[no] stack hello-timeout ***value***	optional
Configure detection delay	[no] stack linkdown-delay	optional
Enter privileged mode	exit	required
Reboot and designate member	reboot ***device-id***	optional

47.2.3 Stack Configuration Examples

1.Network requirements

It is shown as following. SW1 acts as Master， and SW2 acts as Slave.



sketch map of stack networking

2. Configuration procedure

#SW1 configuration:

enable stack

SW1(config)#stack

configure device-id，0 by default，SW1 is optional to modify；

SW1(config)#stack device-id 0

configure SW1 left-port

SW1(config)#stack left-port 0/0/1

SW1(config)#stack left-port 0/0/2

configure the priority of SW1

SW1(config)#stack priority 200

reboot. Reboot to bring stack into effect

SW1(config)#ex

SW1#reboot

#SW2 configuration

SW2(config)#stack
SW2(config)#stack device-id 1
SW2(config)#stack right-port 0/0/1
SW2(config)#stack right-port 0/0/2
SW2(config)#stack priority 100
SW2(config)#exit
SW2#reboot

3. Verification result

After two switches rebooting, perform the connect operation according to network construction. SW2 will be selected as Slave and it will be reboot. After rebooting, the stack light of Master device will keep coruscating whereas the stack light of Slave device will keep bum steady. All configurations can only be performed in Master device.

(1)stack system port information after finishing the stack:

SW1(config)#show interface brief

Port	Desc	Link	shutdn	Speed	Pri	PVID	Mode	TagVlan	UtVlan
e0/0/3		down	false	auto	0	1	hyb		1
e0/0/4		down	false	auto	0	1	hyb		1
e0/0/5		down	false	auto	0	1	hyb		1
e0/0/6		down	false	auto	0	1	hyb		1
e0/0/7		down	false	auto	0	1	hyb		1
……									
e1/0/1		down	false	auto	0	1	hyb		1
e1/0/2		down	false	auto	0	1	hyb		1
e1/0/3		down	false	auto	0	1	hyb		1
e1/0/4		down	false	auto	0	1	hyb		1
e1/0/5		down	false	auto	0	1	hyb		1

(2) Display all stack members' information

SW1(config)#show stack members

Informations of stack devices:

switch 1 <local>

macaddress 00:01:7a:fd:ef:2d device id 0 priority 200

master device left hops 0 right hops 0

stack identity fdef2d003e4a

it's master device 00:01:7a:fd:ef:2d device id 0

switch 2

macaddress 00:01:7a:fd:ee:d2 device id 1 priority 100

slave device left hops 1 right hops infinite

stack identity fdef2d003e4a

it's master device 00:01:7a:fd:ef:2d device id 0

Total entries: 2

(3) Display all neighbors' information

```
SW1(config)#show stack neighbors
   Informations of neighbor devices:
   switch 1 <local>
      macaddress 00:01:7a:fd:ef:2d device id 0 priority 200
      master device left hops 0 right hops 0
      stack identity fdef2d003e4a
      it's master device 00:01:7a:fd:ef:2d device id 0

   switch 2
      macaddress 00:01:7a:fd:ee:d2 device id 1 priority 100
      slave device left hops 1 right hops infinite
      stack identity fdef2d003e4a
      it's master device 00:01:7a:fd:ef:2d device id 0

   Total entries:    2
```

(4)establish vlan，and then add members:

```
SW1(config)#vlan 100
SW1(config-if-vlan)#sw e 0/0/3 e 1/0/3
SW1(config-if-vlan)#show vlan 100
show VLAN information
VLAN ID                          : 100
VLAN status                      : static
VLAN member                        : e0/0/3,e1/0/3.
Static tagged ports        :
      Static untagged Ports     : e0/0/3,e1/0/3.
      Dynamic tagged ports      :

      Total entries: 1 vlan.
```

5)upgrade stack system:

```
      SW1#load application tftp inet 192.168.1.99 host.arj
      Downloading application via TFTP...
         Master device 0 operation complete, successful.
         Slave device 1 operation complete, successful.
      Download application via TFTP successfully.
```

47.2.4 LACP MAD

LACP MAD stack system

operation	command	remark
Enter global configuration mode	configure terminal	-
Enable lacp mad	[no] channel-group ***value*** lacp mad	required
Configure lacp mad domain	[no] lacp mad domain ***value***	optional
Display lacp mad information	show lacp mad	optional

LACP MAD relay device configuration

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Configure LACP expansion field	channel-group ***group-id*** extend-info-relay	required
Display LACP expansion field	show lacp extend-info-relay	optional

Configure MAD to save the port

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-num***	-
Configure mad to save the port	[no] stack mad exclude	required
Display port configuration	show stack mad exclude interface [ethernet ***port-num***]	optional

Note:

LACP MAD only supports dynamic lacp;

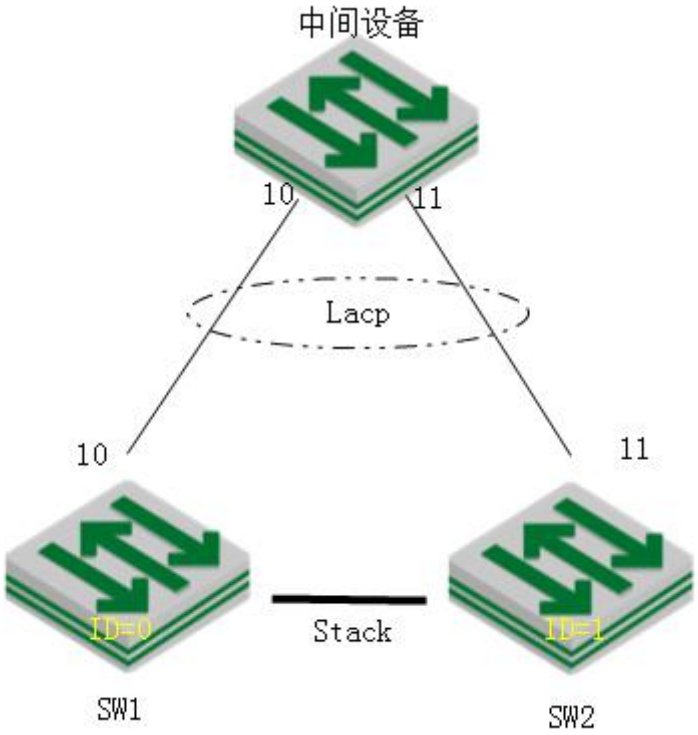
LACP MAD relay device should be able to support LACP expansion field, or it can not take effect；

LACP MAD configuration will not be written in Flash directly；

47.2.5 Configuration Examples for LACP MAD

1.Network requirements

Shown as follow. Two SW form stack system, construct network with the other SW via LACP.



LACP MAD schematic diagram

2.Configuration steps

stack configuration（please refer to stack configuration content for more details ）

configure dynamic LACP

SW1(config)# interface range ethernet 0/0/10 ethernet 1/0/10

SW1(config-if-ethernet-0/0/10)#channel-group 1 mode active

configure LACP MAD in stack system

enable lacp mad

SW1(config)#channel-group 1 lacp mad

configure mad domain (optional configuration, 255 by default)

SW1(config)#lacp mad domain 200

configure the port (optional)

SW1(config)#in e 0/0/12

SW1(config-if-ethernet-0/0/10)#stack mad exclude

configure LACP extension field in relay device configuration

MilddleDUT#configure terminal

MilddleDUT (config)# interface ethernet 0/0/10

MilddleDUT (config-if-ethernet-0/0/10)#channel-group 1 mode active

MilddleDUT (config-if-ethernet-0/0/10)#ex

MilddleDUT(config)#channel-group 1 extend-info-relay

3. verification result:

SW1(config)# logging monitor 0

SW1(config)#debug link_aggregation

SW1(config)#debug stack

manually disconnect all the stack line；

00:26:06: SW1: %stack-5-state: stack port 1/0/1 link down

00:26:08: SW1: %stack-5-state: stack port 0/0/1 link down

00:26:08: SW1: %stack-5-state: isf_devinfo_timeout_hello:device 1 leave stack mac 00:01:7a:fd:ee:d2

the MAD which has been detected ;

00:26:08: SW1: %LINK_AGGREGATION-7-isfMad: A multi-active conflict detected on channel group 1(local ActiveId = 0, peer ActiveId = 1).

only save and linkup port and lacp mad port, shutdown the other ports

SW1(config)#show interface brief

Port	Desc	Link	shutdn	Speed	Pri	PVID	Mode	TagVlan	UtVlan
e0/0/3		down	false	auto	0	1	hyb		1
e0/0/4		down	false	auto	0	1	hyb		1
e0/0/5		down	false	auto	0	1	hyb		1
e0/0/6		down	false	auto	0	1	hyb		1
e0/0/7		down	false	auto	0	1	hyb		1
e0/0/8		down	false	auto	0	1	hyb		1
e0/0/9		down	false	auto	0	1	hyb		1
e0/0/10		up	false	auto-f1000	0	1	hyb		1
e0/0/11		down	false	auto	0	1	hyb		1
e0/0/12		up	false	auto-f1000	0	1	hyb		1
e0/0/13		down	false	auto	0	1	hyb		1
e0/0/14		down	false	auto	0	1	hyb		1
……									
e0/0/21		down	false	auto	0	1	hyb		1
e0/0/22		down	false	auto	0	1	hyb		1
e0/0/23		down	false	auto	0	1	hyb		1

Total entries: 50 .

Note:

All the above are just to demonstrate the effect, the actual use don't need to open the debug.

47.2.6 BFD MAD

Configure BFD MAD—L3 devices

Operation	Command	Remarks
Enter global configuration mode	configure terminal	-
Enable/disable bfd	bfd { enable | disable }	required
Disable STP	no spanning-tree	required
Create a Layer 3 interface	interface vlan-interface ***vlan-id***	optional
Configure member ip of mad detection	[no] mad device-id ip address ***A.B.C.D mask***	required
Enable/disable mad bfd	mad bfd {enable|disable}	required
Display mad bfd information	show mad bfd	optional
Display bfd session information	show bfd session	optional

Configure the MAD reserved port

Operation	Command	Remarks
Enter port configuration mode	interface ethernet ***port-num***	-
Configure mad to save the port	[no] stack mad exclude	required
Display port configuration	show stack mad exclude interface [ethernet ***port-num***]	optional

Note:

1.BFD MAD port asks to disable the spanning tree to avoid detection failure on account of the port in Blocking state and then discard the BFD packet.

2. It asks to use BFD VLAN to isolate MAD port and the other ports so as to avoid the storm.

3.BFD MAD has no requirement on VLAN type of detection port（Access/Trunk/Hybrid）while it requires to ensure the VLAN connectivity.

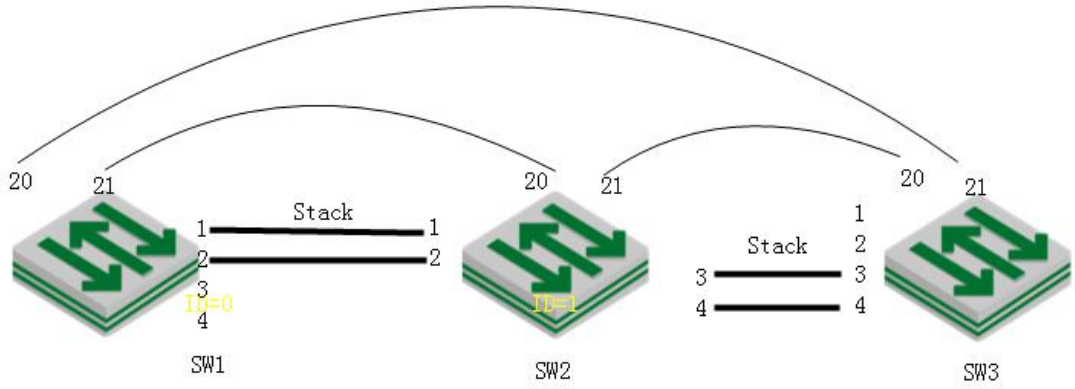
4. It asks to use the command of ***mad member member-id ip address***to configure MAD IP address. Remember do not configure other IP addresses(for example, use the command of ip address to configure normal IP address, interface IP address, VRRP vlan IP address, etc) for fear that it will affect MAD function.

5． It asks to ensure the mesh connection of all stack members when BFD MAD do not use relay device to construct networking; it asks to ensure each member has connected with relay device when using relay device to construct networking.

47.2.7 Configuration Examples for BFD MAD (L3 Devices)

1.Network requirements

Network requirements:



Bfd mad

2.Configuration steps

stack configuration:

SW1(config)#stack

SW1(config)#stack priority 200

SW1(config)#stack device-id 0

SW1(config)#stack left-port 0/0/3

SW1(config)#stack left-port 0/0/4

SW1(config)#stack right-port 0/0/1

SW1(config)#stack right-port 0/0/2

SW2(config)#stack

SW2(config)#stack priority 190

SW2(config)#stack device-id 1

SW2(config)#stack left-port 0/0/1

SW2(config)#stack left-port 0/0/2

SW2(config)#stack right-port 0/0/3

SW2(config)#stack right-port 0/0/4

SW3(config)#stack

SW3(config)#stack priority 170

SW3(config)#stack device-id 2

SW3(config)#stack left-port 0/0/3

SW3(config)#stack left-port 0/0/4

SW3(config)#stack right-port 0/0/1

SW3(config)#stack right-port 0/0/2

MAD BFD port configuration: using vlan100 interface to perform the MAD detect with each SW 20 and SW 21 port, and then save the BFD port；

```
SW1(config)#vlan 100
SW1(config-if-vlan)#switchport    ethernet 0/0/20 to ethernet 0/0/21
SW1(config-if-vlan)#switchport ethernet 1/0/20 to ethernet 1/0/21
SW1(config-if-vlan)#switchport ethernet 2/0/20 to ethernet 2/0/21
SW1(config-if-vlan)#interface range ethernet 0/0/20 ethernet 0/0/21
SW1(config-if-range)#switchport default vlan 100
SW1(config-if-range)#no spanning-tree
SW1(config-if-range)#stack mad exclude
SW1(config-if-range)#interface range ethernet 1/0/20 ethernet 1/0/21
SW1(config-if-range)#switchport default vlan 100
SW1(config-if-range)#no spanning-tree
SW1(config-if-range)#stack mad exclude
SW1(config-if-range)#interface range ethernet 2/0/20 ethernet 2/0/21
SW1(config-if-range)#switchport default vlan 100
SW1(config-if-range)#no spanning-tree
SW1(config-if-range)#stack mad exclude
```

#MAD BFD configuration

```
SW1(config)#interface vlan-interface 100
SW1(config-if-vlanInterface-100)#mad bfd enable
SW1(config-if-vlanInterface-100)#mad device-id 0 ip address 20.20.20.20 255.255.255.0
SW1(config-if-vlanInterface-100)#mad device-id 1 ip address 20.20.20.21 255.255.255.0
SW1(config-if-vlanInterface-100)#mad device-id 2 ip address 20.20.20.22 255.255.255.0
```

3.Result validation

(1) when under normal circumstances, BFD session state will be as following:

```
SW1(config)#show bfd session
Total Session Num: 2
Init Mode: Active
Session Working Under Asynch Mode
LD            SourceAddr      DestAddr        State     Holdtime Interface
0x0add68e8 20.20.20.20     20.20.20.21     DOWN          0ms Vlan100
Init Mode: Active
Session Working Under Asynch Mode
LD            SourceAddr      DestAddr        State     Holdtime Interface
0x0add699c 20.20.20.20     20.20.20.22     DOWN          0ms Vlan100
```

(2) manually disconnect all the STACK line between SW2 and SW3， the conflict that BFD MAD has detected will be shown as following:

```
SW1(config)#s bfd session
Total Session Num: 2
Init Mode: Active
```

Session Working Under Asynch Mode

LD	SourceAddr	DestAddr	State	Holdtime	Interface
0x0add68e8	20.20.20.20	20.20.20.21	DOWN	0ms	Vlan100

Init Mode: Active

Session Working Under Asynch Mode

LD	SourceAddr	DestAddr	State	Holdtime	Interface
0x0add699c	20.20.20.20	20.20.20.22	UP	1890ms	Vlan100

(3) Enter SW3, you can see the Stack processing recovery state:

```
SW1(config)#show stack
  Config in flash:
    enable stack, device id is 2, priority is 200
    left port    2/0/3 2/0/4
    right port    2/0/1 2/0/2

  Config in running:
    enable stack, device id is 2, priority is 170
    left port    2/0/3 2/0/4
    right port    2/0/1 2/0/2

  Local device is master devcie, state is STATE_MASTER

  Linkdown-delay is FALSE, hello-timeout is 30

  Mad status: recovery

  Left-port load-sharing mode: source port
  Right-port load-sharing mode: source port

  Infomation of stack port:
    stack port 2/0/1 is link down speed is unknown
    stack port 2/0/2 is link down speed is unknown
    stack port 2/0/3 is link down speed is unknown
    stack port 2/0/4 is link down speed is unknown
```

48.MPLS Basic Configuration

48.1 MPLS Introduction

Multi-protocol label switching (MPLS) integrates the simplicity of Layer-2 switching and the flexibility of Layer-3 routing, and brings the connection-oriented feature in connectionless IP networks. It is a hot technology in communication field at present. Because of its flexibility and high scalability, MPLS technology has become one of the most promising network technologies in the field of communication after less than ten years of development. It can be used to provide traffic engineering, QoS, Layer 2 and Layer 3 VPN applications.

Compare with the traditional IP network technology, MPLS has the following technical advantages: (1) Link layer protocol independence: From the protocol level, MPLS is located between the network layer and the link layer. The link layer can be Ethernet, ATM, PPP or Frame Relay and other protocols. (2) Flexibility of MPLS extension: By introducing labels, MPLS realizes complex routing and mapping of QoS information to Layer 2 routing information. It can extend the functions through the label stack and support Layer 2 and Layer 3 VPN functions. (3) Universality of the protocol: Unlike IP addresses, labels do not have specific meaning. Labels can be used to map destination IP addresses, Fiber Channel, wavelengths, or even VC channels in SDH / SONET.

In terms of standards, the IETF is the leading standardization organization for MPLS, and has developed a number of standards / drafts for MPLS, among which the more important ones include: RFC3031 (MPLS Architecture), RFC3036 (MPLS Label Stack), RFC3270 (MPLS Support Differentiated Services), and RFC3353 (IP Multicast in MPLS Networks). With the acceleration of the MPLS standardization process, applications based on MPLS technology will be more widely used in existing communication networks, and will become the mainstream of next-generation broadband network technology. MPLS L2 / L3 VPN is the network technology that is being used at present. QoS and traffic engineering are also important applications of MPLS. VPLS is becoming the popular application technology in metropolitan area network construction.

Our switch products provide support for MPLS and related functions, including:

- ✓ LDP (label distribution protocol) conforms to RFC 3036;
- ✓ L2VPN (Note: the function is not yet open) conforms to RFC4664 and related standards;
- ✓ The L3 VPN of MP-BGP conforms to RFC 2547;

48.1.1 MPLS Basic Concepts

1. Forwarding Equivalence Class

As a classification and forwarding technology, MPLS classifies the packets with the same forwarding method into one class, called FEC (Forwarding Equivalence Class).

FEC division is very flexible, and it can be divided into any combination based on source address, destination address, source port, destination port, protocol type, or VPN. For example, in traditional IP forwarding using the longest match algorithm, all packets to the same destination address are one FEC. It is common to assign the same label to a FEC on one device.

2. Label

It is a fixed-length, usually only a short identifier of having local significance. It is usually located between the data link layer encapsulation header and the Layer 3 packet, and is used to uniquely identify the FEC to which a packet belongs.

3. Label Switching Router

LSR (Label Switching Router) is the core switch of the MPLS network, which provides label switching and label distribution.

4. Label Switching Edge Router

At the edge of the MPLS network, the traffic entering the MPLS network is divided into different FECs by the LER, and the corresponding labels are requested for these FECs. It provides traffic classification, label mapping, tag removal.

5. Label Switched Path

An FEC data flow is assigned a certain label at the different nodes, and the data is forwarded according to these labels. The path taken by the data flow is the LSP(Label Switched Path).

6. Label Distribution Protocol

LDP (Label Distribution Protocol) is the control protocol of MPLS, which is equivalent to the signaling protocol in the traditional network. It is responsible for FEC classification, label distribution and LSP establishment and maintenance and a series of operations.

MPLS can use a variety of label distribution protocols, including protocols developed for label distribution, such as LDP and CR-LDP (Constraint-Based Routing using LDP), and protocols that support label distribution after existing protocol extensions, such as BGP(Border Gateway Protocol) and RSVP (Resource ReservationProtocol). In addition, you can configure static LSP manually.

7. Multilayer label stack

If the packet is transmitted in LSP tunnel which is more than one layer, there will be the multilayer label to form the label stack. At the entrance and exit of each tunnel, the label is pushed (PUSH) and popped (POP).

Label stack organizes the labels in the "Last-In-First-Out" manner. MPLS handles label from the top of the stack.

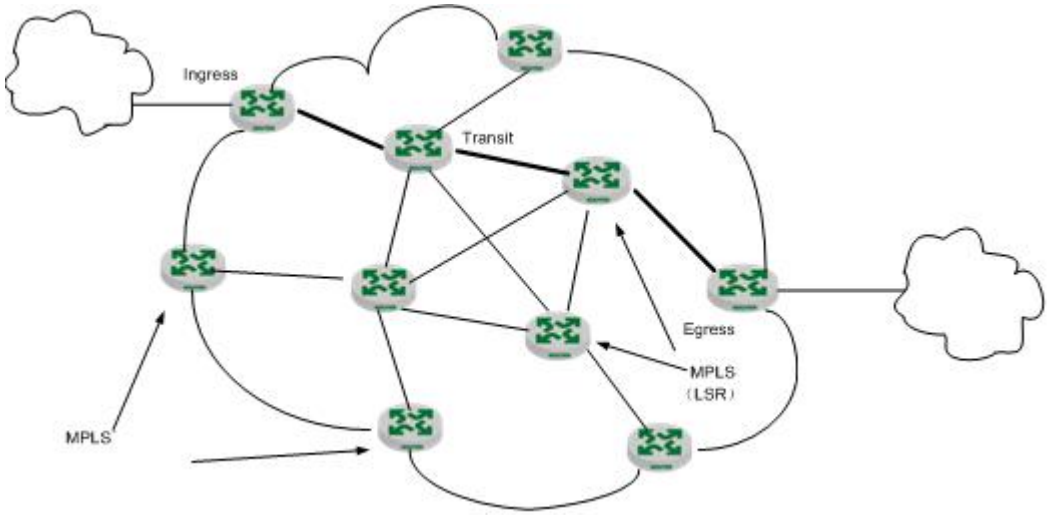
MPLS does not limit the depth of the label stack. If the label stack depth of a packet is m, the label at the bottom of the stack is the 1 level label and the label at the top of the stack is the m level label. Packets that are not pushed into a label can be thought of as empty packets (ie, label stack depth is zero).

48.1.2 MPLS Architecture

As shown below, the basic component of the MPLS network is the LSR, a network consisting of LSRs is called an MPLS domain.

At the edge of the MPLS domain, the LSR connected to other user networks is called the LER (Label Edge Router). The LSR within the area is called the core LSR. MPLS communication is used between the LSRs within the domain. The edge of the MPLS domain is adapted by the LER and legacy IP technologies.

After the ingress LER is pushed into the label, the packet is transported along the LSP consisting of a series of LSRs, where the ingress LER is called the Ingress, the egress LER is the Egress, and the intermediate node is the Transit.



48.1.3 MPLS Basic Working Process

The figure above (MPLS Architecture) provides an overview of the basic MPLS work process:

(1) First, LDP establishes the routing table and LIB (Label Information Base) for the FEC with service requirements in each LSR together with the traditional routing protocol (such as OSPF, ISIS, etc.).

(2) The ingress LER receives the packet, completes the third layer function, determines the FEC to which the packet belongs, and labels the packet to form an MPLS label packet;

(3) Next, in the network formed by the LSR, the LSR forwards the packet based on the label on the packet and the LFIB (Label ForwardingInformation Base), and does not perform any Layer 3 processing on the label packet;

(4) Finally, the labels in the packet are removed by the MPLS egress LER, and the subsequent IP forwarding is continued.

It can be seen, MPLS is not a business or application, it is actually a kind of tunneling technology, but also a routing and switching technology platform that combines label switching and network layer routing. This platform not only supports multiple high-level protocols and services, but also to a certain extent, to ensure the security of information transmission.

48.2 MPLS Basic Configuration Tasks

48.2.1 MPLS Basic Configuration Tasks Introduction

MPLS Basic Configuration Tasks Introduction

Configuration Task		Description	Detailed configuration
Configure the basic MPLS capability		Required	1.2.2
Configure the static LSP		Optional	1.2.3
Configure LDP	Configure the LSR-ID	Required	1.2.4.1
	Enable MPLS LDP	Required	1.2.4.2
	Configure the transport address	Optional	1.2.4.3
	Configure the label distribution control mode	Optional	1.2.4.4
	Configure the label advertisement mode	Optional	1.2.4.5
	Configure the route label re-advertisement function	Optional	1.2.4.6
	Configure the label retention mode	Optional	1.2.4.7

	Configure the penultimate hop pop-up function	Optional	1.2.4.8
	Configure the link Hello hold time	Optional	1.2.4.9
	Configure the link Keepalive hold time	Optional	1.2.4.10
Configure LDP loop detection		Optional	1.2.5
MPLS display and maintenance		Optional	1.2.6

48.2.2 Configure the Basic MPLS Capability

Before using the MPLS-related functions, you must enable MPLS function of the global and the VLAN interface

Configure the Basic MPLS Capability

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enable MPLS	**mpls**	Required. The default is off
Enter interface configuration mode	**interface vlan-interface** *vlan-id*	
Enable MPLS	**mpls**	Required. The default is off
Disable the MPLS function of the system or VLAN interface	**no mpls**	

48.2.3 Configure the Static LSP

You can manually set an LSR as a node on an LSP and can limit the data flow carried on the LSP. Depending on the location in the MPLS domain, the LSR has three nodes: Ingress, Transit, and Egress. Note that the LSRs along the specified LSP must be configured to make the LSP work properly.

Configure the Static LSP

Operation	Command	Remarks
Enter MPLS configuration mode	**mpls**	
Configure the ingress node of the LSP	**static-lsp ingress** *lsp-name* { **destination** *ip-address* { *addr-mask* **|** *mask-length* } } **nexthop** *ip-address* } }	Required

	out-label *out-label-value*	
Configure the intermediate node of the LSP	**static-lsp transit** *lsp-name* **incoming-interface** *interface-type interface-number* **in-label** *in-label-value* **nexthop** *ip-address* **out-label** *out-label-value*	Required
Configure the egress node of the LSP	**static-lsp egress** *lsp-name* **incoming-interface** *interface-type interface-number* **in-label** *in-label-value*	Required
Remove the LSP node configuration	**no static-lsp** *lsp-name*	

48.2.4 LDP Protocol Configuration

48.2.4.1Configure the LSR-ID

In LDP, lsr-id uniquely identifies an LSR, and the value of lsr-id must be globally unique. If you do not manually configure the lsr-id for LDP, the system will automatically select an interface address as the router ID for LDP.

Normally, LDP uses the default MPLS LSR ID. In some networking scenarios where VPN instances are used, if the VPN overlaps the public network address space, you need to configure another LSR ID for the LDP to ensure that the TCP connection can be set up normally. It is recommended to use the switch interface IP as the LSR-ID.

Configure the LSR-ID

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Configure the LSR ID of the local node	**mpls lsr-id** *ip-address*	Required

48.2.4.2Enable MPLS LDP

In global configuration mode, the mpls ldp command enables the LDP capability of the local node and enters the LDP view. Use the no mpls ldp command to disable LDP.

In interface mode, use the mpls ldp command to enable LDP on an interface. Use the no mpls ldp command to remove the ldp capability of an interface.

Enable MPLS LDP

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enable LDP	**mpls ldp**	Required. The default is off
Enter interface configuration mode	**interface vlan-interface** *vlan-id*	
Enable LDP on the interface	**mpls ldp**	Required. The default is off

Note:

This command can be executed only after the MPLS LSR ID is configured and the MPLS capability of the local node is enabled.

48.2.4.3 Configure the Transport Address

The transmission address of LDP is used to establish a TCP connection between peers. In VLAN interface mode, you can specify the IP address of the VLAN interface as the transport address, or specify the IP address of the other VLAN interface as the transport address. By default, the LSR ID of the LSR is the transport address.

Configure the Transport Address

Operation	Command	Remarks
Enter interface configuration mode	**interface vlan-interface** *vlan-id*	-
Specify the IP address of the VLAN interface as the transport address	**mpls ldp transport-address interface**	Optional
Specify the IP address of the other VLAN interface as the transport address	**mpls ldp transport-address vlan-interface** *vlan-id*	Optional
Delete the transport address	**no mpls ldp transport-address**	

Note:

After the interface is enabled with mpls ldp, shut down mpls ldp on the corresponding interface and restart mpls ldp before the new configuration takes effect.

48.2.4.4 Configure the Label Distribution Control Mode

There are two types of label distribution control:

Independent label distribution control mode (independent), An LSR can advertise the label mapping to an LSR connected to it at any time. This approach may result in the upstream label being issued before the downstream label is received.

Ordered label distribution mode (ordered), that is, the LSR sends the label mapping message to the upstream LSR only when the LSR receives a specific label mapping message for the next hop of a specific FEC or if the LSR is the egress node of the LSP of a specific FEC.

Configure the Label Distribution Control Mode

Operation	Command	Remarks
Enter MPLS LDP view mode	**mpls ldp**	-
Configure the label distribution control mode	**label-distribution { independent | orderd }**	Optional. By default, the label distribution control mode is in the ordered mode
Restore the default configuration	**no label-distribution**	

48.2.4.5 Configure the Label Advertisement Mode

There are two types of label advertisement:

DoD （Downstream On Demand）： For a specific FEC, the LSR obtains the label request message from the upstream and then performs label distribution and advertisement.

DU（Downstream Unsolicited）：For a specific FEC, the LSR does not need to obtain the label request message from the upstream，and carries on label distribution and advertisement directly.

Configure the Label Advertisement Mode

Operation	Command	Remarks

Enter interface configuration mode	**interface vlan-interface** *vlan-id*	-
Configure the label advertisement mode	**mpls ldp advertisement { dod | du }**	Optional. By default, the label advertisement mode is DU
Restore the default configuration	**no mpls ldp advertisement**	

Note：

An upstream LSR and a downstream LSR with label advertisement adjacency must use the same label advertisement mode. Otherwise, the LSP can not be established normally.

48.2.4.6 Configure the Route Label Re-advertisement Function

Configure the Route Label Re-advertisement Function

Operation	Command	Remarks
Enter MPLS LDP view mode	**mpls ldp**	
Enable label re-advertisement in DU mode	**du-readvertise**	Optional. Default is not enabled
Set the interval period for periodic re-advertisement of labels in DU mode	**du-readvertise timer** *value*	Optional. The default time is 30s
Disables label re-advertisement in DU mode	**no du-readvertise**	
Restore the default interval period for periodic re-advertisement of labels in DU mode	**no du-readvertise timer**	

48.2.4.7Configure the Label Retention Mode

Label retention mode dictates how to process a label to FEC binding that is received by an LSR but not useful at the moment.

There are two modes of label retention:

Liberal: the label mapping received from the neighbor LSR is retained regardless of whether the neighbor LSR is its own next hop or not.

Conservative: the label mapping received from the neighbor LSR is retained only when the neighbor LSR is its own next hop.

LSR can quickly adapt to route changes by using the liberal label retention mode. With conservative label retention mode, LSR can allocate and store fewer labels.

The conservative label retention mode is usually used with the DoD mode for the LSR with limited label space.

Configure the Label Retention Mode

Operation	Command	Remarks
Enter MPLS-LDP view mode	**mpls ldp**	-
Configure the label retention mode	**label-retention { conservative | liberal }**	Optional. By default, the label retention mode is liberal
Restore the default configuration	**no label-retention**	

48.2.4.8 Configure the Penultimate Hop Pop-up Function

In the MPLS network, the core LSR forwards packets according to the labels on the packets. In the egress node (egress LER), the labels are removed from the packets and IP forwarding is continued. In fact, in a simple MPLS application, the egress node only needs IP forwarding, and the label is no longer used. In this case, to reduce the burden on the egress node and improve the packet processing capability of the MPLS network, you can use the PHP (Penultimate Hop Popping) feature to eject the label at the penultimate node, and the egress node does not perform the label operation.

Configure the Penultimate Hop Pop-up Function

Operation	Command	Remarks
Enter MPLS view mode	**mpls**	-
Configure the second pop-up function	**label advertise { exlicit-null | implicit-null | non-null }**	Optional. By default, PHP is supported, the egress

		node assigns an implicit null label to the penultimate hop.
Restore the default configuration	**no label advertise**	

Note：

explicit-null：Support PHP(Penultimate Hop Popping).The egress node assigns an explicit null label to the penultimate hop, the value is 0.

implicit-null ：Support PHP. The egress node assigns an implicit null label to the penultimate hop, the value is 3.

non-null： PHP is not supported. The egress node assigns a label to the penultimate hop.

48.2.4.9 Configure the Link Hello Hold Time

The LDP neighbor discovery mechanism discovers LDP neighbors by sending hello packets (UDP/prot:646/IP:224.0.0.2) to each other. The LDP-enabled interface periodically sends the message. The hello message sent by the LSR carries the hold time of the locally configured LDP link hello adjacency, and the LSR of the received hello message compares it with the hold time of the locally configured LDP link hello adjacency. The small value is used as the hold time of the adjacency of the LDP link hello. If no link hello message from the adjacency is received within the hold time of the adjacency of the LDP link hello, the adjacency is considered not to exist, and the corresponding session will end and the information of this neighbor will be deleted.

Configure the Link Hello Hold Time

Operation	Command	Remarks
Enter interface configuration mode	**interface vlan-interface** *vlan-id*	-
Configure the link Hello hold time	**mpls ldp timer hello-hold** *value*	Optional. By default, the link Hello hold timer is 15 seconds

48.2.5 Ldp Loopback Detection Configuration

Establishing an LSP in the MPLS domain also needs to prevent loops. The LDP loop detection mechanism detects the presence of LSP loops and avoids loops.

If you perform loop detection on the MPLS domain, you must configure loop detection on all LSRs. However, when establishing LDP sessions, the loop detection configurations on both sides are not required to be consistent.

LDP loop detection has two modes:

1. Maximum number of hops

The hops information is included in the message that passes the label binding (or label request), and the value is incremented by 1 after each hop. When the value reaches the specified maximum value, the loop is considered to be established and the LSP establishment fails.

2. Path Vector

The path information is recorded in a message conveying the label binding (or label request). After each hop, the corresponding device checks whether its LSR ID is in this record.

The LSP is considered to have failed because of a loop in one of the following conditions:

(1) The record of this LSR already exists in the path vector record table;

(2) The hop count of the path reaches the set maximum value.

If the record does not have its own LSR ID, it is added to the record.

Configure the ldp loop detection function

Operation	Command	Remarks
Enter MPLS-LDP view	**mpls ldp**	-
Enable loop detection	**loop-detect**	Required. By default, loop detection is disabled.
Set the maximum hop count for loop detection	**hops-count** *hop-number*	Optional. By default, the maximum hop count for loop detection is 32
Set the maximum value of the path vector	**path-vectors** *pv-number*	Optional. By default, the path vector value is 32

Restore the loopback detection maximum hop count to the default	**no hops-count**	
Restore the maximum value of the path vector to the default	**no path-vectors**	
Disable the loop detection function	**no loop-detect**	

48.2.6 MPLS Display and Maintenance

MPLS Information Display and Debugging Commands

Operation	Command	Remarks
Display information about static LSP	**show mpls static-lsp [lsp-name** *lsp-name*]	Optional
Display information about LDP adjacency	**show mpls ldp neighbor**	Optional
Display the MPLS label usage status	**show mpls label** { **bgp** |**ldp**| **static**| **all** }	Optional
Display information about ILM entries	**show mpls ilm** {**label** |**all**}	Optional
View the NHLFE table information	**show mpls nhlfe** {**token** |**all**}	Optional
Display LSP information created by LDP	**show mpls ldp forwarding-table**	Optional
Debug LDP session information	**debug mpls ldp**	Optional
Packets sent and received by LDP	**ldpdebug{all| encode| enter| trace }**	Optional

48.3 MPLS Basic Configuration Example

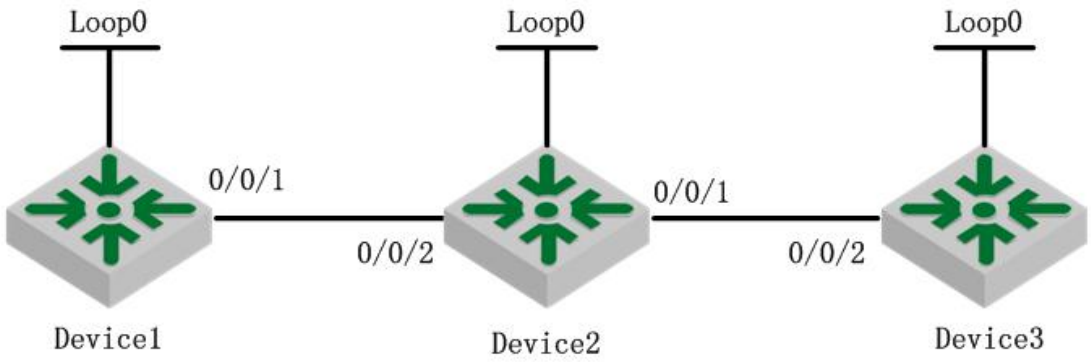
48.3.1 Networking Requirements

（1） Switch1, Switch2, and Switch3 both support MPLS and run OSPF as the IGP on the MPLS backbone;

（2）A local LDP session is established between Switch 1 and Switch 2, and between Switch 2 and Switch 3;

（3）Establish an LSP link between Switch1, Switch2, and Switch3.

The figure below shows the networking diagram



MPLS LDP Configuration

48.3.2 Configuration Steps

1） Configure the VLAN interface IP and subnet mask and loopback interface of each switch according to the figure above (MPLS LDP Configuration). Take SW1 as an example, the configuration of other devices is abbreviated.

```
DEVICE1(config)#interface loopback-interface 0
DEVICE1(config-if-loopBackInterface-0)#ip address 1.1.1.1 255.255.255.255
DEVICE1(config-if-loopBackInterface-0)#exit
DEVICE1(config)#interface vlan-interface 1
DEVICE1(config-if-vlanInterface-1)#ip address 192.168.1.1 255.255.255.0
```

2） Configure the OSPF protocol

Configure DEVICE1

```
DEVICE1(config)#router ospf
DEVICE1(config-router-ospf)#
DEVICE1(config-router-ospf)#network 1.1.1.1 ?
DEVICE1(config-router-ospf)#network 1.1.1.1 0.0.0.0 area 0
DEVICE1(config-router-ospf)#network 192.168.1.1 0.0.0.255 area 0
```

Configure DEVICE2

```
DEVICE2(config)#router ospf
DEVICE2(config-router-ospf)#network 2.2.2.2 255.255.255.255 area 0
DEVICE2(config-router-ospf)#network 192.168.1.2 0.0.0.255 area 0
DEVICE2(config-router-ospf)#network 192.168.2.1 0.0.0.255 area 0
```

Configure DEVICE3

DEVICE3(config)#router ospf
DEVICE3(config-router-ospf)#network 3.3.3.3 0.0.0.0 area 0
DEVICE3(config-router-ospf)#network 192.168.2.2 0.0.0.255 area 0
DEVICE3(config-router-ospf)#

After the configuration, Switch 1 and Switch 2, Switch 2 and Switch 3 should establish an OSPF neighbor relationship with the status of full. Run the show ip route command on each device, and you can see each other have learned to each other's host routing. Take DEVICE1 as an example:

DEVICE1(config)#show ip ospf neighbor
show ip ospf neighbor information

IPAddress	NeighborID	State	Priority	Event	Type
192.168.1.2	2.2.2.2	Full	1	6	dynamic

Total entries: 1 IP ospf neighbor.
DEVICE1(config)#show ip route
Show ip route information

INET route table - vr: 0, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
1.1.1.1	1.1.1.1	UH	0	LOOPBACK-IF0	0	0
2.2.2.2	192.168.1.2	UGH	0	VLAN-IF1	20	0
3.3.3.3	192.168.1.2	UGH	0	VLAN-IF1	30	0
127.0.0.0/8	127.0.0.1	UR	0	lo0	0	0
127.0.0.1	127.0.0.1	UH	4	lo0	0	0
192.168.1.0/24	192.168.1.1	UC	1	VLAN-IF1	0	0
192.168.1.1	192.168.1.1	UH	0	lo0	0	0
192.168.2.0/24	192.168.1.2	UG	0	VLAN-IF1	20	0

Total entries: 8. Printed entries: 8.

3）Configure MPLS basic capability and enable LDP

Configure DEVICE1

DEVICE1(config)#mpls lsr-id 1.1.1.1
DEVICE1(config)#mpls
DEVICE1(config-router-mpls)#exit
DEVICE1(config)#mpls ldp
DEVICE1(config-router-ldp)#interface vlan-interface 1
DEVICE1(config-if-vlanInterface-1)#mpls
DEVICE1(config-if-vlanInterface-1)#mpls ldp
DEVICE1(config-if-vlanInterface-1)#exit
DEVICE1(config)#

Configure DEVICE2

```
DEVICE2(config)#mpls lsr-id 2.2.2.2
DEVICE2(config)#mpls
DEVICE2(config-router-mpls)#exit
DEVICE2(config)#mpls ldp
DEVICE2(config-router-ldp)#interface vlan-interface 1
DEVICE2(config-if-vlanInterface-1)#mpls

DEVICE2(config-if-vlanInterface-1)#mpls ldp
DEVICE2(config-if-vlanInterface-1)#exit
DEVICE2(config)#interface vlan-interface 2
DEVICE2(config-if-vlanInterface-2)#mpls
DEVICE2(config-if-vlanInterface-2)#mpls ldp
DEVICE2(config-if-vlanInterface-2)#exit
```

Configure DEVICE3

```
DEVICE3(config)#mpls lsr-id 3.3.3.3
DEVICE3(config)#mpls
DEVICE3(config-router-mpls)#exit
DEVICE3(config)#mpls ldp
DEVICE3(config-router-ldp)#exit
DEVICE3(config)#interface vlan-interface 2
DEVICE3(config-if-vlanInterface-2)#mpls
DEVICE3(config-if-vlanInterface-2)#mpls ldp
DEVICE3(config-if-vlanInterface-2)#exit
```

After the above configuration, Switch 1 and Switch 2, Switch 2 and Switch 3 establish the local LDP session. Run the display mpls ldp session command on each device, and you can view the establishment of LDP sessions. Run the show mpls ldp neighbor command, and you can view LDP peers. Take DEVICE1 as an example:

```
DEVICE1(config)#show mpls ldp neighbor
mpls ldp neighbor information:
PeerId          TranportAddress    Vlan-If(sw)
2.2.2.2:0       2.2.2.2            sw0

Total entries: 1 mpls ldp neighbor.

DEVICE1(config)#show mpls ldp session
mpls ldp session information:
PeerId          Status          Role        HoldTime    LabelAdvMode    LoopDetection/PVLim
```

2.2.2.2:0 OPERATIONAL PASSIVE 45 DU DISABLE/0

Total entries: 1 mpls ldp session.

48.3.3 Results Verification

A local LDP session is established between Switch 1 and Switch 2, and between Switch 2 and Switch 3, and an LSP link is established between Switch 1, Switch 2 and Switch 3.

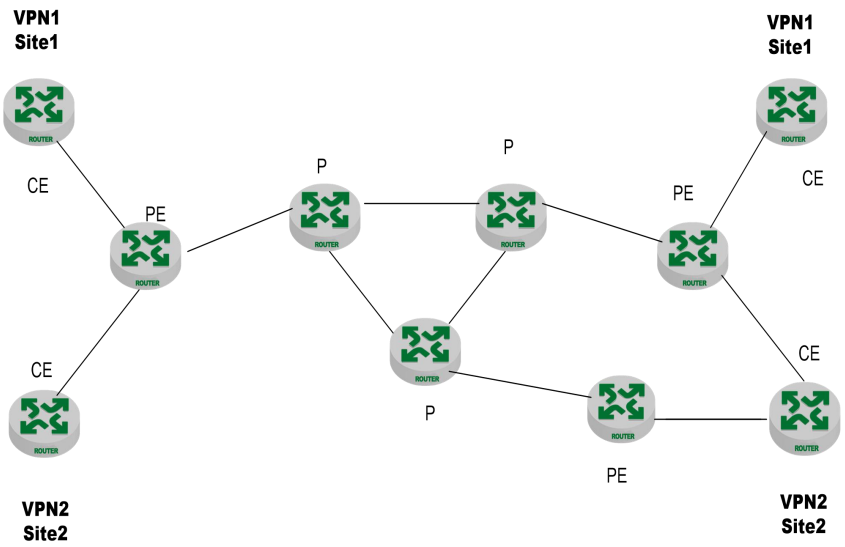
48.4 MPLS L3VPN Introduction

48.4.1 MPLS L3VPN Overview

MPLS L3VPN is a PE-based L3VPN technology in service provider VPN solutions. It uses BGP to advertise VPN routes on the backbone network of the service provider and MPLS to forward VPN packets on the service provider backbone network.

MPLS L3VPN networking is flexible, scalable, and can easily support MPLS QoS and MPLS TE, so it is increasingly used.

48.4.2 MPLS L3VPN Basic Structure



The figure above shows an MPLS L3VPN networking solution, consisting of three parts: CE, PE, and P.

CE （Customer Edge） device: Customer network edge device. An interface is directly connected to the SP (Service Provider). The CE can be a router or a switch, or a host. The CE is not aware of the presence of VPN and does not need to support MPLS.

PE（Provider Edge）router：Service provider edge router is the edge device of the service provider network and is directly connected to the CE of the user. In the MPLS network, all the processing of the VPN takes place on the PE.

P（Provider）router：The backbone router in the service provider network is not directly connected to the CE. The P device only needs to have basic MPLS forwarding capability.

48.4.3 MPLS L3VPN Basic Concept

1. Site

“Site” is often mentioned in the introduction of VPN. The meaning of Site can be understood from the following aspects:

A site is a group of IP systems that have IP connectivity with each other, and the IP connectivity of this group of IP systems does not need to be implemented through the service provider network;

Site partitioning is based on the topological relationship of the device, not the geographical location, although in most cases, the geographical locations of devices in a site are adjacent.

The device in a site can belong to multiple VPN, in other words, a site can belong to multiple VPN;

The Site connects to the service provider network through the CE. A Site can contain multiple CEs, but a CE belongs to only one Site.

For a plurality of Sites connected to the same service provider network, they can be divided into different sets by making policies, and only Sites belonging to the same set can communicate with each other through the service provider network. This set is the VPN.

2. Overlapping Address Spaces

VPN is a private network, different VPN independently manage their own use of the address range, also known as AddressSpace.

Different VPN address space may overlap in a certain range. For example, VPN1 and VPN2 are using the address 10.110.10.0/24 network segment, this causes Overlapping Address Spaces

3.VRF

VPN routing and forwarding instance, referred to as VRF.

In MPLS VPN, a PE router, since it is possible to connect multiple VPN users at the same time, these users (routes) need to be isolated from each other, then this time to use the VRF.

A PE establishes and maintains a VPN instance for each directly connected site. The VPN instance contains VPN membership and routing rules for the corresponding sites. If users in a site belong to multiple VPNs at the same time, the information about all these VPNs will be included in the VPN instance of the site.

To ensure the independence and security of VPN data, each VPN instance on the PE has a relatively independent routing table and LFIB (Label Forwarding Information Base).

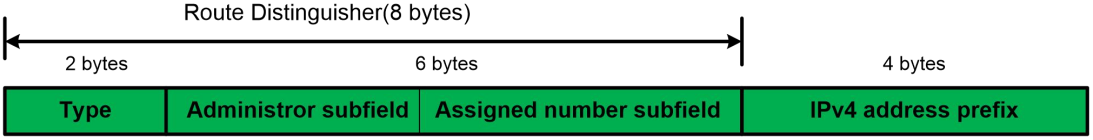
Specifically, the information in the VPN instance includes: label forwarding table, IP routing table, interface bound to the VPN instance, and management information of the VPN instance. The management information of the VPN instance includes the RD (Route Distinguisher), the routing filter policy, the member interface list, and so on.

4. VPN-IPv4 Address

Traditional BGP can not correctly handle VPN routes with overlapped address spaces. Assume that both VPN1 and VPN2 use the IP addresses of 10.110.10.0/24, and each advertises a route to the network segment, BGP will select only one route, resulting in the loss of route to another VPN.

PE routers use MP-BGP to advertise VPN routes and use the VPN-IPv4 address family to solve the above-mentioned problems.

The VPN-IPv4 address consists of 12 bytes, including an 8-byte RD and a 4-byte IPv4 address prefix, as shown below.



After receiving a common IPv4 route from the CE, the PE advertises these private VPN routes to the opposite PE. The independence of private network routes is achieved by attaching RDs to these routes.

SP can allocate RD independently, but must ensure the global uniqueness of RD. In this way, even if VPNs from different service providers use the same IPv4 address space, the PE router can advertise different routes to each VPN.

It is recommended to configure a dedicated RD for each VPN instance on the PE to ensure that the routes to the same CE use the same RD. A VPN-IPv4 address with an RD of 0 corresponds to a globally unique IPv4 address.

The role of RD is to add to a specific IPv4 prefix, making it a globally unique VPN IPv4 prefix.

RD is either associated with an autonomous system number (ASN). In this case, the RD is composed of an autonomous system number and an arbitrary number. Or it is associated with an IP address, in which case the RD is composed of an IP address and an arbitrary number.

RD has two formats, which are distinguished by a 2-byte Type field:

When Type is 0, the Administrator subfield occupies 2 bytes and the Assigned number subfield occupies 4 bytes. The format is 6 bits AS number: 32 bits User-defined number. For example: 100: 1

When Type is 1, the Administrator subfield occupies 4 bytes and the Assigned number subfield occupies 2 bytes. The format is as follows: 32 bits IPv4 address: 16 bits User-defined number. For example: 172.1.1.1: 1

To ensure the global uniqueness of RD, it is not recommended to set the value of the Administrator subfield to a private AS number or private IP address.

5. VPN Target Attribute

MPLS L3VPN uses the BGP extended community attribute--VPN target (also known as Route Target) to control the advertisement of VPN routing information.

A VPN instance on a PE router has two types of VPN target attributes:

Export Target attribute: Before the local PE advertises VPN-IPv4 routes learned from its own directly connected sites to other PEs, set the export target attribute for these routes.

Import Target attribute: The PE checks its export target attribute when it receives VPN-IPv4 routes advertised by other PEs. Only when this attribute matches the Import Target attribute of the VPN instance on the PE, the route is added to the corresponding VPN routing table.

That is, the VPN target attribute defines which sites a VPN-IPv4 route can be received by and which sites the PE router can receive routes from.

Similar to RD, VPN Target has two formats:

16bits autonomous system number: 32bits user-defined number. For example: 100: 1.

32bits IPv4 address: 16bits user-defined number. For example: 172.1.1.1: 1.

6. MP-BGP

MP-BGP (Multiprotocol extensions for BGP-4) propagates VPN composition information and routes between PE routers.

MP-BGP is backward compatible. It supports both traditional IPv4 address family and other address families (such as VPN-IPv4 address family). Using MP-BGP ensures that VPN private network routes are advertised only within the VPN and communication between MPLS VPN members is implemented.

48.4.4 Packet Forwarding of MPLS L3VPN

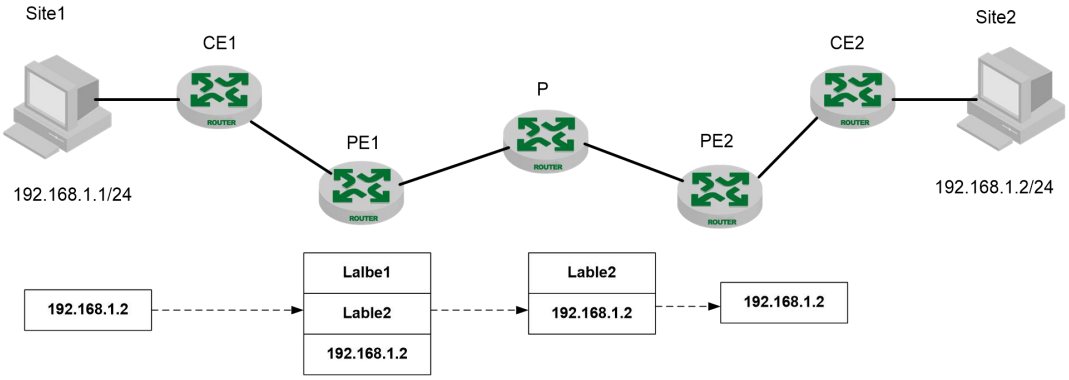
In basic MPLS L3VPN applications, VPN packets are forwarded with two layers of labels:

The first layer (Outer layer) labels are exchanged inside the backbone network and indicate an LSP from the PE to the opposite PE. VPN packets use this label to reach the opposite PE along the LSP.

The second layer (inner layer) label is used when reaching the CE from the opposite PE to indicate to which site the packet should be sent or, more specifically, to which CE. In this way, the opposite PE can find the interface to forward packets according to the inner label.

In the special case, two sites that belong to the same VPN are connected to the same PE. In this case, you only need to know how to reach the opposite CE.

The figure below shows the forwarding of VPN packets:



MPLS L3VPN Packet Forwarding Diagram

(1) Site 1 sends an IP packet with destination address 192.168.1.2 and CE 1 sends the packet to PE 1.

(2) PE 1 searches the routing table of the VPN instance based on the interface and destination address where the packet arrives. After matching, the packets are forwarded and labeled with inner and outer labels.

(3) The MPLS network uses the outer label of the packet to transmit the packet to PE 2. （If PHP is enabled on the P router, the packet is stripped of the outer label when it reaches the previous hop of PE 2, and only contains the inner label.）

(4) PE 2 searches the VPN instance routing table based on the inner label and the destination address to determine the outgoing interface of the packet and forwards the packet to the CE.

(5) CE 2 transmits the packet to the destination according to the normal IP forwarding process.

48.5 MPLS L3VPN Configuration Tasks

48.5.1 MPLS L3VPN Configuration Tasks Introduction

MPLS L3VPN Configuration Tasks Introduction

Configure the task		Explanation	Detailed configuration
Configure a VPN instance	Create a VPN instance	Required	2.2.2.1
	Associate a VPN instance with an interface	Required	2.2.2.2
	Configure the route-related attributes of the VPN instance	Required	2.2.2.3

Configure PE-CE and PE-PE routes	Configure route switching between PE and CE	Required	2.2.3.1
	Configure Route Switching between PEs	Required	2.2.3.2
MPLS L3VPN display and maintenance		Optional	2.2.4

48.5.2 Configure a VPN Instance

The VPN instance is used to isolate VPN private network routes from public network routes. In all MPLS L3VPN networking scenarios, you need to configure a VPN instance.

VPN instance can isolate VPN private network route from public network route. The routes of different VPN instances are also isolated from each other. This feature makes the use of VPN instances not limited to MPLS L3VPN.

The configuration of a VPN instance is performed on the PE.

48.5.2.1 Create a VPN instance

The VPN instance is associated with the Site in the implementation. A VPN instance does not directly correspond to a VPN. A VPN instance integrates the VPN membership and routing rules of its corresponding site.

A VPN instance takes effect only after RD is configured.

Create a VPN instance

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Create a VPN instance	**ip vpn-instance** *vpn-instance-name*	Required
Configure the RD of the VPN instance	**ip vpn-instance** *vpn-instance-name* **route-distinguisher** *route-distinguisher*	Required

48.5.2.2 Associate a VPN Instance with an Interface

After the VPN instance is configured, you need to associate it with the VLAN interface of the CE.

Associate a VPN Instance with an Interface

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the VLAN interface to be associated	**interface vlan-interface** *vlan-id*	
Configure the RD of the VPN instance	**ip binding vpn-instance** *vpn-instance-name*	Required.By default, an interface is not associated with any VPN instance

Note：

If the interface to be bound is configured with an IP address, you need to delete the interface IP address before executing the ip binding vpn-instance command.

48.5.2.3 Configure the Route-related Attributes of the VPN Instance

VPN route release control process is as follows:

When a VPN route learned from CE is imported into BGP, BGP associates it with a list of VPN target extended community attributes. Usually, this list is the output routing attribute list of the VPN instance associated with the CE.

The VPN instance determines the routes that can be accepted and imported into the VPN instance according to the import-extcommunity in the VPN target.

The VPN instance performs the modification of the VPN target attribute to the advertised routes according to the export-extcommunity in the VPN target.

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Associate the current VPN instance with one or more VPN targets	**ip vpn-instance** *vpn-instance-name* **vpn-target** *vpn-target* &<1-31> [**both** |**export-extcommunity** |**import-extcommunity**]	Required. By default, an interface is not associated with any VPN instance

48.5.3 Configure PE-CE and PE-PE Routes

48.5.3.1 Configure Route Switching between PE and CE

Static routes, RIP, OSPF, EBGP, and other routing protocols can be used to configure route switching between PE and CE. Which protocol to use depends on the actual needs.

1. Configure a static route between PE and CE

Configure a static route between PE and CE

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Configure a static route for the specified VPN instance	**ip route** *dip mask nexthop* [**vpn-instance** *vpn-instance-name*]	Optional. This configuration takes effect on PE. The configuration of CE is the same as that of ordinary static routes

Note：

2. Configure RIP between PE and CE

A RIP process can belong to only one VPN instance. If you do not bind to a VPN instance when the RIP process is started, the process belongs to the public network process.

Configure RIP between PE and CE

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter the RIP process configuration mode	**router rip**	
Enable the rip protocol on the VPN interface and advertise the routes	**network** *ip* [**vpn-instance** *vpn-instance-name*]	Optional. The configuration takes effect on the PE and the common RIP is configured on the CE.

3. Configure EBGP between PE and CE

Configure EBGP on the PE

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter BGP view	**router bgp** *AS- number*	Optional
Enter BGP-VPN instance view	**address-family ipv4 vpn-instance** *vpn-instance-name*	Optional
Configure the CE as a	**neighbor** *ipaddress* **remote-as** *AS- number*	Optional

VPN private network peer		
Configure the tcp connection interface	**neighbor** *ipaddress* **update-source loopback-interface** *LoopBack-interface-number*	Optional

Configure BGP on the CE

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter BGP view	**router bgp** *AS- number*	Optional
Declare the route advertised by BGP	**network** *ipaddress*	Optional
Configure the PE as a VPN private network peer	**neighbor** *ipaddress* **remote-as** *AS- number*	Optional

48.5.3.2 Configure Route Switching between PEs

Configure Route Switching between PEs

Operation	Command	Remarks
Enter the global configuration mode	**configure terminal**	-
Enter BGP view	**router bgp** *AS- number*	Optional
Configure the opposite PE as peer	**neighbor** *ipaddress* **remote-as** *AS- number*	Optional
Specify the source interface for route update packets	**neighbor** *ipaddress* **update-source loopback-interface** *LoopBack-interface-number*	Optional
Enter BGP-VPNv4	**address-family vpnv4**	Optional

sub-address family view		
Activate BGP neighbors	**neighbor** *ipaddress* **activate**	Peers exchange BGP-VPNv4 routing information

48.5.4 MPLS L3VPN Display and Maintenance

MPLS L3VPN Display and Maintenance

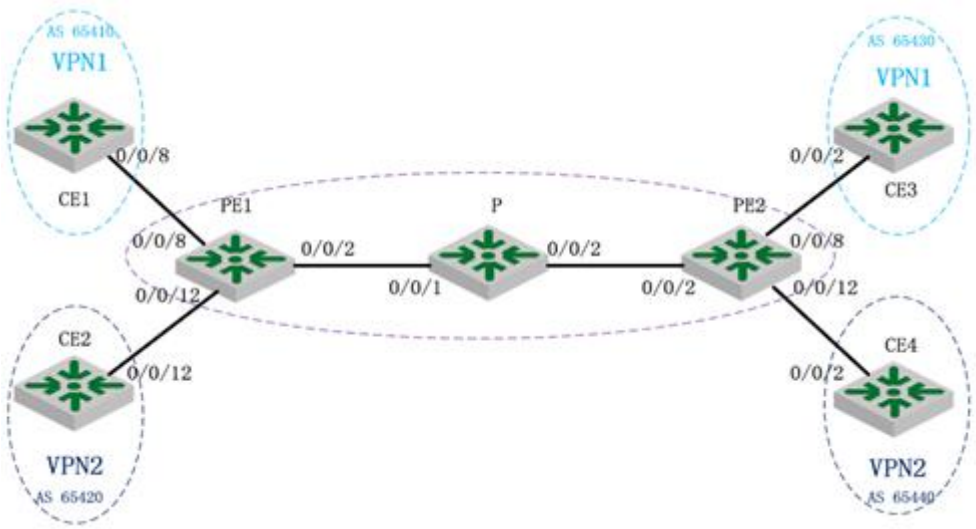
Operation	Command	Remarks
View the current vpn instance	**show ip vpn-instance** [*vpn-instance-name*]	Optional. If no instance name is specified, all vpn instances are displayed
View the routing table of vpn1	**Show ip route vpn-instance** *vpn-instance-name*	Optional
Check the validity and reachability of IP in the destination VPN	**ping** [*options*] *dip* [**vpn-instance** *vpn-instance-name*]	Optional
Trace the route to the destination IP	**tracert** [*options*] *dip* [**vpn-instance** *vpn-instance-name*]	Optional. Peers exchange BGP-VPNv4 routing information

48.6 MPLS L3VPN Configuration Example

48.6.1 Requirement and Networking

CE1 and CE3 belong to VPN1, CE2 and CE4 belong to VPN2. VPN1 uses the VPN target attribute value of 111: 1, and VPN2 uses the VPN target attribute value of 222: 2. CE and PE use EBGP to exchange VPN routes. Use OSPF to implement internal interworking between PEs and

use MP-IBGP to exchange VPN routing information. The figure below shows the networking diagram.



MPLS L3VPN Configuration Example

Device	Port	Vlan-interface	IP
PE1	0/0/2	3	172.168.1.1
	0/0/8	2	192.168.2.1
	0/0/12	10	192.168.3.1
	Loop	0	1.1.1.1
p	0/0/1	3	172.168.1.2
	0/0/2	4	172.168.2.1
	Loop	0	2.2.2.2
PE2	0/0/2	4	172.168.2.2
	0/0/8	13	192.168.4.1
	0/0/12	14	192.168.5.1
	Loop	0	3.3.3.3
CE1	0/0/8	2	192.168.2.2
CE2	0/0/12	10	192.168.3.2
CE3	0/0/2	13	192.168.4.2
CE4	0/0/2	14	192.168.5.2

List of Switch Interface Configurations

48.6.2 Configuration steps

1) Configure the IP and subnet mask of each interface, including the loopback interface, according to the table above (List of Switch Interface Configurations).
2) Configure the IGP protocol (OSPF) on the MPLS backbone network.

#PE1 configuration

```
PE1(config)#interface vlan-interface 3
PE1(config-if-vlanInterface-3)#ip address 172.168.1.1 255.255.255.0
PE1(config-if-vlanInterface-3)#interface vlan-interface 2
PE1(config-if-vlanInterface-2)#ip address 192.168.2.1 255.255.255.0
PE1(config-if-vlanInterface-2)#interface vlan-interface 10
PE1(config-if-vlanInterface-10)#ip address 192.168.3.1 255.255.255.0
PE1(config-if-vlanInterface-10)#interface loopback-interface 0
PE1(config-if-loopBackInterface-0)#ip address 1.1.1.1 255.255.255.255
PE1(config)#router ospf
PE1(config-router-ospf)#network 1.1.1.1 0.0.0.0 area 0
PE1(config-router-ospf)#network 172.168.1.1 0.0.0.255 area 0
PE1(config-router-ospf)#exit
```

#P configuration

```
P(config)#interface vlan-interface 3
P(config-if-vlanInterface-3)#ip address 172.168.1.2 255.255.255.0
P(config-if-vlanInterface-3)#exit
P(config)#interface vlan-interface 4
P(config-if-vlanInterface-4)#ip address 172.168.2.1 255.255.255.0
P(config-if-vlanInterface-4)#exit
P(config)#interface loopback-interface 0
P(config-if-loopBackInterface-0)#ip address 2.2.2.2 255.255.255.255
P(config-if-loopBackInterface-0)#exit
P(config)#router ospf
P(config-router-ospf)#network 2.2.2.2 0.0.0.0 area 0
P(config-router-ospf)#network 172.168.1.2 0.0.0.255 area 0
P(config-router-ospf)#network 172.168.2.1 0.0.0.255 area 0
P(config-router-ospf)#exit
```

#PE2 configuration

```
PE2(config)#interface vlan-interface 4
PE2(config-if-vlanInterface-4)#ip address 172.168.2.2 255.255.255.0
PE2(config-if-vlanInterface-4)#exit
PE2(config)#interface vlan-interface 13
PE2(config-if-vlanInterface-13)#ip address 192.168.4.1 255.255.255.0
PE2(config-if-vlanInterface-13)#exit
```

PE2(config)#interface vlan-interface 14
PE2(config-if-vlanInterface-13)#ip address 192.168.5.1 255.255.255.0
PE2(config-if-vlanInterface-14)#exit
PE2(config)#interface loopback-interface 0
PE2(config-if-loopBackInterface-0)#ip address 3.3.3.3 255.255.255.255
PE2(config-if-loopBackInterface-0)#exit
PE2(config)#router ospf
PE2(config-router-ospf)#network 3.3.3.3 0.0.0.0 area 0
PE2(config-router-ospf)#network 172.168.2.2 0.0.0.255 area 0
PE2(config-router-ospf)#exit

3) Configure MPLS and MPLS LDP on the MPLS backbone network to establish LDP LSP
#PE1 configuration
PE1(config)#mpls
PE1(config-router-mpls)#label advertise non-null
PE1(config-router-mpls)#exit
PE1(config)#mpls lsr-id 1.1.1.1
PE1(config)#interface vlan-interface 3
PE1(config-if-vlanInterface-3)#mpls
PE1(config-if-vlanInterface-3)#exit
PE1(config)#mpls ldp
PE1(config-router-ldp)#exit
PE1(config)#interface vlan-interface 3
PE1(config-if-vlanInterface-3)#mpls ldp
PE1(config-if-vlanInterface-3)#exi
PE1(config)#

#P configuration
P(config)#mpls
P(config-router-mpls)#label advertise non-null
P(config-router-mpls)#exit
P(config)#mpls lsr-id 2.2.2.2
P(config)#interface vlan-interface 3
P(config-if-vlanInterface-3)#mpls
P(config-if-vlanInterface-3)#exit
P(config)#interface vlan-interface 4
P(config-if-vlanInterface-4)#mpls
P(config-if-vlanInterface-4)#exit
P(config)#mpls ldp
P(config-router-ldp)#exit
P(config)#interface vlan-interface 3
P(config-if-vlanInterface-3)#mpls ldp
P(config-if-vlanInterface-3)#exit
P(config)#interface vlan-interface 4

P(config-if-vlanInterface-4)#mpls ldp
P(config-if-vlanInterface-4)#exit

#PE2 configuration
PE2(config)#mpls
PE2(config-router-mpls)#label advertise non-null
PE2(config-router-mpls)#exit
PE2(config)#mpls lsr-id 3.3.3.3
PE2(config)#interface vlan-interface 4
PE2(config-if-vlanInterface-4)#mpls
PE2(config-if-vlanInterface-4)#exit
PE2(config)#mpls ldp
PE2(config-router-ldp)#exit
PE2(config)#interface vlan-interface 4
PE2(config-if-vlanInterface-4)#mpls ldp
PE2(config-if-vlanInterface-4)#exit

After the above configuration, PE1-P-PE2 should establish the LDP session and LSP link correctly. Use the show mpls ldp session command to view the LDP session between PE and P. Use the show mpls ldp forwarding-table command to view the LDP forwarding entries

PE1(config)#show mpls ldp session
mpls ldp session information:

PeerId	Status	Role	HoldTime	LabelAdvMode	LoopDetection/PVLim
2.2.2.2:0	OPERATIONAL	PASSIVE	45	DU	DISABLE/0

Total entries: 1 mpls ldp session.

PE1(config)#show mpls ldp forwarding-table
mpls ldp lsp information:

DestIp/PrefixLen	In/OutLabel	NextHop	Interface(sw)
2.2.2.2/32	NULL/10000	172.168.1.2	Device1
172.168.2.0/24	NULL/10016	172.168.1.2	Device1
1.1.1.1/32	10000/NULL	1.1.1.1	lo0
172.168.1.0/24	10008/NULL	172.168.1.1	lo0
192.168.2.0/24	10016/NULL	192.168.2.1	lo0
192.168.3.0/24	10024/NULL	192.168.3.1	lo0
3.3.3.3/32	NULL/10024	172.168.1.2	Device1

Total entries: 7 mpls ldp lsp information.

P(config)#show mpls ldp neighbor
mpls ldp neighbor information:
PeerId TranportAddress Vlan-If(sw)
1.1.1.1:0 1.1.1.1 sw0
3.3.3.3:0 3.3.3.3 Device1

Total entries: 2 mpls ldp neighbor.

P(config)#show mpls ldp forwarding-table
mpls ldp lsp information:
DestIp/PrefixLen In/OutLabel NextHop Interface(sw)
2.2.2.2/32 10000/NULL 2.2.2.2 lo0
172.168.1.0/24 10008/NULL 172.168.1.2 lo0
172.168.2.0/24 10016/NULL 172.168.2.1 lo0
3.3.3.3/32 10024/10000 172.168.2.2 Device2
1.1.1.1/32 10032/10000 172.168.1.1 Device1

Total entries: 5 mpls ldp lsp information.

PE2(config)#show mpls ldp session
mpls ldp session information:
PeerId Status Role HoldTime LabelAdvMode LoopDetection/PVLim
2.2.2.2:0 OPERATIONAL ACTIVE 45 DU DISABLE/0

Total entries: 1 mpls ldp session.

PE2(config)#show mpls ldp forwarding-table
mpls ldp lsp information:
DestIp/PrefixLen In/OutLabel NextHop Interface(sw)
3.3.3.3/32 10000/NULL 3.3.3.3 lo0
172.168.2.0/24 10008/NULL 172.168.2.2 lo0
192.168.4.0/24 10016/NULL 192.168.4.1 lo0
192.168.5.0/24 10024/NULL 192.168.5.1 lo0
1.1.1.1/32 NULL/10032 172.168.2.1 Device1
2.2.2.2/32 NULL/10000 172.168.2.1 Device1
172.168.1.0/24 NULL/10008 172.168.2.1 Device1

Total entries: 7 mpls ldp lsp information.

4) Configure VPN instance on the PE and connect the CE to the PE

#PE1 configuration
PE1(config)#ip vpn-instance 1
PE1(config)#ip vpn-instance 1 route-distinguisher 100:1

PE1(config)#ip vpn-instance 1 vpn-target 111:1 both
PE1(config)#ip vpn-instance 2
PE1(config)#ip vpn-instance 2 route-distinguisher 100:2
PE1(config)#ip vpn-instance 2 vpn-target 222:2 both
PE1(config)#interface vlan-interface 2
PE1(config-if-vlanInterface-2)#no ip address
PE1(config-if-vlanInterface-2)#ip binding vpn-instance 1
PE1(config-if-vlanInterface-2)#ip address 192.168.2.1 255.255.255.0
PE1(config-if-vlanInterface-2)#exit
PE1(config)#interface vlan-interface 10
PE1(config-if-vlanInterface-10)#no ip address
Router id has changed into "1.1.1.1". If needed, reboot ospf to make new router id available.
Delete ipaddress successfully!
PE1(config-if-vlanInterface-10)#ip binding vpn-instance 2
PE1(config-if-vlanInterface-10)#ip address 192.168.3.1 255.255.255.0

#PE2 configuration
PE2(config)#ip vpn-instance 1
PE2(config)#ip vpn-instance 1 route-distinguisher 100:1
PE2(config)#ip vpn-instance 1 vpn-target 111:1 both
PE2(config)#ip vpn-instance 2
PE2(config)#ip vpn-instance 2 route-distinguisher 100:2
PE2(config)#ip vpn-instance 2 vpn-target 222:2 both
PE2(config)#
PE2(config)#interface vlan-interface 13
PE2(config-if-vlanInterface-13)#no ip address
PE2(config-if-vlanInterface-13)#ip binding vpn-instance 1
PE2(config-if-vlanInterface-13)#ip address 192.168.4.1 255.255.255.0
PE2(config-if-vlanInterface-13)#exit
PE2(config)#interface vlan-interface 14
PE2(config-if-vlanInterface-14)#no ip address
Router id has changed into "3.3.3.3". If needed, reboot ospf to make new router id available.
Delete ipaddress successfully!
PE2(config-if-vlanInterface-14)#ip binding vpn-instance 2
PE2(config-if-vlanInterface-14)#ip address 192.168.5.1 255.255.255.0

After the above configuration, you can view the VPN instance on the PE through the show ip vpn-instance command

PE1(config)#show ip vpn-instance 1
VPN-Instance Name and ID : 1, 1
Route Distinguisher : 100:1
Export VPN Targets : 111:1
Import VPN Targets : 111:1

Interfaces : Vlanif2

PE1(config)#show ip vpn-instance 2
VPN-Instance Name and ID : 2, 2
Route Distinguisher : 100:2
Export VPN Targets : 222:2
Import VPN Targets : 222:2
Interfaces : Vlanif10

PE2(config)#show ip vpn-instance 1
VPN-Instance Name and ID : 1, 1
Route Distinguisher : 100:1
Export VPN Targets : 111:1
Import VPN Targets : 111:1
Interfaces : Vlanif13

PE2(config)#show ip vpn-instance 2
VPN-Instance Name and ID : 2, 2
Route Distinguisher : 100:2
Export VPN Targets : 222:2
Import VPN Targets : 222:2
Interfaces : Vlanif14

PE2(config)#

5) Establish EBGP peers on PEs and CEs and import VPN routes
Each CE configuration
CE1(config)#show run bgp
![BGP]
router bgp 65410
network 192.168.2.0
neighbor 192.168.2.1 remote-as 100
exit

CE1(config)#show run if
![IF]
interface vlan-interface 2
ip address 192.168.2.2 255.255.255.0
exit

CE2(config)#show run if bgp
![IF]
interface vlan-interface 10
ip address 192.168.3.2 255.255.255.0

exit
![BGP]
router bgp 65420
network 192.168.3.0
neighbor 192.168.3.1 remote-as 100
exit
CE3(config)#show run if bgp
![IF]
interface vlan-interface 13
ip address 192.168.4.2 255.255.255.0
exit
![BGP]
router bgp 65430
network 192.168.4.0
neighbor 192.168.4.1 remote-as 100
exit

CE4(config)#show run if bgp
![IF]
interface vlan-interface 14
ip address 192.168.5.2 255.255.255.0
exit
![BGP]
router bgp 65440
network 192.168.5.0
neighbor 192.168.5.1 remote-as 100
exit

#PE1 configuration
PE1(config-router-bgp)#address-family ipv4 vpn-instance 1
PE1(config-router-vpn-1)#network 192.168.2.0
PE1(config-router-vpn-1)#neighbor 192.168.2.2 remote-as 65410
PE1(config-router-vpn-1)#exit-address-family
PE1(config-router-bgp)#address-family ipv4 vpn-instance 2
PE1(config-router-vpn-2)#network 192.168.3.0
PE1(config-router-vpn-2)#neighbor 192.168.3.2 remote-as 65420
PE1(config-router-vpn-2)#exit-address-family

#PE2 configuration
PE2(config-router-bgp)#address-family ipv4 vpn-instance 1
PE2(config-router-vpn-1)#network 192.168.4.0
PE2(config-router-vpn-1)#neighbor 192.168.4.2 remote-as 65430
PE2(config-router-vpn-1)#exit-address-family
PE2(config-router-bgp)#address-family ipv4 vpn-instance 2

PE2(config-router-vpn-2)#network 192.168.5.0
PE2(config-router-vpn-2)#neighbor 192.168.5.2 remote-as 65440
PE2(config-router-vpn-2)#exit-address-family

After the above configuration, you should be able to query the EBGP peer on the PE and learn VPN routes

PE1(config)#show ip bgp summary

Neighbor	VR	V	AS	MsgRcvd	MsgSent	Up/Down	State/PfxRcd
192.168.2.2	1	4	65410	17	17	00:07:41	Established
192.168.3.2	2	4	65420	17	15	00:07:41	Established

Total number of neighbors 2

PE1(config)#show ip bgp vpn-instance 1
VPN-instance 1 (RD type 0 100:1, VR 1)
Autonomous System number 100, local router ID 192.168.2.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.2.0/24	192.168.2.2			65410 i
*>		0.0.0.0	0		i

Total number of best entries 1

PE1(config)#show ip bgp vpn-instance 2
VPN-instance 2 (RD type 0 100:2, VR 2)
Autonomous System number 100, local router ID 192.168.2.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.3.0/24	192.168.3.2			65420 i
*>		0.0.0.0	0		i

Total number of best entries 1

PE2(config)#show ip bgp summary

Neighbor	VR	V	AS	MsgRcvd	MsgSent	Up/Down	State/PfxRcd
192.168.4.2	1	4	65430	10	10	00:04:19	Established

192.168.5.2 2 4 65440 10 9 00:04:19 Established

Total number of neighbors 2

PE2(config)#show ip bgp vpn-instance 1
VPN-instance 1 (RD type 0 100:1, VR 1)
Autonomous System number 100, local router ID 192.168.4.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.4.0/24	192.168.4.2			65430 i
*>		0.0.0.0	0		i

Total number of best entries 1

PE2(config)#show ip bgp vpn-instance 2
VPN-instance 2 (RD type 0 100:2, VR 2)
Autonomous System number 100, local router ID 192.168.4.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.5.0/24	192.168.5.2			65440 i
*>		0.0.0.0	0		i

Total number of best entries 1

6) Establish MP-IBGP peers between PEs
#PE1 configuration
PE1(config)#router bgp 100
Been in current protocol mode.
PE1(config-router-bgp)#neighbor 3.3.3.3 remote-as 100
PE1(config-router-bgp)#neighbor 3.3.3.3 update-source loopback-interface 0
PE1(config-router-bgp)#address-family vpnv4
PE1(config-router-vpnv4)#neighbor 3.3.3.3 activate
PE1(config-router-vpnv4)#exit-address-family
PE1(config-router-bgp)#exit
PE1(config)#
#PE2 configuration
PE2(config)#router bgp 100
Been in current protocol mode.

PE2(config-router-bgp)#bgp router-id 192.168.4.1
PE2(config-router-bgp)#neighbor 1.1.1.1 remote-as 100
PE2(config-router-bgp)#neighbor 1.1.1.1 update-source loopback-interface 0
PE2(config-router-bgp)#address-family vpnv4
PE2(config-router-vpnv4)#neighbor 1.1.1.1 activate
PE2(config-router-vpnv4)#exit-address-family
PE2(config-router-bgp)#exit

After the above configurations, PE1 and PE2 establish the peer relationship and learn the private network routes of the local CE and remote users. The CE also learns the private network routes of the same VPN user at the remote side.

PE1(config)#show ip bgp summary

Neighbor	VR	V	AS	MsgRcvd	MsgSent	Up/Down	State/PfxRcd
3.3.3.3	0	4	100	3	2	00:00:16	Established
192.168.2.2	1	4	65410	29	29	00:13:36	Established
192.168.3.2	2	4	65420	29	28	00:13:36	Established

Total number of neighbors 3

PE1(config)#show ip bgp vpn-instance 1
VPN-instance 1 (RD type 0 100:1, VR 1)
Autonomous System number 100, local router ID 192.168.2.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.2.0/24	192.168.2.2			65410 i
*>		0.0.0.0	0		i
*>i	192.168.4.0/24	3.3.3.3		100	i

Total number of best entries 2

PE1(config)#show ip bgp vpn-instance 2
VPN-instance 2 (RD type 0 100:2, VR 2)
Autonomous System number 100, local router ID 192.168.2.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.3.0/24	192.168.3.2			65420 i
*>		0.0.0.0	0		i
*>i	192.168.5.0/24	3.3.3.3		100	i

Total number of best entries 2

PE1(config)#show ip route vpn-instance 1
Show ip route information

INET route table - vr: 1, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
127.0.0.1	127.0.0.1	UH	0	lo0	0	0
192.168.2.0/24	192.168.2.1	UC	1	VLAN-IF2	0	0
192.168.2.1	192.168.2.1	UH	0	lo0	0	0
192.168.4.0/24	mpls#1004	ULS	0	*	0	0

Total entries: 4. Printed entries: 4.

PE1(config)#show ip route vpn-instance 2
Show ip route information

INET route table - vr: 2, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
127.0.0.1	127.0.0.1	UH	0	lo0	0	0
192.168.3.0/24	192.168.3.1	UC	1	VLAN-IF10	0	0
192.168.3.1	192.168.3.1	UH	0	lo0	0	0
192.168.5.0/24	mpls#1005	ULS	0	*	0	0

Total entries: 4. Printed entries: 4.

PE1(config)#show ip route
Show ip route information

INET route table - vr: 0, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
1.1.1.1	1.1.1.1	UH	0	LOOPBACK-IF0	0	0
2.2.2.2	mpls#1001	UHLS	360	*	0	0
3.3.3.3	mpls#1003	UHLS	193	*	0	0
127.0.0.0/8	127.0.0.1	UR	0	lo0	0	0
127.0.0.1	127.0.0.1	UH	5	lo0	0	0
172.168.1.0/24	172.168.1.1	UC	1	VLAN-IF3	0	0
172.168.1.1	172.168.1.1	UH	0	lo0	0	0

172.168.2.0/24 mpls#1002 ULS 26 * 0
0

Total entries: 8. Printed entries: 8.

PE2(config)#show ip bgp summary

Neighbor	VR	V	AS	MsgRcvd	MsgSent	Up/Down	State/PfxRcd
1.1.1.1	0	4	100	3	2	00:00:10	Established
192.168.4.2	1	4	65430	21	21	00:09:38	Established
192.168.5.2	2	4	65440	21	20	00:09:38	Established

Total number of neighbors 3

PE2(config)#show ip bgp vpn-instance 1
VPN-instance 1 (RD type 0 100:1, VR 1)
Autonomous System number 100, local router ID 192.168.4.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*>i	192.168.2.0/24	1.1.1.1		100	i
*	192.168.4.0/24	192.168.4.2			65430 i
*>		0.0.0.0	0		i

Total number of best entries 2

PE2(config)#show ip bgp vpn-instance 2
VPN-instance 2 (RD type 0 100:2, VR 2)
Autonomous System number 100, local router ID 192.168.4.1
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*>i	192.168.3.0/24	1.1.1.1		100	i
*	192.168.5.0/24	192.168.5.2			65440 i
*>		0.0.0.0	0		i

Total number of best entries 2

PE2(config)#show ip route vpn-instance 1
Show ip route information

INET route table - vr: 1, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
127.0.0.1	127.0.0.1	UH	0	lo0	0	0
192.168.2.0/24	mpls#1004	ULS	0	*	0	0
192.168.4.0/24	192.168.4.1	UC	1	VLAN-IF13	0	0
192.168.4.1	192.168.4.1	UH	0	lo0	0	0

Total entries: 4. Printed entries: 4.

PE2(config)#show ip route vpn-instance 2
Show ip route information

INET route table - vr: 2, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
127.0.0.1	127.0.0.1	UH	0	lo0	0	0
192.168.3.0/24	mpls#1005	ULS	0	*	0	0
192.168.5.0/24	192.168.5.1	UC	1	VLAN-IF14	0	0
192.168.5.1	192.168.5.1	UH	0	lo0	0	0

Total entries: 4. Printed entries: 4.

PE2(config)#show ip route
Show ip route information

INET route table - vr: 0, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
1.1.1.1	mpls#1001	UHLS	142	*	0	0
2.2.2.2	mpls#1002	UHLS	391	*	0	0
3.3.3.3	3.3.3.3	UH	0	LOOPBACK-IF0	0	0
127.0.0.0/8	127.0.0.1	UR	0	lo0	0	0
127.0.0.1	127.0.0.1	UH	5	lo0	0	0
172.168.1.0/24	mpls#1003	ULS	70	*	0	0
172.168.2.0/24	172.168.2.2	UC	1	VLAN-IF4	0	0
172.168.2.2	172.168.2.2	UH	0	lo0	0	0

Total entries: 8. Printed entries: 8.

CE1(config)#show ip route
Show ip route information

INET route table - vr: 0, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
127.0.0.0/8	127.0.0.1	UR	0	lo0	0	0
127.0.0.1	127.0.0.1	UH	4	lo0	0	0
192.168.2.0/24	192.168.2.2	UC	1072	VLAN-IF2	0	0
192.168.2.2	192.168.2.2	UH	1350	lo0	0	0
192.168.4.0/24	192.168.2.1	UG	0	VLAN-IF2	0	0

Total entries: 5. Printed entries: 5.

CE1(config)#show ip bgp
Autonomous System number 65410, local router ID 192.168.2.2
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.2.0/24	192.168.2.1			100 i
*>		0.0.0.0	0		i
*>	192.168.4.0/24	192.168.2.1			100 i

Total number of best entries 2

CE2(config)#show ip bgp
Autonomous System number 65420, local router ID 192.168.3.2
Status codes: s suppressed, * valid, > best, i internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	NextHop	Metric	LocalPref	Path
*	192.168.3.0/24	192.168.3.1			100 i
*>		0.0.0.0	0		i
*>	192.168.5.0/24	192.168.3.1			100 i

Total number of best entries 2

CE2(config)#show ip route
Show ip route information

INET route table - vr: 0, table: 254

Destination	Gateway	Flags	Use	Interface	Metric	MTU
127.0.0.0/8	127.0.0.1	UR	0	lo0	0	0
127.0.0.1	127.0.0.1	UH	4	lo0	0	0
192.168.3.0/24	192.168.3.2	UC	1108	VLAN-IF10	0	0
192.168.3.2	192.168.3.2	UH	1395	lo0	0	0

192.168.5.0/24 192.168.3.1 UG 0 VLAN-IF10 0 0

Total entries: 5. Printed entries: 5.

48.6.3 Results Verification

Users in the same VPN can communicate with each other. Different VPN users can not communicate with each other. For example, CE1 can only ping ce3 successfully, CE2 can only ping ce4 successfully.

CE1(config)#ping 192.168.4.2

PING 192.168.4.2: with 32 bytes of data:

reply from 192.168.4.2: bytes=32 time<10ms TTL=62
reply from 192.168.4.2: bytes=32 time<10ms TTL=62
reply from 192.168.4.2: bytes=32 time=20ms TTL=62
reply from 192.168.4.2: bytes=32 time<10ms TTL=62
reply from 192.168.4.2: bytes=32 time<10ms TTL=62

----192.168.4.2 PING Statistics----
5 packets transmitted, 5 packets received, 0% packet loss
round-trip (ms) min/avg/max = 0/4/20

CE1(config)#ping 192.168.5.2

PING 192.168.5.2: with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.
Request timed out.
Request timed out.
no answer from 192.168.5.2

CE1(config)#ping 192.168.3.2

PING 192.168.3.2: with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.

Request timed out.

Request timed out.

no answer from 192.168.3.2

49.PBR Configuration

49.1 PBR Overview

Policy-based routing is a mechanism for routing policies based on user-defined policies. In contrast to IP address-based routing table forwarding, policy routing can flexibly route routes based on information such as the source address of the packets.

In general, a PBR takes precedence over a common route. That is, when a device forwards a packet, it compares the packet with the policy-based routing rule. If the match criterion is matched, the packet is forwarded according to PBR. If the packet does not match the conditions of the policy-based route, the packet is forwarded according to the common route.

49.1.1 PBR Mode（policy-based-route）

PBR specifies matching rules by ACLs. It supports the next hop, priority, and default next hop of packets. Currently, only IPv4 unicast policy routing is supported.

PBR supports dynamic modification of ACL rules to dynamically update matching rules to implement flexible control of services.

- In PBR mode, if the packet matches the policy-based routing condition but the configured next-hop address does not exist, the packet will continue to be forwarded.
- If the default next hop is configured, the normal route is forwarded. If the route cannot be matched, the route will be forwarded.
- The default next-hop command (def-route) is used when the next hop is not configured in the policy-based route or the next hop is invalid. The destination IP address of the packet is not found in the routing table. , And the default next hop configured by PBR will be used.

49.1.2 Configure PBR Policy

Configure PBR

Operation	Command	Mark
enter global configuration mode	**configure terminal**	-
Configure ACL	**access-list [**num**] permit [**ip info**] def-route**	optional
Configure PBR	**policy-based-route** { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] } { [**next-hop [**ip address**]**] }	optional