FSOS

ERRP Configuration

Contents

1. ERRP Configuration

1.1 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.

1.1.1 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.

1.1.2 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.

1.1.3 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.

1.1.4 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.

1.2 ERRP Configuration

1.2.1 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

1.2.2 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

1.2.3 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

1.2.4 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

1.2.5 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***]	

1.2.6 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

1.2.7 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

1.2.8 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

1.2.9 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

1.2.10 Clear Protocol Message Statistic

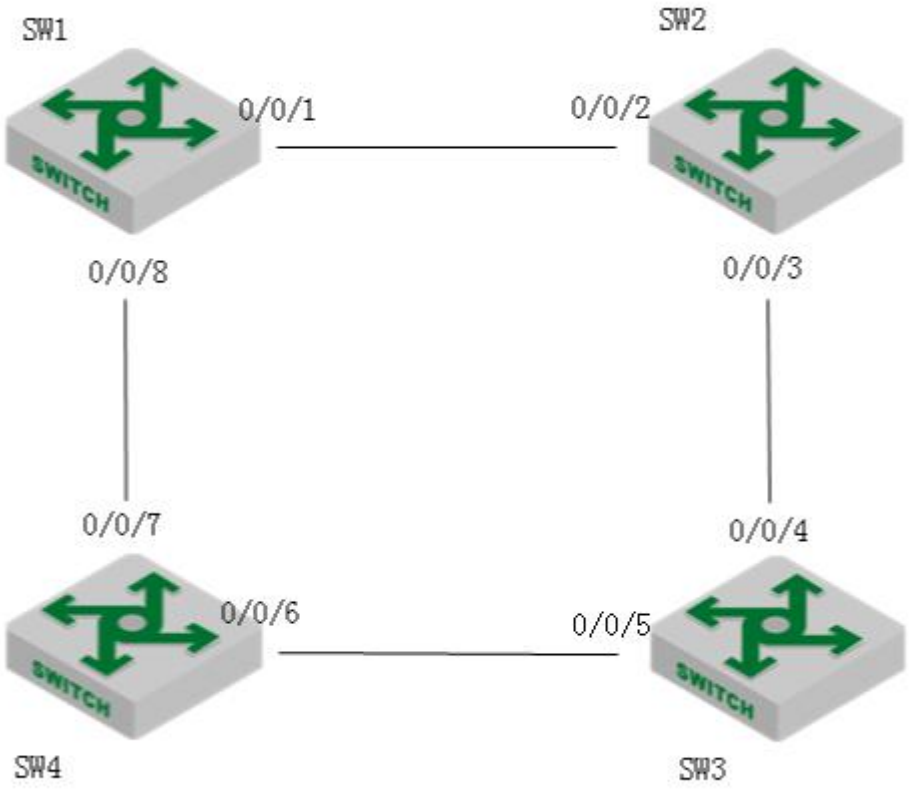
Clear protocol message statistic

operation	command	remark
Enter the global configuration mode	configure terminal	-
Clear the statistics	clear errp [domain ***domain-id*** [ring ***ring-id***]]	optional

1.3 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
```