

QSW-3900 Ethernet Switch
User's Manual



QTECH
МИР ДОСТУПНЕЕ

Content

Chapter 1	Accessing Switch.....	1-14
1.1	Command Line Interface	1-14
1.1.1	Command Line Configuration Mode	1-14
1.1.2	Command Syntax Comprehension	1-16
1.1.3	Syntax Help	1-16
1.1.4	History command	1-17
1.1.5	Symbols in command.....	1-17
1.2	Command Symbols Description	1-18
1.2.1	Command Parameter Categories	1-18
1.3	User management	1-19
1.3.1	System default user name	1-19
1.3.2	Add user	1-19
1.3.3	Modify password	1-19
1.3.4	Modify privilege	1-20
1.3.5	Remove user name.....	1-20
1.3.6	View system user information	1-20
1.4	Remote authentication of administrator	1-21
1.4.1	Start RADIUS/TACACS+ remote authentication	1-21
1.4.2	Display authentication configuration	1-21
1.5	Ways of managing switch.....	1-21
1.5.1	Manage switch by hyper terminal	1-21
1.5.2	Manage switch by telnet	1-22
1.6	Brief introduction of SSH	1-22
1.7	SSH Configuration list	1-23
1.7.1	Enable/disable SSH function of the device	1-23
1.7.2	SSH key configuration	1-23
1.7.3	Others	1-24
Chapter 2	Switch Manage and Maintenance	2-25
2.1	System IP configuration	2-25
2.2	Configure manage IP interface	2-25
2.3	Configuration ip address by manual operation	2-25
2.4	Configuration Files Management.....	2-25
2.4.1	Edit configuration files.....	2-25
2.4.2	Modify and save current configuration.....	2-25
2.4.3	Erase saved configuration	2-26
2.4.4	Execute saved configuration.....	2-26
2.4.5	Display saved configuration	2-26
2.4.6	Display current configuration	2-26
2.4.7	Configure file executing mode shift.....	2-27
2.5	Online Loading Upgrade Program	2-27
2.5.1	Upload and download files by TFTP	2-27
2.5.2	Upload and download files by FTP	2-28
2.5.3	Download files by Xmodem	2-28
2.6	Reboot.....	2-29
2.7	System Maintenance.....	2-29
2.7.1	Basic Configuration and Management	2-30
2.7.2	Network connecting test command	2-30
2.7.3	Loopback test command.....	2-31
2.7.4	Remote access restriction	2-31
2.7.5	The number of Telnet user restriction	2-32
2.7.6	Routing tracert command	2-32

2.7.7	Packets rate limit to CPU	2-32
2.8	Monitor system by SNMP	2-33
2.8.1	Brief introduction of SNMP	2-33
2.9	SNMP Mechanism	2-33
2.10	SNMP Protocol Version	2-34
2.11	MIB Overview	2-34
2.12	SNMP Configuration	2-34
2.12.1	Configure community name and accessing right.	2-35
2.12.2	Configure sysContact.....	2-35
2.12.3	Configure Trap destination host address.....	2-36
2.12.4	Configure sysLocation	2-36
2.12.5	Configure sysName	2-37
2.12.6	Configure notify.....	2-37
2.12.7	Configure engine id.....	2-37
2.12.8	Configure view	2-38
2.12.9	Configure group	2-38
2.12.10	Configure user	2-39
2.13	Enable/disable dlf forward packet.....	2-40
2.14	CPU Alarm Configuration	2-40
2.14.1	Brief introduction of CPU alarm	2-40
2.14.2	CPU alarm configuration list.....	2-40
2.14.3	Enable/disable CPU alarm.....	2-41
2.14.4	Configure CPU busy or unbusy threshold	2-41
2.14.5	Display CPU alarm information.....	2-41
2.15	Anti-DOS Attack.....	2-42
2.15.1	IP segment anti-attack	2-42
Chapter 3	MAC address table management.....	3-43
3.1	Introduction to Bridging	3-43
3.2	Major Functionalities of Bridges.....	3-43
3.2.1	Maintaining the bridge table.....	3-43
3.2.2	Forwarding and filtering	3-45
3.3	Brief introduction of MAC address table management	3-47
3.4	MAC address table management list	3-47
3.5	Configure system MAC address aging time	3-48
3.6	Configure MAC address item	3-48
3.6.1	Add MAC address.....	3-48
3.6.2	Add blackhole MAC address	3-48
3.6.3	Delete MAC address item.....	3-49
3.6.4	Display MAC address table	3-49
3.6.5	Enable/disable MAC address learning	3-49
3.6.6	Display MAC address learning	3-50
3.6.7	Modify MAC address learning mode	3-50
Chapter 4	Port Configuration.....	4-51
4.1	Port configuration introduction.....	4-51
4.2	Port Configuration	4-51
4.2.1	Port related configuration.....	4-51
4.2.2	Enter interface configuration mode.....	4-51
4.2.3	Enable/disable specified interface	4-51
4.2.4	Configure interface duplex mode and speed rate	4-52
4.2.5	Interface Priority Configuration	4-52
4.2.6	Interface description configuration.....	4-52
4.2.7	Ingress/egress bandwidth-control configuration.....	4-53

4.2.8	Enable/disable VLAN filtration of receiving packet of interface.....	4-53
4.2.9	Interface ingress acceptable-frame configuration	4-53
4.2.10	Enable/disable interface flow-control.....	4-53
4.2.11	Port mode configuration.....	4-54
4.2.12	Trunk allowed VLAN configuration	4-54
4.2.13	The default vlan-id of trunk port configuration	4-54
4.2.14	Add access port to specified VLAN	4-55
4.2.15	Display interface information	4-55
4.2.16	Display/ clear interface statistics information.....	4-55
4.3	Interface mirror.....	4-55
4.3.1	Brief introduction of interface mirror	4-55
4.3.2	Interface mirror configuration.....	4-56
4.4	Brief introduction of Port LACP.....	4-56
4.4.1	LACP.....	4-57
4.4.2	Manual Link Aggregation	4-58
4.4.3	Static LACP link aggregation	4-58
4.5	Load-Balance in a Link Aggregation Group.....	4-59
4.6	Aggregation Port Group	4-59
4.7	Link aggregation configuration.....	4-59
4.8	Interface BPDU-rate configuration.....	4-61
4.8.1	Brief introduction of interface CAR	4-61
4.8.2	Port CAR configuration command list.....	4-61
4.8.3	Enable/disable interface globally	4-61
4.8.4	Enable/disable interface CAR on interface.....	4-61
4.8.5	Configure the reopen time of the port shutdown by port-car.....	4-62
4.8.6	Configure the port-car-rate	4-62
4.8.7	Display port-car information.....	4-62
4.9	Port Alarm Configuration	4-62
4.9.1	Brief introduction of port alarm configuration.....	4-62
4.9.2	Port alarm configuration list	4-62
4.9.3	Enable/disable port alarm globally.....	4-63
4.9.4	Enable/disable port alarm on the port.....	4-63
4.9.5	Configure the exceed threshold and normal threshold of port alarm	4-63
4.9.6	Display port alarm	4-63
4.10	Shutdown-control feature	4-64
4.11	Interface shutdown-control configuration list	4-64
4.11.1	Configuration mode and time.....	4-64
4.11.2	Configuration interface shutdown-control	4-64
4.11.3	Display shutdown-control.....	4-65
4.12	Port isolation configuration	4-65
4.13	Strom control configuration.....	4-65
Chapter 5	VLAN Configuration.....	5-67
5.1	Introduction to VLAN	5-67
5.1.1	VLAN Overview.....	5-67
5.1.2	VLAN Fundamental.....	5-68
5.1.3	VLAN Classification	5-68
5.1.4	VLAN Interface.....	5-68
5.1.5	Port-Based and 802.1Q VLAN.....	5-69
5.1.6	Port link type	5-69
5.1.7	Default VLAN	5-69
5.1.8	Super VLAN	5-69
5.1.9	VLAN interface type.....	5-69

5.1.10	Default VLAN	5-70
5.2	VLAN configuration list	5-70
5.2.1	Create/delete VLAN	5-70
5.2.2	Add/delete VLAN interface	5-71
5.2.3	Specify/restore VLAN description	5-71
5.2.4	Configure interface type	5-71
5.2.5	Configure interface default vlan ID	5-72
5.2.6	Configure tag vlan	5-72
5.2.7	Display VLAN information	5-72
5.3	Brief introduction of GVRP	5-73
5.3.1	GARP protocol	5-73
5.3.2	Brief introduction of GVRP	5-73
5.3.3	GARP messages and timers	5-73
5.4	GVRP Configuration list	5-75
5.4.1	Enable/disable global GVRP	5-75
5.4.2	Enable/disable GVRP on a port	5-76
5.4.3	Display GVRP	5-76
5.4.4	Add/delete vlan that can be dynamic learnt by GVRP	5-76
5.4.5	Display vlan that can be learnt by GVRP	5-77
5.4.6	Examples for GVRP configuration	5-77
5.5	Brief introduction of QinQ	5-77
5.5.1	Introduction to QinQ	5-77
5.5.2	Implementations of QinQ	5-78
5.5.3	Adjustable TPID Value of QinQ Frames	5-78
5.6	QinQ configuration list	5-79
5.6.1	Configure global QinQ	5-79
5.6.2	Configure QinQ mode of interface	5-80
5.6.3	Configure interface dynamic QinQ	5-80
5.6.4	Enable/disable vlan-swap	5-81
5.6.5	Configure global vlan-swap	5-81
5.6.6	Configure rewrite-outer-vlan	5-81
5.6.7	Display dynamic QinQ	5-82
5.6.8	Display vlan-swap	5-82
5.6.9	Display rewrite-outer-vlan	5-82
Chapter 6	Layer 3 Configuration	6-84
6.1	Brief Introduction of Layer 3 switching	6-84
6.2	Layer 3 Cnfiguration list	6-84
6.2.1	VLAN division and the creation of layer 3 interface	6-84
6.2.2	Transmission mode configuration	6-84
6.2.3	Create VLAN interface for normal VLAN	6-85
6.2.4	Create superVLAN interface and add VLAN to superVLAN	6-85
6.2.5	Configure IP address for VLAN interface or superVLAN interface	6-85
6.2.6	Configure accessing IP address range of VLAN or superVLAN interface	6-85
6.2.7	ARP proxy configuration	6-86
6.2.8	Display interface configuration	6-86
6.3	Brief introduction of static routing	6-86
6.3.1	Default Route	6-86
6.3.2	Application Environment of Static Routing	6-87
6.4	Static routing configuration list	6-87
6.4.1	Add/delete static route	6-87
6.4.2	Display route table information	6-87

Chapter 7	VRRP	7-89
7.1	Overview.....	7-89
7.1.1	Background	7-89
7.1.2	Benefits	7-89
7.2	Introduction to VRRP	7-90
7.2.1	Concepts	7-90
7.2.2	Introduction to Virtual Router	7-91
7.2.3	VRRP Working Process.....	7-91
7.2.4	Backup's Monitoring of the Master State	7-93
7.3	Application Scenarios.....	7-93
7.3.1	4Master/Backup	7-93
7.3.2	4.2 Load Balancing	7-94
7.3.3	Master's Monitoring of Uplinks Through BFD/NQA.....	7-95
7.3.4	Backup's Monitoring of Master State Using BFD	7-96
7.4	References	7-96
Chapter 8	RIP Configuration	8-99
8.1	Brief introduction of RIP	8-99
8.2	RIP Overview.....	8-99
8.2.1	RIP Working Mechanism	8-100
8.2.2	RIP Version	8-101
8.2.3	RIP Message Format.....	8-101
8.2.4	TRIP	8-103
8.2.5	Protocols and Standards.....	8-103
8.3	RIP configuration list	8-103
8.3.1	Enable RIP	8-104
8.3.2	Specify IP network to run RIP protocol	8-104
8.3.3	RIP working status of specified interface.....	8-104
8.3.4	RIP version of specified interface	8-104
8.3.5	Enable host routing	8-105
8.3.6	Enable route convergence.....	8-105
8.3.7	Configure authentication to RIP packet	8-105
8.3.8	Configure split	8-105
8.3.9	Configure metricin.....	8-106
8.3.10	Define prefix list	8-106
8.3.11	Configure redistribution.....	8-106
8.3.12	Configure distribute-list	8-107
8.3.13	Display RIP configuration.....	8-107
Chapter 9	BFD.....	9-108
9.1	Overview.....	9-108
9.1.1	Background	9-108
9.1.2	Benefits	9-108
9.2	BFD Implementation	9-108
9.2.1	Mechanism.....	9-108
9.2.2	BFD Packets	9-110
9.2.3	BFD Session Establishment	9-111
9.2.4	Timer Negotiation.....	9-112
9.2.5	Fault Detection	9-114
9.3	Application Scenarios.....	9-114
9.3.1	Configuring BFD for Routing Protocols	9-114
9.3.2	Configuring BFD for Fast Reroute	9-115
9.3.3	Configuring BFD for VRRP	9-116
Chapter 10	OSPF Configuration	10-120

10.1	Brief introduction of OSPF.....	10-120
10.1.1	Basic Concepts	10-121
10.1.2	OSPF Area Partition and Route Summarization	10-122
10.1.3	Classification of OSPF Networks.....	10-127
10.1.4	DR and BDR	10-128
10.1.5	OSPF Packet Formats.....	10-129
10.1.6	Supported OSPF Features	10-136
10.1.7	Protocols and Standards.....	10-138
10.2	OSPF Configuration list.....	10-138
10.2.1	Enable/disable OSPF.....	10-139
10.2.2	Configure router ID	10-139
10.2.3	Specify interface and area id	10-139
10.2.4	Configure area authentication type.....	10-139
10.2.5	Configure interface type.....	10-140
10.2.6	Configure interface cost.....	10-140
10.2.7	Configure priority when selecting DR.....	10-141
10.2.8	Configure Hello time interval.....	10-141
10.2.9	Configure interface invalid time of neighbour routers.....	10-142
10.2.10	Configure retransmission LSA time interval of neighbor router.....	10-142
10.2.11	Configure time needed when interface sending link state update packet	10-143
10.2.12	Configure packet authentication key	10-143
10.2.13	Configure STUB area of OSPF.....	10-143
10.2.14	Configure route convergence in OSPF.....	10-144
10.2.15	Configure OSPF virtual connection	10-144
10.2.16	Configure route introduced by OSPF other route protocol.....	10-145
10.2.17	Configure OSPF introduced default route	10-146
10.2.18	Configure external route parameter received by OSPF	10-146
10.2.19	OSPF monitor and maintain	10-146
Chapter 11	BGP Configuration.....	11-148
11.1	Brief Introduction of BGP.....	11-148
11.1.1	BGP Message Type	11-149
11.1.2	BGP Route Attributes.....	11-151
11.1.3	BGP Routing Policy	11-156
11.1.4	Problems in Large-Scale BGP Networks.....	11-156
11.1.5	Protocol Standard	11-160
11.2	BGP Configuration Task List.....	11-160
11.3	BGP Configuration	11-161
11.3.1	Enable/disable BGP.....	11-161
11.3.2	Configure BGP peer.....	11-162
11.3.3	Configure BGP timer.....	11-163
11.3.4	Configure local preference.....	11-163
11.3.5	Configure AS MED.....	11-164
11.3.6	Compare MED from different AS neighbors	11-164
11.3.7	Configure BGP route aggregation	11-164
11.3.8	Configure route information of IGP protocol introduced by BGP	11-164
11.3.9	Configure BGP distribution list.....	11-165
11.3.10	Define AS path list.....	11-165
11.3.11	BGP monitor and maintenance.....	11-166
Chapter 12	Multicast Protocol Configuration	12-167
12.1	Overview.....	12-167
12.1.1	Background.....	12-167

12.1.2	Benefits	12-167
12.2	Multicast Implementation.....	12-167
12.2.1	Multicast Addressing Mechanism	12-168
12.2.2	Multicast Address Mapping.....	12-169
12.2.3	Group Membership Management.....	12-169
12.2.4	Multicast Packet Forwarding.....	12-171
12.2.5	Multicast Routing Protocols	12-171
12.2.6	Multicast Models	12-174
12.3	GMRP Overview	12-174
12.3.1	GMRP Configuration.....	12-175
12.4	IGMP Snooping Overview	12-176
12.4.1	IGMP Snooping.....	12-176
12.4.2	Basic Concepts in IGMP Snooping.....	12-177
12.4.3	How IGMP Snooping Works	12-178
12.4.4	Processing of Multicast Protocol Messages.....	12-180
12.4.5	Protocols and Standards.....	12-182
12.5	IGMP Snooping configuration.....	12-185
12.5.1	IGMP Snooping multicast interface aging time configuration.....	12-185
12.5.2	IGMP Snooping max-response-time configuration.....	12-186
12.5.3	IGMP Snooping interface fast-leave configuration	12-186
12.5.4	Configure the number of the multicast group allowed learning.....	12-186
12.5.5	IGMP Snooping permit/deny group configuration	12-186
12.5.6	IGMP Snooping route-port forward configuration.....	12-187
12.5.7	Enable/disable IGMP Snooping querier	12-187
12.5.8	Configure IGMP Snooping query-interval.....	12-187
12.5.9	Configure IGMP Snooping querier vlan.....	12-187
12.5.10	Configure IGMP Snooping query max response.....	12-188
12.5.11	Configure IGMP Snooping query source IP	12-188
12.5.12	Configure IGMP Snooping route port aging	12-188
12.5.13	Add IGMP Snooping route port.....	12-188
12.6	Static Multicast Configuration	12-189
12.6.1	Brief introduction of Static Multicast.....	12-189
12.6.2	Static Multicast Configuration	12-189
12.6.3	Create multicast group.....	12-189
12.6.4	Add interfaces to multicast group	12-189
12.6.5	Display multicast group information.....	12-190
12.6.6	Delete interface members from multicast group.....	12-190
12.6.7	Delete multicast group	12-190
12.7	Cross-VLAN multicast Configuration	Ошибка! Закладка не определена.
12.7.1	Brief Introduction of Cross-Vlan multicast	Ошибка! Закладка не определена.
12.7.2	Cross-VLAN Multicast Configuration	Ошибка! Закладка не определена.
12.7.3	Enable/disable cross-vlan multicast	Ошибка! Закладка не определена.
12.7.4	Configure tag/untag attribution of multicast packet transmission and vlan-id of the tagged attribution	Ошибка! Закладка не определена.
12.7.5	Display cross-vlan multicast	Ошибка! Закладка не определена.
Chapter 13	DHCP Configuration	13-205
13.1	Brief introduction of DHCP	13-205
13.2	Technical details.....	13-206
13.2.1	DHCP discovery.....	13-206
13.2.2	DHCP offers	13-206
13.2.3	DHCP requests	13-206
13.2.4	DHCP acknowledgement.....	13-207

13.2.5	DHCP information	13-207
13.2.6	DHCP releasing	13-207
13.2.7	Client configuration parameters.....	13-207
13.2.8	Options.....	13-207
13.2.9	DHCP IP Address Assignment.....	13-207
13.3	DHCP server configuration list.....	13-210
13.3.1	Enable DHCP relay.....	13-211
13.3.2	Configure DHCP server	13-211
13.3.3	Specify DHCP server for layer 3 interface	13-211
13.3.4	Display DHCP server configuration	13-211
13.3.5	Hide DHCP server	13-212
13.4	Local IP Address Pool Configuration.....	13-212
13.4.1	Enter IP address pool configuration mode	13-212
13.4.2	Configure gateway and netmask of local IP address pool	13-213
13.4.3	Configure local IP address pool network interface	13-213
13.4.4	Disable/enable specified IP address in IP address pool	13-213
13.4.5	Configure lease time.....	13-214
13.4.6	Configure DNS.....	13-214
13.4.7	Configure WINS.....	13-214
13.4.8	Display IP address pool configuration	13-214
13.4.9	Configure ip-bind.....	13-215
13.4.10	Display ip-bind	13-215
13.4.11	Add dhcp client	13-215
13.4.12	Show dhcp client.....	13-215
13.5	Introduction to DHCP Relay Agent.....	13-216
13.5.1	Usage of DHCP Relay Agent.....	13-216
13.5.2	DHCP Relay Agent Fundamentals	13-216
13.5.3	Option 82 Supporting.....	13-217
13.6	DHCP relay configuration list.....	13-219
13.6.1	Enable DHCP relay.....	13-219
13.6.2	Configure vlan interface.....	13-220
13.6.3	Support relay option82.....	13-220
13.7	Introduction DHCP snooping	13-220
13.8	DHCP snooping configuration list	13-222
13.8.1	Enable DHCP snooping.....	13-222
13.8.2	Configure trust ports	13-222
13.8.3	Configure max host number	13-222
13.8.4	Configure IP source guard.....	13-222
13.8.5	Show DHCP snooping of ports	13-222
13.8.6	Show DHCP snooping configuration of VLANs.....	13-222
13.8.7	Show information of clients.....	13-222
Chapter 14	ARP Configuration.....	14-223
14.1	Brief Introduction of ARP	14-223
14.1.1	ARP announcements	14-224
14.1.2	ARP probe.....	14-224
14.1.3	ARP mediation	14-224
14.1.4	Variants of the protocol	14-224
14.1.5	Inverse ARP and Reverse ARP	14-224
14.2	ARP spoofing.....	14-225
14.2.1	How ARP spoofing works?.....	14-225
14.2.2	ARP Spoofing/poising Animation	14-225
14.3	ARP-Proxy.....	14-225

14.4	Anti-flood ARP	14-226
14.5	ARP configuration list	14-226
14.5.1	Add and delete ARP table item	14-226
14.5.2	Display ARP table item	14-227
14.5.3	Configure ARP aging time	14-227
14.5.4	Display ARP aging time	14-227
14.5.5	Display ARP table item	14-227
14.5.6	Enable/disable ARP anti-flood attack.....	14-227
14.5.7	Configure deny action and threshold of ARP anti-flood.....	14-228
14.5.8	Configure ARP anti-flood recover-time	14-228
14.5.9	ARP anti-flood MAC recover.....	14-228
14.5.10	Display ARP anti-flood attack information.....	14-229
14.5.11	Bind blackhole mac generated by arp anti-flood to be general	14-229
14.5.12	Enable/disable ARP anti-spoofing	14-229
14.5.13	Configure unknown ARP packet handling strategy	14-230
14.5.14	Enable/disable ARP anti-spoofing valid-check	14-230
14.5.15	Enable/disable ARP anti-spoofing deny-disguiser.....	14-230
14.5.16	Display ARP anti-spoofing	14-231
14.5.17	Configure trust port of ARP anti-attack	14-231
Chapter 15	ACL Configuration	15-232
15.1	ACL Overview	15-232
15.1.1	ACL Match Order	15-232
15.1.2	Ways to Apply ACL on a Switch.....	15-233
15.1.3	ACLs Based on Time Ranges.....	15-233
15.2	Configuring ACL.....	15-234
15.2.1	Matching order configuration	15-234
15.2.2	ACL support	15-234
15.3	ACL configuration.....	15-235
15.3.1	Configuration list	15-235
15.3.2	Configure time range	15-235
15.3.3	Standard ACL.....	15-236
15.3.4	Define extended ACL.....	15-236
15.3.5	Define layer 2 ACL.....	15-237
15.3.6	Activate ACL.....	15-238
15.3.7	Monitor and maintenance of ACL	15-238
Chapter 16	QOS Configuration	16-243
16.1	Brief introduction of QOS	16-243
16.1.1	Flow.....	16-243
16.1.2	Traffic classification.....	16-243
16.1.3	Access control list	16-243
16.1.4	Packet filtration	16-243
16.1.5	Flow monitor.....	16-244
16.1.6	Interface speed limitation.....	16-244
16.1.7	Redirection	16-244
16.1.8	Priority mark.....	16-244
16.1.9	Choose interface outputting queue for packet.....	16-244
16.1.10	Queue scheduler.....	16-244
16.1.11	cos-map	16-245
16.1.12	Flow mirror	16-245
16.1.13	Statistics based on flow.....	16-245
16.1.14	Copy packet to CPU	16-245
16.2	QOS Configuration.....	16-245

16.2.1	QoS Configuration list.....	16-245
16.2.2	Packet redirection configuration	16-245
16.2.3	Priority configuration	16-246
16.2.4	Queue-scheduler configuration.....	16-246
16.2.5	The cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol.....	16-247
16.2.6	Flow mirror configuration	16-247
16.2.7	Flow statistic configuration.....	16-247
16.3	Monitor and maintenance of QoS	16-248
Chapter 17	STP Configuration	17-249
17.1	Brief introduction of STP Configuration.....	17-249
17.1.1	Introduction to STP	17-249
17.1.2	Introduction to MSTP	17-256
17.1.3	Protocols and Standards.....	17-260
17.2	STP Configuration.....	17-261
17.2.1	STP Configuration list	17-261
17.2.2	Enable/disable STP	17-261
17.2.3	Enable/disable interface STP	17-261
17.2.4	Configure STP priority.....	17-262
17.2.5	Configure switch Forward Delay.....	17-262
17.2.6	Configure Hello Time	17-262
17.2.7	Configure Max Age	17-263
17.2.8	Configure path cost of specified interfaces.....	17-263
17.2.9	Configure STP priority of specified port.....	17-263
17.2.10	Configure spanning-tree root-guard	17-264
17.2.11	Configure interface to force to send rstp packet.....	17-264
17.2.12	Configure link type of specified interface.....	17-264
17.2.13	Configure the current port as an edge port	17-264
17.2.14	Configure the speed limit of sending BPDU of specified interface.....	17-265
17.2.15	STP monitor and maintenance	17-265
17.2.16	Enable/disable STP remote-loop-detect.....	17-266
17.3	Brief Introduction of MSTP	17-266
17.4	MSTP Configuration.....	17-267
17.4.1	MSTP configuration list.....	17-267
17.4.2	Configure MSTP timer parameter.....	17-267
17.4.3	Configure MSTP configuration mark	17-267
17.4.4	Configure MSTP netbridge priority	17-268
17.4.5	Configure MSTP interface edge interface status.....	17-268
17.4.6	Configure MSTP interface link type	17-268
17.4.7	Configure MSTP interface path cost.....	17-268
17.4.8	Configure MSTP interface priority.....	17-269
17.4.9	Configure spanning-tree mst root-guard.....	17-269
17.4.10	Display MSTP configuration information	17-269
17.4.11	Enable/disable digest snooping.....	17-270
17.4.12	Configure Ignore of VLAN.....	17-270
Chapter 18	802.1X Configuration Command.....	20-275
18.1	Brief introduction of 802.1X configuration	20-275
18.2	802.1X Configuration	20-275
18.2.1	AAA configuration mode	20-275
18.3	RADIUS and TACACS+ Server Configuration.....	20-275
18.3.1	System default user	20-276
18.3.2	User's authentication	20-276

18.4	Local authentication configuration	20-276
18.4.1	Add users	20-276
18.4.2	Change password	20-277
18.4.3	Modify User's Privilege Level	20-278
18.4.4	Delete User	20-278
18.4.5	Show users	20-278
18.5	Remote authentication configuration	20-279
18.5.1	Configure RADIUS to be remote authentication server	20-279
18.5.2	Configure TACACS+ remote authentication	20-279
18.5.3	802.1X Configuration	20-280
Chapter 19	SNTP Client Configuration	21-282
19.1	Brief introduction of SNTP protocol	21-282
19.2	SNTP client configuration	21-282
19.2.1	Enable/disable SNTP client	21-282
19.2.2	SNTP client working mode configuration	21-282
19.2.3	SNTP client unicast server configuration	21-283
19.2.4	SNTP client broadcast delay configuration	21-283
19.2.5	SNTP client multicast TTL configuration	21-283
19.2.6	SNTP client poll interval configuration	21-283
19.2.7	SNTP client retransmit configuration	21-284
19.2.8	SNTP client valid server configuration	21-284
19.2.9	SNTP client MD5 authentication configuration	21-284
Chapter 20	Syslog Configuration	22-285
20.1	Brief introduction of Syslog	22-285
20.2	Syslog Configuration	22-285
20.2.1	Enable/disable Syslog	22-286
20.2.2	Syslog sequence number configuration	22-286
20.2.3	Syslog time stamps configuration	22-286
20.2.4	Syslog terminal outputting configuration	22-286
20.2.5	Syslog logging buffered outputting configuration	22-287
20.2.6	Syslog Flash storage outputting configuration	22-287
20.2.7	Syslog logging host outputting configuration	22-288
20.2.8	Syslog SNMP Agent outputting configuration	22-289
20.2.9	Module debug configuration	22-289
Chapter 21	LLDP configuration	23-290
21.1	Brief introduction of LLDP protocol	23-290
21.1.1	LLDP Overview	23-290
21.2	LLDP configuration	23-291
21.2.1	LLDP configuration list	23-291
21.2.2	Enable/disable global LLDP	23-291
21.2.3	Configure LLDP hello-time	23-291
21.2.4	Configure LLDP hold-time	23-291
21.2.5	Interface LLDP packet receiving/sending mode configuration	23-292
21.2.6	Display LLDP information	23-292
Chapter 22	ERRP Command Configuration	24-294
22.1	Brief introduction of ERRP	24-294
22.2	ERRP Overview	24-294
22.3	Basic Concepts in ERRP	24-294
22.3.1	ERRP domain	24-294
22.3.2	ERRP ring	24-295
22.3.3	Control VLAN and data VLAN	24-295
22.3.4	Node	24-295

22.3.5	Primary port and secondary port	24-295
22.3.6	Common port and edge port.....	24-296
22.3.7	Multi-domain intersection common port.....	24-296
22.3.8	Timers	24-296
22.3.9	ERRP Packets	24-296
22.4	Typical ERRP Networking	24-297
22.4.1	Single ring	24-297
22.4.2	Multi-domain tangent rings	24-298
22.4.3	Single-domain intersecting rings.....	24-299
22.4.4	Dual homed rings.....	24-299
22.4.5	Multi-domain intersecting rings.....	24-300
22.5	How ERRP Works.....	24-300
22.5.1	Polling mechanism.....	24-300
22.5.2	Link down alarm mechanism	24-300
22.5.3	Ring recovery.....	24-301
22.5.4	Broadcast storm suppression mechanism in a multi-homed subring in case of primary ring link failure.....	24-301
22.5.5	Protocols and Standards.....	24-301
22.6	ERRP Configuration.....	24-301
22.6.1	ERRP Configuration list	24-301
22.6.2	ERRP configuration	24-301
22.6.3	Configure ERRP timer	24-302
22.6.4	Enter ERRP configuration mode.....	24-302
22.6.5	Configure control-vlan of ERRP domain.....	24-302
22.6.6	Create ERRP ring	24-303
22.6.7	Enable/disable ERRP ring	24-303
22.6.8	Display ERRP domain and ring information	24-303
Chapter 23	PPPoE Plus Configuration	25-304
23.1	Brief Introduction of PPPoE Plus	25-304
23.2	PPPoE Plus Configuration.....	25-304
23.2.1	PPPoE Plus Configuration list	25-304
23.2.2	Enable/disable PPPoE Plus.....	25-304
23.2.3	Configure PPPoE Plus type.....	25-305
Chapter 24	CFM Configuration	26-306
24.1	Brief introduction of CFM.....	26-306
24.2	Connectivity fault management overview.....	26-306
24.3	Basic Concepts in Connectivity Fault Detection	26-306
24.3.1	Maintenance domain.....	26-306
24.3.2	Maintenance association	26-306
24.3.3	Maintenance point.....	26-306
24.3.4	Basic Functions of Connectivity Fault Management	26-307
24.3.5	Protocols and Standards.....	26-308
24.4	CFM Configuration.....	26-308
24.4.1	CFM Configuration list	26-308
24.4.2	Configure cfm domain.....	26-308
24.4.3	Configure cfm mep level	26-309
24.4.4	Configure cfm mip level	26-309
24.4.5	Configure remote cfm rmep level	26-309
24.4.6	Configure cfm cc interval	26-310
24.4.7	Enable/disable VLAN sending cfm cc enable level	26-310
24.4.8	cfm ping.....	26-310
24.4.9	cfm traceroute	26-311

24.4.10	Display cfm domain.....	26-311
24.4.11	Display cfm maintenance-points local	26-311
24.4.12	Display cfm maintenance-points remote	26-312
24.4.13	Display cfm cc database	26-312
24.4.14	Display cfm errors	26-312

Chapter 1 Accessing Switch

This chapter is the basic knowledge for system management, including :

- 1) Command line interface
- 2) Command syntax comprehension
- 3) Syntax help
- 4) History command
- 5) Symbols in command
- 6) Parameter in command
- 7) User management
- 8) Ways for switch management

1.1 Command Line Interface

System provides a series of configuration command and command line interface. User can configure and manage switch by command line. Command line interface has the features as following :

- 1) Local configuration by Console interface
- 2) Local or remote configuration by TelNet
- 3) Configure command classification protection to guarantee unauthorized user illegal accessing.
- 4) Input “?” at any moment to obtain help information
- 5) Provide such network test command as ping to diagnose network fault
- 6) Provide FTP, TFTP, Xmodem to download and upload files
- 7) Keywords partial matching searching is adopted by command line convertor for user to input non-conflicting key words, such as : interface command can only input “interf”

1.1.1 Command Line Configuration Mode

System command line adopts classification protection to prevent illegal accessing of unauthorized user. Each command mode is for different configuration with the connection and distinction. For example, after successful accessing, user of all level can enter common user mode which can only see the system operation information; administrator can input “enable” to enter privileged mode; input “configure terminal” to enter global configuration mode from privileged mode which can enter related configuration mode according to inputting different configuration command. For example :

Command line provides command mode as following :

- 1) User mode
- 2) Privileged mode
- 3) Global configuration mode
- 4) Interface configuration mode
- 5) VLAN configuration mode
- 6) AAA configuration mode

- 7) RADIUS configuration mode
- 8) Domain configuration mode

The function and details of each command mode are as following :

Command Line Configuration Mode

Command line mode	Function	Prompt character	Command for entering	Command for exiting
User mode	See switch operation information	QTECH>	Connect with switch after inputting user name and password	exit disconnect with switch
Privileged mode	See switch operation information and manage system	QTECH#	Input enable in user mode	exit return to user mode quit disconnect with switch
Global configuration mode	Configure global parameter	QTECH(config)#	Input configure terminal in privileged mode	exit, end return to privileged mode quit disconnect with switch
Interface configuration mode	Configure interface parameter	QTECH(config-if-ethernet-0/1)#	Input "interface Ethernet 0/1" in global configuration mode, interface configuration can enter other interface mode and VLAN configuration mode without inputting "exit".	end return to privileged mode exit return to global configuration mode quit disconnect with switch
VLAN configuration mode	Configure VLAN parameter	QTECH(config-if-vlan)#	Input " vlan 2 " in global configuration mode, VLAN configuration mode can enter other VLAN mode and interface configuration mode without inputting "exit".	
AAA configuration mode	Create domain	QTECH(config-aaa)#	Input "aaa" in global configuration mode	
RADIUS configuration mode	Configure RADIUS server parameter	QTECH(config-radius-default)#	Input "radius host default" in global configuration mode	end return to privileged mode exit return to AAA configuration mode quit disconnect with switch
Domain configuration mode	Configure domain parameter	QTECH(config-aaa-test.com)#	Input "domain test.com" in AAA configuration mode	
VLAN Interface mode	Configure VLAN L3 interface	QTECH(config-if-vlan-interface-22)#	Input "interface vlan-interface 22" in global configuration mode	end return to privileged mode exit return to global configuration mode quit disconnect with switch
SuperVLAN Interface mode	Configure SuperVLAN L3 interface	QTECH(config-if-super-vlan-interface-1)#	Input "interface supervlan-interface 1" in global configuration mode	end return to privileged mode exit return to global configuration mode quit disconnect with switch
RIP configuration mode	Configure RIP parameter	QTECH(config-router-rip)#	Input "route rip" in global configuration mode	end return to privileged mode exit return to global configuration mode quit disconnect with switch
OSPF configuration mode	Configure OSPF parameter	QTECH(config-router-ospf)#	Input "route ospf" in global configuration mode	end return to privileged mode exit return to global configuration mode quit disconnect with switch

PIM configuration mode	Configure PIM parameter	QTECH(config-router-pim#	Input "pim" in global configuration mode	end return to privileged mode exit return to global configuration mode quit disconnect with switch
------------------------	-------------------------	--------------------------	--	---

1.1.2 Command Syntax Comprehension

This chapter describes the steps needed for command configuration. Please read this section and related detail information of command line interface in the following sections carefully.

The logging in identity verification of the system console of this switch is used to verify the identity of the operating user. It permits and refuses the logging in by matching recognizing user name and password.

Step 1. Following are showed when entering command line interface,

Username(1-32 chars) :

Please input user name, press Enter button, and then the prompt is as following :

Password (1-16 chars) :

Input password. If it is correct, enter the user mode with the following prompt :

QTECH>

 Note : Defaulted login and password is admin/123456.

In switch system, there are 2 different privileges. One is administrator, and the other is common user. Common user only can see the configuration information of switch without right to modify it but administrator can manage and configure the switch by specified command.

Logging in as administrator can enter privileged mode from user mode.

QTECH>enable

Step 2 : Input command

Skip to step 3, if the command needs input the parameter. Continue this step if the command need input the parameter.

If the command needs a parameter, please input it. When inputting a parameter, keyword is needed.

The parameter of the command is specified which is the number or character string or IP address in a certain range. Input "?" when you are uncomprehending, and input the correct keyword according to the prompt. Keyword is what is to be operated in command.

If more than one parameter are needed, please input keywords and each parameter in turn according to the prompt until "<enter>" is showed in prompt to press enter button.

Step 3 : Press enter button after inputting complete command.

For example :

! User need not input parameter

QTECH#quit

"quit" is a command without parameter. The name of the command is quit. Press enter button after inputting it to execute this command.

! User need input parameter

QTECH(config)#vlan 3

"vlan 3" is a command with parameter and keyword, vlan of which is command keyword and 3 of which is parameter.

1.1.3 Syntax Help

There is built-in syntax help in command line interface. If you are not sure about the syntax of some command, obtain all command and its simple description of the current mode by inputting "?" or help command; list all keywords beginning with the current character string by inputting "?" closely after the command character string; input "?" after space, if "?" is in the same location of the keyword, all keywords and its simple description

will be listed, if “?” is in the same location of parameter, all the parameter description will be listed, and you can continue to input command according to the prompt until the prompt command is “<enter>” to press enter button to execute command.

For example :

Directly input “?” in privileged mode

QTECH#?

System mode commands :

cls clear screen

help description of the interactive help

ping ping command

quit disconnect from switch and quit

.....

Input “?” closely after keyword

QTECH(config)#interf?

interface

Input “?” after command character string and space

QTECH(config)#spanning-tree ?

forward-time config switch delaytime

hello-time config switch hellotime

max-age config switch max agingtime

priority config switch priority

<enter> The command end.

- Parameter range and form

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

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

- Command line end prompt

QTECH(config)#spanning-tree ?

<enter> The command end.

1.1.4 History command

Command line interface will save history command inputted by user automatically so that user can invoke history command saved by command line interface and re-execute it. At most 100 history commands can be saved by command line interface for each user. Input “Ctrl+P” to access last command, and “Ctrl+N” for next command.

1.1.5 Symbols in command

There are all kinds of symbols in command syntax which is not a part of command but used to describe how to input this command. Table 1-2 makes a brief description of these symbols.

1.2 Command Symbols Description

Command Symbols Description

Symbol	Description
Vertical bars	Vertical bars () means coordinate, together using with braces ({ }) and square brackets ([]).
Square brackets []	Square brackets ([]) mean optional elements. For example : show vlan [vlan-id]
Braces { }	Braces ({ }) group required choices, and vertical bars () separate the alternative elements. Braces and vertical bars within square brackets ([{ }]) mean a required choice within an optional element.

1.2.1 Command Parameter Categories

There are 5 categories command parameter as following :

- Scale

Two numerical value linked by hyphen in angle brackets (< >) means this parameter is some number in the range of those two numbers.

For example :

INTEGER<1-10> means user can input any integer between 1 and 10 (include 1 and 10), such as 8 is a valid number.

- IP address

The prompt which is in the form of A.B.C.D. means the parameter is an IP address. A valid IP address is needed to input.

For example :

192.168.0.100 is a valid IP address.

- MAC address

The prompt which is in the form of H : H : H : H : H means the parameter is a MAC address. A valid MAC address is needed to input. If a multicast MAC address is needed, there will be related prompt.

For example :

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

- Interface list

The prompt of interface list is STRING<3-4>. Interface parameter interface-num is in the form of interface-type + interface-number. Interface-type is Ethernet and interface-number is slot-num/port-num, in which slot-num is in the range of 0 to 2, and port-num is in the range of 1 to 24. Seriate interfaces with the same type can be linked by to keyword, but the port number to the right of the to keyword must be larger than the one to the left of the keyword, and this argument only can be repeated for 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 displaying spanning-tree information of interface ethernet 0/0/1 ethernet 0/0/3 to ethernet 0/0/5

- Character string

The prompt which is in the form of STRING<3-4> means the parameter is a character string which is in the form of 1 to 19 characters. "?" can be inputted to display the concrete command description.

1.3 User management

There are 2 privileges for user :

- 1) administrator
- 2) normal user

Normal user can only enter user mode not privileged mode after logging in, so that he can only see system information but not to configure it. Administrator has the right to enter all modes, and query and configure all parameters.

1.3.1 System default user name

There is a system default built-in user name called **admin**, and the initial password is **123456**. It is suggested modifying password when logging in switch for the first time to avoid leaking it. This user name cannot be deleted and the privilege cannot be modified either. It also possesses the right to manage other users. Please remember your modified password.

1.3.2 Add user

Log in with the identity of system administrator admin to enter privileged mode, then global configuration mode by using username command. Input user name, user's privilege, password to add new user according to system prompt or by using the following command.

username *username* [**privilege level**] { **password encryption-type password** }

username : User name of new users and existed users ranges from 1 to 32 printable characters excluding such wildcards as '/', ':', '*', '?', '\\', '<', '>', '|', '"' etc.

privilege : Privilege of new user ranges from 0 to 15. 0 to 1 means user while 2 to 15 means administrator.

encryption-type : the value of it is 0 or 7. 0 means non-encryption and 7 means encryption (It is not supported now).

password : Log in password for new user and modified password of the existed user ranges from 1 to 16 characters or numbers.

If the privilege doesn't configure, the default privilege is ordinary user. At most 8 users are supported.

Caution : User name supports case insensitivity while password doesn't support case sensitivity.

! Add a new administrator "red", configure privilege to be 3, and password to be 1234

```
QTECH(config)#username qtech privilege 3 password 0 1234
```

1.3.3 Modify password

In global configuration mode, system administrator admin can use the following command to modify password of his or other user. Other user can only modify his own password.

username change-password

For example :

! Modify the password of user "red" to be 123456

```
QTECH(config)#username change-password
```


```
please input you login password : *****
```

```
please input username : red
```

```
Please input user new password : *****
```

```
Please input user comfirm password : *****
```

change user qtech password success.

 **Caution** : For restoration default password of “admin” user, please refer to support@qtech.ru.

1.3.4 Modify privilege

In global configuration mode, only administrator admin can use following command to modify the privilege of other user.

username *username* [**privilege level**] { **password encryption-type password** }


username : User name of new users and existed users ranges from 1 to 32 printable characters excluding such wildcards as '/', ':', '*', '?', '\\', '<', '>', '|', '"' etc.

privilege : Privilege of new user or the modified privilege of existed user ranges from 0 to 15. 0 to 1 means user while 2 to 15 means administrator. **Caution** : the privilege of administrator cannot be modified.

encryption-type : the value of it is 0 or 7. 0 means non-encryption and 7 means encryption (It is not supported now).

password : Log in password for new user and modified password of the existed user ranges from 1 to 16 characters or numbers.

If inputting nothing to modify the privilege of existed user, the privilege doesn't modify.

 **Caution** : User name supports case insensitivity while password doesn't support case sensitivity.

For example :

! Modify the privilege of administrator “qtech” to be 1, and password to be 1234

```
QTECH(config)#username qtech privilege 1 password 0 1234
```

1.3.5 Remove user name

System administrator admin can use following command to remove user name in global configuration mode

no username *username*

Username is the user name to be deleted.

For example :

! Remove user qtech

```
QTECH(config)#no username qtech
```

1.3.6 View system user information

View user list, and input

show username

command or

show username [username]

command in any configuration mode to display information of all users.

For example :

! Display information of user qtech

```
QTECH(config)#show username qtech
```

display user information

user name role

1.4 Remote authentication of administrator

After authentication, user's default privilege is normal user. Only when there is Service-Type field in authentication accepting packet the value of which is Administrative, user's privilege is administrator.



Caution : Admin user only supports local database authentication.

1.4.1 Start RADIUS/TACACS+ remote authentication

Use following command in global configuration mode :

```
muser { local | { radius radiusname/tacacs+ tacacsname { pap | chap } [ local ] } }
```

It can be configured to authenticate only by RADIUS/TACACS+ remote authentication or by local database authentication after no response of RADIUS/TACACS+ server caused by failing connection.

1.4.2 Display authentication configuration

Use following command to display authentication configuration.

```
show muser
```

1.5 Ways of managing switch

System provides following ways of management :

- By hyper terminal accessing command-line interface(CLI)
- By telnet/ssh managing system
- By SNMP managing software management system
- By Web browser such as Internet Explorer managing system

1.5.1 Manage switch by hyper terminal

Use hyper terminal (or simulation terminal software) connect to Console to access system command line interface (CLI) by hyper terminal.

Configuration : Open "file" -> "attribute" menu, popping up a window. Enter configuration to restore it to default value, and click "setting" and then choose "auto-detect" in the pulldown list of "terminal simulation" and click [ok]. After the successful connection and seeing logging in interface of operation system in terminal, configure switch by command line interface. The steps are as following :

Step 1 : Connect switch Console with computer serial port;

Step 2 : After the switch power on and system successful booting, logging in prompt can be seen :

Username(1-32 chars) :

Step 3 : Input correct user name, press enter button, then input corresponding password. If it is the first time to logging in switch, use default user name admin and its password 123456 to log in and operate as system administrator. If your own user name and password exist, log in with your own user name and password;

Step 4 : After successfully logging in, following information is displayed :

```
QTECH>
```

Step 5 : As administrator, after entering privileged mode, use copy running-config startup-config command to save configuration.

```
QTECH#copy running-config startup-config
```

When following information is displayed :

```
Startup config in flash will be updated, are you sure(y/n)? [n]y
```

```
Building, please wait...
```

It means system is saving configuration. Please wait, then the prompt is :

```
Build successfully.
```

It means current configuration is saved successfully.

Following information is displayed when system booting :

```
Ready to load startup-config, press ENTER to run or CTRL+C to cancel :
```

Press enter button to make saved configuration be effective, and press CTRL+C to restore system default configuration.

Step 6 : Administrator can use stop connection when overtime, while normal user can use this function in user mode. Input timeout command to configure the overtime of user's logging in to be 20 minutes. And use no timeout command to configure overtime to be non-over timing.

Step 7 : Input following command after finishing operation to switch :

```
QTECH#quit
```

It is used to exit user interface.

1.5.2 Manage switch by telnet

Step 1 : Establish configuration environment by connecting computer by network to switch interface;

Step 2 : Run Telnet program in computer;

Step 3 : After switch is power on, input switch IP address to connect to switch, and input configured logging in password according to the prompt, then the command line prompt is displayed (such as QTECH>). It will be disconnected after 1 minute when there is not any input before successfully logging in or wrong inputting of user name and password for 5 times. If there is such prompt as "Sorry, session limit reached.", please connect later (At most 5 telnet users are allowed to log in at the same time.);

Step 4 : Use related command to configure switch system parameter or view switch operation. If you want to enter privileged mode, user must possess the privilege of administrator. If you need any help, please input "?" at any moment. For concrete command, please refer to following chapters.

Step 5 : If you want to exit telnet, use quit or exit command to exit in user mode, and quit command to exit in other mode. Administrator can use stop username command in privileged mode to exit logging in.

1.6 Brief introduction of SSH

SSH is short for Secure Shell. Users can access to the device via standard SSH client, and sent up safe connection with device. The Data that transmitted via SSH connection are encrypt, which assure the transmitted sensitive data, management data and configuration data, such as password, between the users and devices will not be wiretapped or acquired illegally by the third party.

SSH can replace Telnet, providing users with means of safely management and device configuration.

1.7 SSH Configuration list

The configuration task list of SSH is as follows :

- 1) Enable/disable SSH function of the device
- 2) SSH secret key configuration
- 3) Others

1.7.1 Enable/disable SSH function of the device

Enable/disable SSH function of the device in global mode, users can not access to the devices via SSH client when SSH function is closed. To access to the device via SSH client, users need to configure correct secret key and upload the secret key in the device besides opening up the SSH function.

Configuration command is as following :

ssh

no ssh

Example :

! Enable SSH

QTECH(config)#ssh

1.7.2 SSH key configuration

Use SSH secret key in privileged mode. User cannot use SSH client to log in if there is no secret key or the key is incorrect or the key is not load. In order to log in by SSH client, configure correct key and load it with SSH enabling.

The configured secret key should be RSA. There are two kinds of keys : public and private. It can use the default key and also can download keyfile to device by tftp and ftp. Configured key can be used after loading. Configured key is stored in Flash storage which will be load when system booting. It also can load the key stored in Flash storage by command line when system booting.

If configured key is not RSA key or public and private key are not matched, user cannot log in by SSH.

Keyfile contains explanation and key explain line and the key. Explain line must contain “ : ” or space. Key contains the key coded by Base64, excluding “ : ” and space. Private keyfile cannot contain public key. Private keyfile cannot use password to encrypt.

1.7.2.1 Configure default key.

The command is as following :

crypto key generate rsa

Example :

! Configure SSH key to be default key

QTECH#crypto key generate rsa

1.7.2.2 Download or upload key by tftp or ftp.

The command is as following :

load keyfile { public | private } tftp *server-ip filename*

load keyfile { public | private } ftp *server-ip filename username passwd*

upload keyfile { public | private } tftp *server-ip filename*

upload keyfile { public | private } ftp *server-ip filename username passwd*

Example :

! Download keyfile pub.txt from tftp server 1.1.1.1 to be SSH public key

```
QTECH#load keyfile public tftp 1.1.1.1 pub.txt
```

1.7.2.3 Clear configured key.

This command will clear all keyfiles stored in Flash storage. The configuration command is as following :

crypto key zeroize rsa

Example :

! Clear configured SSH key

```
QTECH#crypto key zeroize rsa
```

1.7.2.4 Load new key.

After configuring new SSH key, it restored in Flash storage without loading. This command can read configured key from Flash storage and update the current key. When system booting, it will detect Flash storage, if SSH key is configured, it will load automatically. The configuration command is as following :

crypto key refresh

Example :

! Load new SSH key :

```
QTECH#crypto key refresh
```

1.7.3 Others

Use following command to display SSH configuration

show ssh

This command is used to display SSH version number, enabling/disabling SSH and SSH keyfile. The SSH keyfile is “available” when the key is configured and loaded.

Use following command to display configured keyfile

show keyfile { public | private }

Use following command to display logged in SSH client

show users

This command is used to display all logged in Telnet and SSH client.

Use following command to force logged in SSH client to stop

stop username

This command can force logged in SSH client to stop. Username is the logged in user name.

It allows at most 5 SSH clients to logged in. If Telnet client has logged in, the total number of SSH and Telnet clients is no more than 5. For example, if there are 2 Telnet clients in device, at most 3 SSH clients can log in.

Chapter 2 Switch Manage and Maintenance

2.1 System IP configuration

IP address means a unique address of 32 bits which is distributed to host in Internet. IP address consists of network number and host number. The structure of IP address can make us easy to address in Internet.

2.2 Configure manage IP interface

It must be existed VLAN or SuperVLAN IP interface. For more details please refer to [Create VLAN interface for normal VLAN](#)

2.3 Configuration ip address by manual operation

Use `ipaddress` command in vlan interface configuration mode to configuration ip address and netmask by manual operation :

ip address *ip-address mask*

ip-address means system ip address. *Mask* means netmask.

For example :

! Configure IP address of VLAN 100 to be 192.168.0.100, netmask to be 255.255.0.0.

```
QTECH(config-if-vlan)#ip address 192.168.0.100 255.255.0.0.
```

2.4 Configuration Files Management

2.4.1 Edit configuration files

Configuration files adopts text formatting which can be upload to PC from devices by FTP and TFTP protocol. Use text edit tool (such as windows notepad) to edit uploaded configuration files.

System is defaulted to execute configuration files in global configuration mode, so there are two initial commands : “enable”, and “configure terminal”. There is entering symbol after each command.

2.4.2 Modify and save current configuration

User can modify and save system current configuration by command line interface to make current configuration be initial configuration of system next booting.

copy running-config startup-config

This command is needed to save current configuration. When executing configuration files, if there is un-executed command, it will be displayed as “[Line : xxxx]invalid : commandString”. If there is command with executing failure, it will be displayed as “[Line : xxxx]failed : commandString”. If there is a command beyond 512 characters, it will be displayed as “[Line : xxxx]failed : too long command : commandString”, and only first 16 characters of

this command will be displayed, and end up with ..., in which "xxxx" means the line number of the command, and commandString means command character string. Un-executive command includes command with grammar fault and un-matching pattern. Use following command in privileged mode.

```
QTECH#copy running-config startup-config
```

2.4.3 Erase saved configuration

Use **clear startup-config** command to clear saved configuration. After using this command to clear saved configuration and reboot switch. The switch will restore to original configuration. Use this command in privileged mode.

```
QTECH#clear startup-config
```

2.4.4 Execute saved configuration

User can restore saved configuration by commang line interface by using that command in privileged mode to execute saved configuration :

```
copy startup-config running-config
```

2.4.5 Display saved configuration

User can display syatem saved configuration information in the form of text by command line interface. Use following command to display system saved configuration :

```
show startup-config [ module-list ]
```

module-list : Optional module. If the module name is unoptioned, all information of configuration files will be displayed. If choose one or same of the modules, the specified information will be displayed. This command can be used in any configuration mode.

For example :

```
! Display all saved configuration
```

```
QTECH#show running-config
```

```
! Display saved configuration of GARP and OAM module
```

```
QTECH#show running-config garp oam
```

2.4.6 Display current configuration

User can display syatem current configuration information in the form of text by command line interface. Use following command to display system current configuration :

```
show running-config [ module-list ]
```

module-list : Optional module. If the module name is unoptioned, all information of configuration files will be displayed. If choose one or same of the modules, the specified information will be displayed.

For example :

```
! Display all configurations
```

```
QTECH#show running-config
```

```
! Display configuration of GARP and OAM module
```

```
QTECH#show running-config garp oam
```

2.4.7 Configure file executing mode shift

User can change executing mode of configuration file by command line interface. System saved configuration files can be executed in stop and continue mode. When coming across errors, the executing will not stop; it will display errors and continue executing. It is defaulted to be non-stop mode. Use **buildrun mode stop** to configure executing mode to be stopped. Use **buildrun mode continue** command to configure buildrun mode to be continue. Use these commands in privileged mode.

For example :

! Configure buildrun mode to be stop.

QTECH#buildrun mode stop

! Configure buildrun mode to be continue

QTECH#buildrun mode continue

2.5 Online Loading Upgrade Program

System can upgrade application program and load configuration files on line by TFTP, FTP, Xmodem, and can upload configuration files, logging files, alarm information by TFTP and FTP.

2.5.1 Upload and download files by TFTP

Use following command to upload files by TFTP :

upload { alarm | configuration | logging } **tftp** *tftpserver-ip filename*

Use following command to download files by TFTP :

load { application | configuration | whole-bootrom } **tftp** *tftpserver-ip filename*

tftpserver-ip is the IP address of TFTP server. *filename* is the file name to be loaded which cannot be system key words (such as con cannot be file name in windows operation system). Open TFTP server and set file upload path before use this command.

Suppose IP address of TFTP server is 192.168.0.100, file name is abc. Open TFTP server to configure upload and download path in privileged mode.

For example :

! Upload configuration to 192.168.0.100 by FTP and saved as abc

QTECH#upload configuration ftp 192.168.0.100 abc username password

Configuration information saved when uploading is successful.

! Download configuration program abc to 192.168.0.100 by TFTP

QTECH#load configuration ftp 192.168.0.100 abc

Reboot the switch after successful download and run new configuration program.

! Upload alarm to 192.168.0.100 by TFTP and saved as abc

QTECH#upload alarm tftp 192.168.0.100 abc

! Upload logging to 192.168.0.100 by TFTP and saved as abc

QTECH#upload logging tftp 192.168.0.100 abc

! Download application program app.arj to 192.168.0.100 by TFTP

QTECH#load application tftp 192.168.0.100 app.arj

Reboot the switch after successful download and run new application program.

! Download whole-bootrom abc to 192.168.0.100 by TFTP

QTECH#load whole-bootrom tftp 192.168.0.100 rom3x26.bin

2.5.2 Upload and download files by FTP

Use following command to upload files by FTP :

upload { alarm | configuration | logging } **ftp** *ftpserver-ip filename username userpassword*

Use following command to download files by FTP :

load { application | configuration | whole-bootrom } **ftp** *ftpserver-ip filename username userpassword*

ftpserver-ip is the IP address of FTP server. *Filename* is the file name to be loaded which cannot be system key words (such as con cannot be file name in windows operation system). Open FTP server and set username, password and file upload path before use this command.

Suppose IP address of TFTP server is 192.168.0.100, file name is abc. Open TFTP server to configure username to be user, password to be 1234 and file download path in privileged mode.

For example :

! Upload configuration to 192.168.0.100 by FTP and saved as abc

```
QTECH#upload configuration ftp 192.168.0.100 abc user 1234
```

Configuration information saved when uploading is successful.

! Download configuration program abc to 192.168.0.100 by FTP

```
QTECH#load configuration ftp 192.168.0.100 abc user 1234
```

Reboot the switch after successful download and run new configuration program.

! Download application program abc to 192.168.0.100 by FTP

```
QTECH#load application ftp 192.168.0.100 abc user 1234
```

Reboot the switch after successful download and run new application program.

! Upload alarm to 192.168.0.100 by FTP and saved as abc

```
QTECH#upload alarm ftp 192.168.0.100 abc user 1234
```

! Upload logging to 192.168.0.100 by FTP and saved as abc

```
QTECH#upload logging ftp 192.168.0.100 abc user 1234
```

! Download whole-bootrom abc to 192.168.0.100 by FTP

```
QTECH#load whole-bootrom ftp 192.168.0.100 abc user 1234
```

2.5.3 Download files by Xmodem

Use load application xmodem command to load application program by Xmodem protocol.

load application xmodem

Input following command in privileged mode :

```
QTECH#load application xmodem
```

Choose “send” -> “send file” in super terminal, and input full path and filename of the file in filename dialog box, and choose Xmodem protocol in “protocol”, then click **【send】** .

Reboot the switch after successful download and run new application program.

Use load configuration xmodem command to load configuration program by Xmodem protocol.

load configuration xmodem

Input following command in privileged mode :

```
QTECH#load configuration xmodem
```

Choose “send” -> “send file” in super terminal, and input full path and filename of the file in filename dialog box, and choose Xmodem protocol in “protocol”, then click **【send】** .

Reboot the switch after successful download and run new application program.

Use load whole-bootrom xmodem command to load whole bootrom by xmodem protocol.

load whole-bootrom xmodem

Input following command in privileged mode :

```
QTECH#load whole-bootrom xmodem
```

Choose “send” -> “send file” in super terminal, and input full path and filename of the file in filename dialog box, and choose Xmodem protocol in “protocol”, then click **【send】** .

Reboot the switch after successful download and run new BootRom program.

2.6 Reboot

Use the command in privileged mode to reboot switch :

reboot

2.7 System Maintenance

Use show command to check system information. Show command can be divided into following categories :

- 4) Command of displaying system configuration
- 5) Command of displaying system operation
- 6) Command of displaying system statistics

Show command related to all protocols and interfaces refers to related chapters. Followings are system show commands.

Use following commands in any configuration mode :

show version	Display system version
show username	Display administrator can be logged in
show users	Display administrators logged in
show system	Display system information
show memory	Display memory
show clock	Display system clock
show cpu	Display cpu information

For example :

! Display system version

```
QTECH(config)#sh ver
```

```
software platform      : Broadband NetWork Platform Software
software version      : QTECH QSW-3900 V100R001B01D003P001SP9
copyright             : Copyright (c) 2001-2009
compiled time         : Jul 16 2009 10 : 10 : 00
processor              : PPC 8245, 400MHz
SDRAM (bytes)         : 128M
flash memory (bytes)  : 8192k
MAC address           : 00 : 1f : ce : 11 : 87 : 6f
product serial number : 0105000508060200000430
hardware version      : V2.0
bootrom version       : V1.32
```

2.7.1 Basic Configuration and Management

System basic configuration and management includes :

2.7.1.1 Configure host name

Use hostname command in global configuration mode to configure system command line interface prompt.
Use no hostname command to restore default host name.

Configure system command line interface prompt.

hostname *hostname*

hostname : character strings range from 1 to 32, these strings can be printable, excluding such wildcards as '/', ':', '*', '?', '\\', '<', '>', '|', ""etc.

Use no hostname command in global configuration mode to restore default host name to be QTECH.

For example :

! Configure hostname to be QSW-3900

```
QTECH(config)#hostname QSW-3900
```

```
QSW-3900(config)#
```

2.7.1.2 Configure system clock

Use clock set command in privileged mode to configure system clock.

clock set *HH : MM : SS YYYY/MM/DD*

For example :

! Configure system clock to be 2001/01/01 0 : 0 : 0

```
QTECH#clock set 0 : 0 : 0 2001/01/01
```

2.7.2 Network connecting test command

Use ping command in privileged mode or user mode to check the network connection.

ping [-c *count*] [-s *packetsize*] [-t *timeout*] *host*

Parameter :

-c count : The number of packet sending.

-s packetsize : The length of packet sending, with the unit of second

-t timeout : the time of waiting for replying after packet is sent, with the unit of second

For example :

! Ping 192.168.0.100

```
QTECH#ping 192.168.0.100
```

```
PING 192.168.0.100 : with 32 bytes of data :
```

```
reply from 192.168.0.100 : bytes=32 time<10ms TTL=127
```

```
reply from 192.168.0.100 : bytes=32 time<10ms TTL=127
```

```
reply from 192.168.0.100 : bytes=32 time<10ms TTL=127
```

```
reply from 192.168.0.100 : bytes=32 time<10ms TTL=127
```

```
reply from 192.168.0.100 : bytes=32 time<10ms TTL=127
```

```
----192.168.0.100 PING Statistics----
```

5 packets transmitted, 5 packets received, 0% packet loss

round-trip (ms) min/avg/max = 0/0/0

2.7.3 Loopback test command

In global configuration mode, loopback command is used to test exterior of all interfaces; in interface configuration mode, loopback command is used to test whether the interface is normal, and it can be divided into interior and exterior. When exterior testing, exterior wire must be inserted (receiving and sending lines of RJ 45 connected directly). Use 4 different wires when the speed is less than 100M.

Using loopback command to do the loopback test, interface cannot transmit data packet correctly, and it will be automatically ended after a certain time. If shutdown command is executed, loopback test fails; when loopback test is executing, speed, duplex, mdi, vct and shutdown operations are forbidden. After exterior test, pull out the exterior wire to avoid abnormal communication.

Loopback on all interfaces :

loopback { internal | external }

Loopback on specified interface :

loopback { external | internal }

External means external loopback and internal means internal loopback

For example :

! Loopback on interface Ethernet 0/0/1

QTECH(config-if-ethernet-0/0/1)#loopback external

! Loopback on all interfaces

QTECH(config)#loopback internal

2.7.4 Remote access restriction

You can restrict host IP address or some network interface of switch by restricting web, telnet and snmp agent, but other IP address without configuration cannot manage switch. By default, three servers possess an address interface of 0.0.0.0, so users of any IP address can manage switch. Different IP address and mask mean different information. The mask in reverse which is 0.0.0.0 means host address, or it means network interface. 255.255.255.255 means all hosts. When enabling a configuration, an item of 0.0.0.0 must be deleted. When receiving a packet, judge the IP address whether it is in the range of managed IP address. If it does not belong to it, drop the packet and shutdown telnet connection.

login-access-list { web | snmp | telnet | telnet-limit } *ip-address wildcard*

Web means accessing IP address restriction of web server; snmp means accessing IP address restriction of snmp agent; telnet means accessing IP address restriction of telnet; ipaddress means IP address; wildcard means mask wildcard which is in the form of mask in reverse. 0 means mask this bit, and 1 means does not mask this bit. When mask in reserve is 0.0.0.0, it means host address, and 255.255.255.255 means all hosts. Use the no command to delete corresponding item.

For example :

! Configure ip address allowed by telnet management system to be 192.168.0.0/255.255.0.0

QTECH(config)#login-access-list telnet 192.168.0.0 0.0.255.255

QTECH(config)#no login-access-list telnet 0.0.0.0 255.255.255.255

Use show login-access-list command to display all ip address allowed by web, snmp, telnet management system.

show login-access-list

2.7.5 The number of Telnet user restriction

Configure the max number of Telnet users. This function can restrict the number of Telnet user (0-5) to enter privileged mode at the same time. The user logged in without entering privileged mode will not be restricted but restricts by the max number. Administrator and super user will not be restricted and can be logged in through series interface. Display the configuration by show users command.

Configure it in global configuration mode :

login-access-list telnet-limit *limit-no*

no login-access-list telnet-limit

Example :

! Configure only 2 Telnet users can enter privileged mode

```
QTECH(config)#login-access-list telnet-limit 2
```

2.7.6 Routing tracer command

Tracert is used for routing detecting and network examination. Configure it in privileged mode :

tracert [**-u** | **-c**] [**-p** *udpport* | **-f** *first_ttl* | **-h** *maximum_hops* | **-w** *time_out*] *target_name*

Parameter :

-u means sending udp packet,

-c means sending echo packet of icmp. It is defaulted to be -c;

-p *udpport* : destination interface address for sending udp packet which is in the range of 1 to 65535 and defaulted to be 62929;

-f *first_ttl* : initial ttl of sending packet which is in the range of 1 to 255 and defaulted to be 1;

-h *maximum_hops* : the max ttl of sending packet which is in the range of 1 to 255 and defaulted to be 30;

-w *time_out* : the overtime of waiting for the response which is in the range of 10 to 60 with the unit of second and default to be 10 seconds;

target_name : destination host or router address

Example :

! Tracert 192.168.1.2

```
QTECH#tracert 192.168.1.2
```

```
Tracing route to 192.168.1.2 [192.168.1.2]
```

```
over a maximum of 30 hops :
```

```
  1    20 ms  <10 ms <10 ms    192.168.0.1
```

```
  1    20 ms  <10 ms  30 ms    192.168.1.2
```

```
tracert complete.
```

2.7.7 Packets rate limit to CPU

Command `cpu-car` is used to configure cpu rate for receiving packet. This packets can be like IGMP, BPDU, DHCP etc. Configure it in global configuration mode :

cpu-car *target-rate*

no cpu-car is used to restore to default cpu rate for receiving packet.

Parameter :

target-rate : cpu rate for receiving packet , which is in the range of 1 to 1000pps and the default rate is 50pps..

Example :

! Configure cpu rate for receiving packet to be 100pps

```
QTECH(config)#cpu-car 100
```

2.8 Monitor system by SNMP

2.8.1 Brief introduction of SNMP

SNMP(Simple Network Management Protocol)is an important network management protocol in TCP/IP network. It realizes network management by exchanging information packets. SNMP protocol provides possibility of concentrated management to large sized network. Its aim is guaranteeing packet transmission between any two points to be convenient for network administrator to search information, modify and search fault, finish fault diagnosing, capacity planning and creation reporting at any network node. It consists of NMS and Agent. NMS (Network Management Station), is the working station of client program running, and Agent is server software running in network devices. NMS can send GetRequest, GetNextRequest and SetRequest packet to Agent. After receiving requirement packet of NMS, Agent will Read or Write management variable according to packet type and create Response packet, and return it to NMS. On the other hand, the Trap packet of abnormality of cold boot or hot boot of devices will send to NMS.

QTECH company is present it own QTECH NMS and Agent server. Please refer to the [http :
//www.QTECH.ru/support/software.htm](http://www.QTECH.ru/support/software.htm)

System supports SNMP version of v1, v2c and v3. v1 provides simple authentication mechanism which does not support the communication between administrator to administrator and v1 Trap does not possess authentication mechanism. V2c strengthens management model (security), manages information structure, protocol operation, the communications between managers, and it can create and delete table, and strengthen communication capacity of managers, and reduce the storage operation of agency. V3 realizes user distinguishing mechanism and packet encryption mechanism, and greatly improves security of SNMP protocol.

Simple Network Management Protocol (SNMP) offers a framework to monitor network devices through TCP/IP protocol suite. It provides a set of basic operations in monitoring and maintaining the Internet and has the following characteristics :

- Automatic network management : SNMP enables network administrators to search information, modify information, find and diagnose network problems, plan for network growth, and generate reports on network nodes.
- SNMP shields the physical differences between various devices and thus realizes automatic management of products from different manufacturers. Offering only the basic set of functions, SNMP makes the management tasks independent of both the physical features of the managed devices and the underlying networking technology. Thus, SNMP achieves effective management of devices from different manufactures, especially so in small, fast and low cost network environments.

2.9 SNMP Mechanism

An SNMP enabled network is comprised of network management station (NMS) and Agent.

NMS is a station that runs the SNMP client software. It offers a user friendly human computer interface, making it easier for network administrators to perform most network management tasks. Currently, the most commonly used NMSs include Quidview, Sun NetManager, and IBM NetView.

Agent is a program on the device. It receives and handles requests sent from the NMS. Only under certain circumstances, such as interface state change, will the Agent inform the NMS.

NMS manages an SNMP enabled network, whereas Agent is the managed network device. They exchange

management information through the SNMP protocol.

SNMP provides the following four basic operations :

Get operation : NMS gets the value of a certain variable of Agent through this operation.

Set operation : NMS can reconfigure certain values in the Agent MIB (Management Information Base) to make the Agent perform certain tasks by means of this operation.

Trap operation : Agent sends Trap information to the NMS through this operation.

Inform operation : NMS sends Trap information to other NMSs through this operation.

2.10 SNMP Protocol Version

Currently, SNMP agents support SNMPv3 and are compatible with SNMPv1 and SNMPv2c.

SNMPv1 and SNMPv2c authenticate by means of community name, which defines the relationship between an SNMP NMS and an SNMP Agent. SNMP packets with community names that did not pass the authentication on the device will simply be discarded. A community name performs a similar role as a key word and can be used to regulate access from NMS to Agent.

SNMPv3 offers an authentication that is implemented with a User-Based Security Model (USM for short), which could be authentication with privacy, authentication without privacy, or no authentication no privacy. USM regulates the access from NMS to Agent in a more efficient way.

2.11 MIB Overview

Management Information Base (MIB) is a collection of all the objects managed by NMS. It defines the set of characteristics associated with the managed objects, such as the object identifier (OID), access right and data type of the objects.

MIB stores data using a tree structure. The node of the tree is the managed object and can be uniquely identified by a path starting from the root node. As illustrated in the following figure, the managed object B can be uniquely identified by a string of numbers {1.2.1.1}. This string of numbers is the OID of the managed object B.

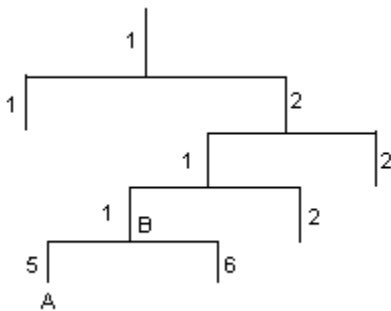


Figure 1 MIB tree

2.12 SNMP Configuration

SNMP configuration command list includes :

- 1) Configure community
- 2) Configure sysContact
- 3) Configure Trap destination host address
- 4) Configure sysLocation
- 5) Configure sysName

- 6) Configure notify
- 7) Configure engine id
- 8) Configure view
- 9) Configure group
- 10) Configure user
- 11) Configure community

SNMP adopts community authentication. The SNMP packets which are not matching the authenticated community name will be dropped. SNMP community name is a character string. Different community can possess the accessing right of read-only or read-write. Community with the right of read-only can only query system information, but the one with the right of read-write can configure system. System can configure at most 8 community names. It is defaulted to configure without community name. Configure it in global configuration mode.

2.12.1 Configure community name and accessing right.

This command can also be used to modify community attribution with character string community-name being the same.

snmp-server community *community-name* { ro | rw } { deny | permit } [**view** view-name]

community-name is a printable character string of 1 to 20 characters; ro|rw means read only or can be read and write; permit, deny means community can or cannot be activated;

View-name is view configured for community. The default configuration view is iso.

Delete community name and accessing right

no snmp-server community *community-name*

community-name is existed community name.

For example :

! Add community qtech, and configure privilege to be rw, and permit

```
QTECH(config)#snmp-server community qtech rw permit
```

! Remove community qtech

```
QTECH(config)#no snmp-server community qtech
```

Display community name in any mode

show snmp community

For example :

! Display SNMP community information

```
QTECH(config)#show snmp community
```

2.12.2 Configure sysContact

sysContact is a managing variable in system group in MIB , the content of which is the contact way of the administrator. Configure it in global configuration mode :

snmp-server contact *syscontact*

no snmp-server contact

syscontact : Contact way to administrator ranges from 1 to 255 printable characters. Use the no command to restore default way of contacting to administrator.

For example :

! Configure administrator contact way to be support@QTECH.ru

```
QTECH(config)#snmp-server contact support@QTECH.ru
```



Caution : Use quotation mark to quote space in character string.

Use show snmp contact command in any configuration mode to display how to contact administrator :

show snmp contact

For example :

! Display how to contact with administrator

```
QTECH(config)#show snmp contact
```

```
manager contact information : support@QTECH.ru
```

2.12.3 Configure Trap destination host address

Use this configuration to configure or delete IP address of destination host. Configure it in global configuration mode.

Configure notify destination host address

```
snmp-server host host-addr [version {1 | 2c | 3 [auth | noauth | priv]}] community-string [udp-port port] [ notify-type [ notifytype-list ] ]
```

Delete notify destination host address

```
no snmp-server host ip-address community-string { 1 | 2c | 3 }
```

ip-address and snmp-server means IP address in SNMP server notify sending list.

community-string means the security name IP corresponded in snmp-server notify table item.

Security name is the community name for snmpv1 and snmp v2c, and username for snmpv3. 1, 2c, 3 mean SNMP versions. Port means the port number sent to. Notifytype-list means optional notify list. If it is unoptioned, default to choose all type. Only optionaed type will be sent to destination host.

For example :

! Configure SNMP server, the IP address is configured to be 192.168.0.100, and SNMP version to be 2c, and community name to be user

```
QTECH(config)#snmp-server host 192.168.0.100 version 2c user
```

! Delete the item with the notify destination host being 192.168.0.100 and community name being user

```
QTECH(config)#no snmp-server host 192.168.0.100 user
```

Display snmp-server notify item in any configuration mode : :

show snmp host

! Display Trap information of snmp

```
QTECH(config)#show snmp host
```

2.12.4 Configure sysLocation

sysLocation is a managing variable in system group of MIB which is used to denote location of devices be managed. Configure it in global configuration mode :

```
snmp-server location syslocation
```

Syslocation is the character string of system location ranges from 1 to 255 printable characters.

For example :

! Configure system location to be sample sysLocation factory.

```
QTECH(config)#snmp-server location "sample sysLocation factory"
```

Use quotation mark to quote space in character string.

Use show snmp location command in any configuration mode to display system location :

show snmp location

2.12.5 Configure sysName

sysName is a managing variable in system group of MIB which is switch name. Configure it in global configuration mode :

snmp-server name *sysname*

no snmp-server name

Sysname means the character string of system name ranges from 1 to 255 printable characters.

For example :

! Configure system name to be QSW-3900

```
QTECH(config)#snmp-server name "QSW-3900"
```



Caution : Use quotation mark to quote space in character string.

2.12.6 Configure notify

Enable/disable sending all kinds of notify types by configuring notify sending. The defaulted notify sending is trap. After disabling notify sending, trap will not be sent. Notify sending is defaulted to disable. Configure it in global configuration mode :

snmp-server enable traps [*notificationtype-list*]

no snmp-server enable traps [*notificationtype-list*]

notificationtype-list : Notificationtype list defined by system. To enable or disable specified notification type by choose one or several type. If the keyword is vacant, all types of notification are enabled or disabled.

Notify types are as following :

- 1) bridge : Enable/disable STP
- 2) interfaces : interface LinkUp/LinkDown
- 3) snmp : accessing control; cold boot/heat boot of system
- 4) gbnsavecfg : save configuration
- 5) rmon : RMON trap
- 6) gbn : self-define Trap, such as interface Blocking, CAR, loopback detect

For example :

! Enable notificationtype gbn

```
QTECH(config)# snmp-server enable traps gbn
```

2.12.7 Configure engine id

This configuration is used to configure local engine-id or recognizable remote engine-id.

Default local engine id is 27514000000000000000000000000000 which cannot be deleted but modified. It is defaulted to have no recognizable remote engine-id which can be added and deleted. Once delete a recognizable remote engine the corresponded user can also be deleted. At most 32 engines can be configured. Use no snmp-server engineID command to restore default local engine-id or remove remote engine-id. Configure it in global configuration mode :

snmp-server engineID { **local** *engineid-string* | **remote** *ip-address* [**udp-port** *port-number*] *engineid-string* }

no snmp-server engineID { **local** | **remote** *ip-address* [**udp-port** *port-number*] }

Display current engine configuration in any configuration mode :

show snmp engineID [local | remote]

engineid-string is an engine id that can only be recognized in a network. This system only supports printable characters of engine id which excludes space.

Ip-address is remote engine ip address. Local ip address is not allowed to input.

Port-number is remote engine port number. Default port number is 162

For example :

! Configure local engine id to be 12345

```
QTECH(config)# snmp-server engineid local 12345
```

! Configure remote engine that can be recognized locally. Configure remote engine ip to be 1.1.1.1, and port number to be 888, and id to be 1234

```
QTECH(config)# snmp-server engineid remote 1.1.1.1 udp-port 888 1234
```

! Display local engine configuration

```
QTECH(config)# show snmp engineid local
```

2.12.8 Configure view

Use `snmp-server view` command to configure view and its subtree. Iso, internet and sysview are the default views. At most 64 views can be configured. View Internet must not delete and modify. Configure it in global configuration mode :

```
snmp-server view view-name oid-tree { included | excluded }
```

```
no snmp-server view view-name [ oid-tree ]
```

View-name means the name of the view to be added. It ranges from 1 to 32, excluding space.

Oid-tree means the subtree of the view which corresponds to such a mib node as "1.3.6.1"; The substring of OID must be the integer between 0 and 2147483647.

In the view name string of character contains the character integer adds on which OID to contain the node integer adds on 2 again and do not surpass 64.

The sum of the number of characters in view name string and the number of oid nodes should not be more than 62.

When configuring view subtree to be exclude, the node in this subtree cannot be accessed which does not mean the node excluded this subtree can be accessed. When configuring notify destination host, if the security name is the community, sending notify is not effected on view; if the user with the security name being SNMPv3, sending notify is controlled by notify view of this user. What this notify view controlled is the accessing of the node that variable belongs to and it is not influence accessing attribution of trap OID that notify belonged to. If notify does not contain binded variable, sending notify is not effected on view.

For example :

! Add view "view1", and configure it to have a subtree "1.3.6.1"

```
QTECH(config)# snmp-server view view1 1.3.6.1 include
```

! Add a subtree "1.3.6.2" for existed view "view1"

```
QTECH(config)# snmp-server view view1 1.3.6.2 include
```

! Remove existed view "view1"

```
QTECH(config)# no snmp-server view view1
```

! Display configured view

```
QTECH(config)# show snmp view
```

2.12.9 Configure group

Use this configuration to configure a accessing conreol group. Folowing groups are default to exist : (1) security model is v3, the security level is differentiated group initial ; (2) security model is v3, the security level is differentiated encrypt group initial. At most 64 groups can be configured. Configure it in global configuration mode :

snmp-server group *groupname* { **1** | **2c** | **3** [**auth** | **noauth** | **priv**] [**context** *context-name*]} [**read** *readview*] [**write** *writeview*] [**notify** *notifyview*]

no snmp-server group *groupname* { **1** | **2c** | **3** [**auth** | **noauth** | **priv**] [**context** *context-name*]}

Display configured group in any configuration mode :

show snmp group

groupname means group name, which ranges from 1 to 32 characters, excluding space.

Readview is a view name, which means the right to read in the view. If the keyword is vacant, it is default not to include readable view.

Writeview is a view name, which means the right to read and write in the view. If the keyword is vacant, it is default not to include readable and writable view.

Notifyview is a view name, which means the right to send notification in the view. If the keyword is vacant, it is default not to include notify sending view.

Context-name is facility context. If the keyword is vacant, it is default to be local facility.

For example :

! Add group “group1” to local facility, using security model 1, and configure read, write, and notify view to be internet

```
QTECH(config)# snmp-server group group1 1 read internet write internet notify Internet
```

! Remove group “group1” from local facility

```
QTECH(config)# no snmp-server group group1 1
```

! Display current group configuration.

```
QTECH(config)# show snmp group
```

2.12.10 Configure user

Use this configuration to configure user for local engine and recognizable remote engine. Following users are default to exist : (1)initialmd5(required md5 authentication), (2) initialsha(required sha authentication), (3) initialnone(non- authentication). The above three users are reserved for system not for user. The engine the user belonged to must be recognizable. When deleting recognizable engine, contained users are all deleted. At most 64 users can be configured. Configure it in global configuration mode :

snmp-server user *username groupname* [*remote host* [*udp-port port*]] [**auth** { **md5** | **sha** } { **authpassword** { **encrypt-authpassword** *authpassword* | *authpassword* } | **authkey** { **encrypt-authkey** *authkey* | *authkey* } }] [**priv** **des** { **privpassword** { **encrypt-privpassword** *privpassword* | *privpassword* } | **privkey** { **encrypt-privkey** *privkey* | *privkey* } }]

no snmp-server user *username* [*remote host* [*udp-port port*]]

Display configured user in any configuration mode :

show snmp user

Username is the username to be configured. It ranges from 1 to 32 characters, excluding space.

Groupname is the groupname that user going to be added. It ranges from 1 to 32 characters, excluding space.

Host is remote engine ip address. If it is vacant, it is default to be local engine.

Port is the port number of remote engine. If it is vacant, it is default to be 162.

Authpassword is authentication password. Unencrypted password ranges from 1 to 32 characters. To avoid disclosing, this password should be encrypted. To configured encrypted password needs client-side which supports encryption to encrypt password, and use encrypted cryptograph to do the configuration. Cryptograph is different by different encryption. Input cryptograph in the form of hexadecimal system, such as “a20102b32123c45508f91232a4d47a5c”

Privpassword is encryption password. Unencrypted password ranges from 1 to 32 characters. To avoid disclosing, this password should be encrypted. To configured encrypted password needs client-side which supports encryption to encrypt password, and use encrypted cryptograph to do the configuration. Cryptograph is different by different encryption. Input cryptograph in the form of hexadecimal system, such as “a20102b32123c45508f91232a4d47a5c”

Authkey is authentication key. Unauthenticated key is in the range of 16 byte (using md5 key folding) or 20 byte (using SHA-1 key folding). Authenticated key is in the range of 16 byte (using md5 key folding) or 24 byte

(using SHA-1 key folding).

Privkey is encrypted key. Unencrypted key ranges from 16 byte, and encrypted key ranges from 16 byte.

Keyword encrypt-authpassword, encrypt-authkey, encrypt-privpassword, encrypt-privkey are only used in command line created by compile to prevent leaking plain text password and key. When deconfiguring SNMP, user cannot use above keywords.

For example :

! Add user "user1" for local engine to group "grp1", and configure this user not to use authentication and encryption.

```
QTECH(config)# snmp-server user user1 grp1
```

! Add user "user2" for local engine to group "grp2", and configure this user to use md5 authentication and non-encryption with the auth-password to be 1234

```
QTECH(config)# snmp-server user user2 grp2 auth md5 auth-password 1234
```

! Add user "user3" for local engine to group "grp3", and configure this user to use md5 authentication and des encryption with the auth-password to be 1234 and privpassword to be 4321

```
QTECH(config)# snmp-server user user3 grp3 auth md5 auth-password 1234 priv des
priv-password 4321
```

2.13 Enable/disable dlf forward packet

Use dlf-forward command to enable dlf forward.

```
dlf-forward { multicast | unicast }
```

```
no dlf-forward { multicast | unicast }
```

Use dlf-forward command in global configuration mode or interface configuration mode to enable dlf forward. Use no dlf-forward command to disable dlf forward :

```
dlf-forward { multicast | unicast }
```

```
no dlf-forward { multicast | unicast }
```

For example :

! Disable dlf forward for unicast

```
QTECH(config)#no dlf-forward unicast
```

! Disable dlf forward for multicast

```
QTECH(config)#no dlf-forward multicast
```

2.14 CPU Alarm Configuration

2.14.1 Brief introduction of CPU alarm

System can monitor CPU usage. If CPU usage rate is beyond cpu busy threshold, cpu busy alarm is sent because the cpu is busy. In this status, if cpu is below cpu unbusy threshold, cpu unbusy alarm is sent. This function can report current CPU usage to user.

2.14.2 CPU alarm configuration list

CPU alarm configuration command includes :

- Enable/disable CPU alarm
- Configure CPU busy or unbusy threshold
- Display CPU alarm information

2.14.3 Enable/disable CPU alarm

Configure it in global configuration mode :
Enable CPU alarm

alarm cpu

Disable CPU alarm

no alarm cpu

by default, CPU alarm enables.

For example :

! Enable CPU alarm

```
QTECH(config)#alarm cpu
```

2.14.4 Configure CPU busy or unbusy threshold

Use alarm cpu threshold command in global configuration mode to configure CPU busy or unbusy threshold :

Configure CPU busy or unbusy threshold

alarm cpu threshold [busy *busy*] [unbusy *unbusy*]

busy > unbusy. Default CPU busy threshold is 90%, and CPU unbusy threshold is 60%.

For example :

! Configure CPU busy threshold to be 30%, and CPU unbusy threshold to be 10%

```
QTECH(config)#alarm cpu threshold busy 30 unbusy 10
```

2.14.5 Display CPU alarm information

Use show alarm cpu command in any mode to display cpu alarm information :

show alarm cpu

For example :

! Display CPU alarm information

```
QTECH(config)#show alarm cpu
```

```
CPU status alarm      : enable
```

```
CPU busy threshold(%) : 90
```

```
CPU unbusy threshold(%) : 60
```

```
CPU status           : unbusy
```

2.15 Anti-DOS Attack

2.15.1 IP segment anti-attack

The IP segment packet number which can be received by system do not occupy resources of all receiving packets, which can normally handle other non-segment packets when receiving IP segment attack and the range of IP segment receiving number can be configured. 0 means system will not handle IP segment packet so that system can avoid the influence on segment attack.

Configure it in global configuration mode

anti-dos ip fragment *maxnum*

Display related information

show anti-dos

Chapter 3 MAC address table management

3.1 Introduction to Bridging

A bridge is a store-and-forward device that connects and transfers traffic between local area network (LAN) segments at the data-link layer. In some small-sized networks, especially those with dispersed distribution of users, the use of bridges can reduce the network maintenance costs, without requiring the end users to perform special configurations on the devices.

In applications, there are four major kinds of bridging technologies : transparent bridging, source-route bridging (SRB), translational bridging, and source-route translational bridging (SR/TLB).

Transparent bridging is used to bridge LAN segments of the same physical media type, primarily in Ethernet environments. Typically, a transparent bridging device keeps a bridge table, which contains mappings between destination MAC addresses and outbound interfaces.

Presently the devices support the following transparent bridging features :

- Bridging over Ethernet
- Bridging over point-to-point (PPP) and high-level data link control (HDLC) links
- Bridging over X.25 links
- Bridging over frame relay (FR) links
- Inter-VLAN transparent bridging
- Routing and bridging are simultaneously supported

3.2 Major Functionalities of Bridges

3.2.1 Maintaining the bridge table

A bridge relies on its bridge table to forward data. A bridge table consists two parts : MAC address list and interface list. Once connected to a physical LAN segment, a bridge listens to all Ethernet frames on the segments. When it receives an Ethernet frame, it extracts the source MAC address of the frame and creates a mapping entry between this MAC address and the interface on which the Ethernet frame was received.

As shown in I. Figure 1, Hosts A, B, C and D are attached to two LAN segments, of which LAN segment 1 is attached to bridge interface 1 while LAN segment 2 is connected with bridge interface 2. When Host A sends an Ethernet frame to Host B, both bridge interface 1 and Host B receive this frame.

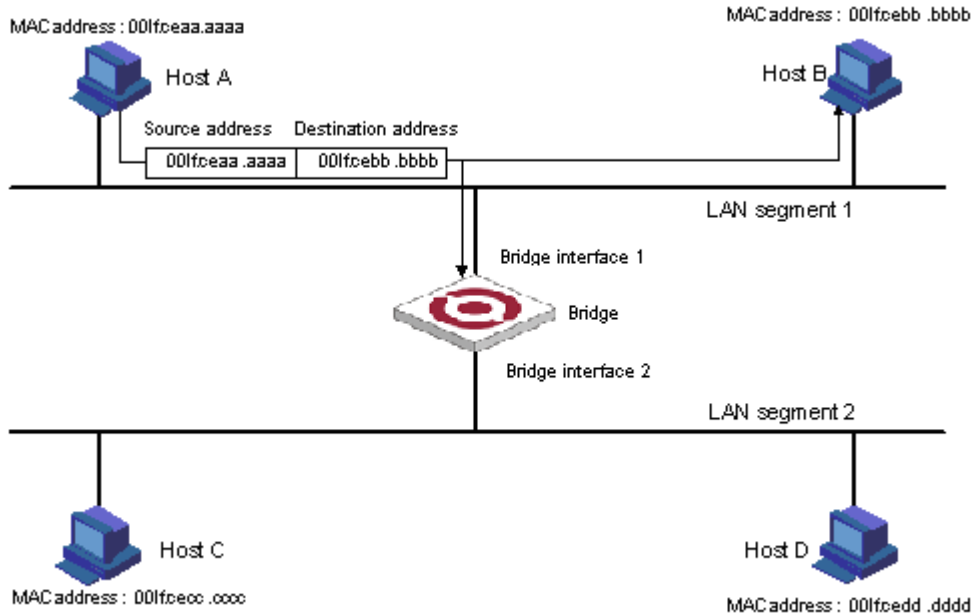


Figure 1. Host A sends an Ethernet frame to Host B on LAN segment 1

As the bridge receives the Ethernet frame on bridge interface 1, it determines that Host A is attached to bridge interface 1 and creates a mapping between the MAC address of Host A and bridge interface 1 in its bridge table, as shown in Figure 2.

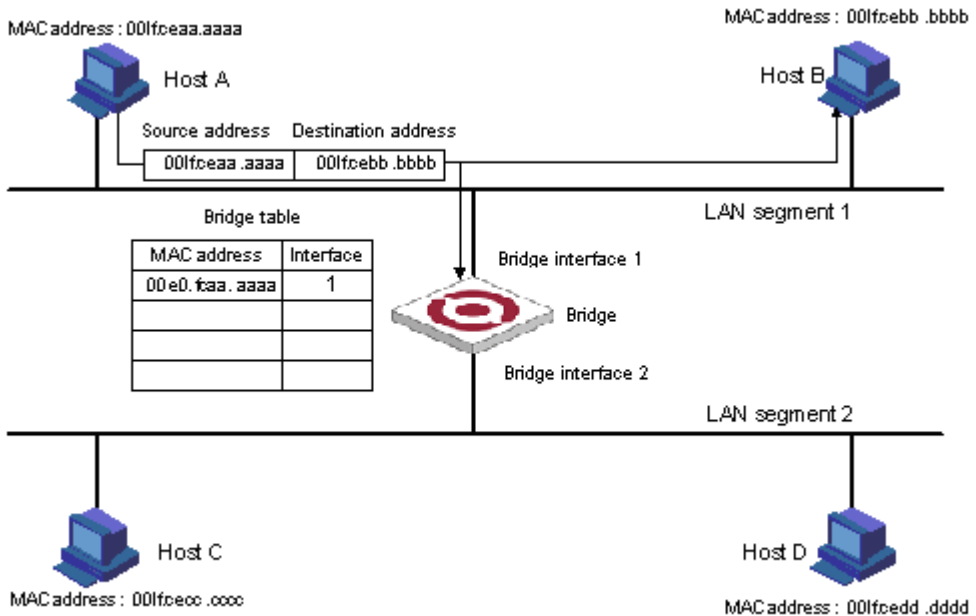


Figure 2 The bridge determines that Host A is attached to interface 1

When Host B responds to Host B, the bridge also hears the Ethernet frame from Host B. As the frame is received on bridge interface 1, the bridge determines that Host B is also attached to bridge interface 1, and creates a mapping between the MAC address of Host B and bridge interface 1 in its bridge table, as shown in Figure 3.

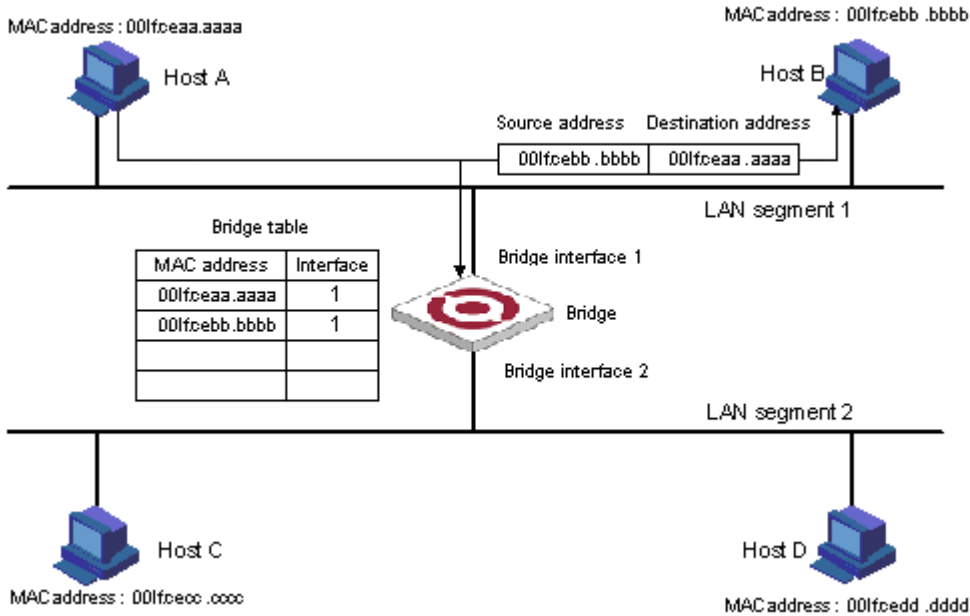


Figure 3 The bridge determines that Host B is also attached to interface 1

Finally, the bridge obtains all the MAC-interface mappings (assume that all hosts are in use), as shown in Figure 4.

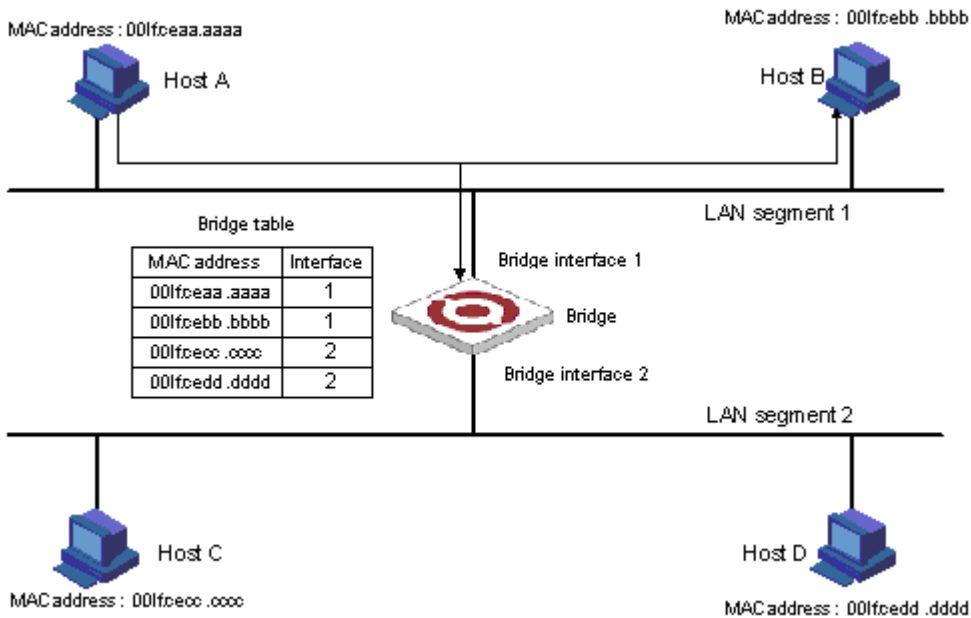


Figure 4 The final bridge table

3.2.2 Forwarding and filtering

The bridge makes data forwarding or filtering decisions based on the following scenarios :

When Host A sends an Ethernet frame to Host C, the bridge searches its bridge table and finds out that Host C is attached to bridge interface 2, and forwards the Ethernet frame out of bridge interface 2, as shown in II. Figure 5.

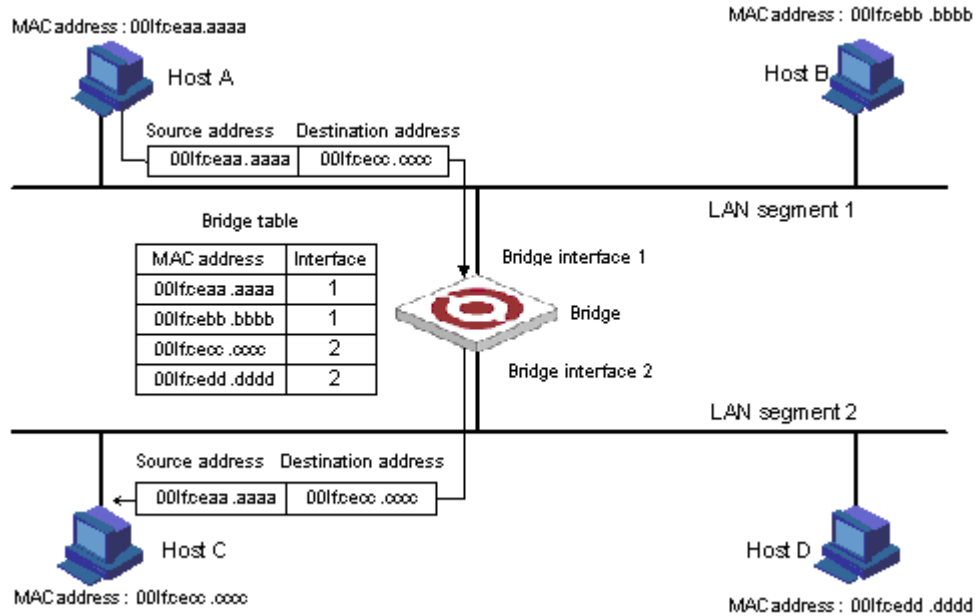


Figure 5 Forwarding

When Host A sends an Ethernet frame to Host B, as Host B is on the same LAN segment with Host A, the bridge filters the Ethernet frame instead of forwarding it, as shown in II. Figure 6.

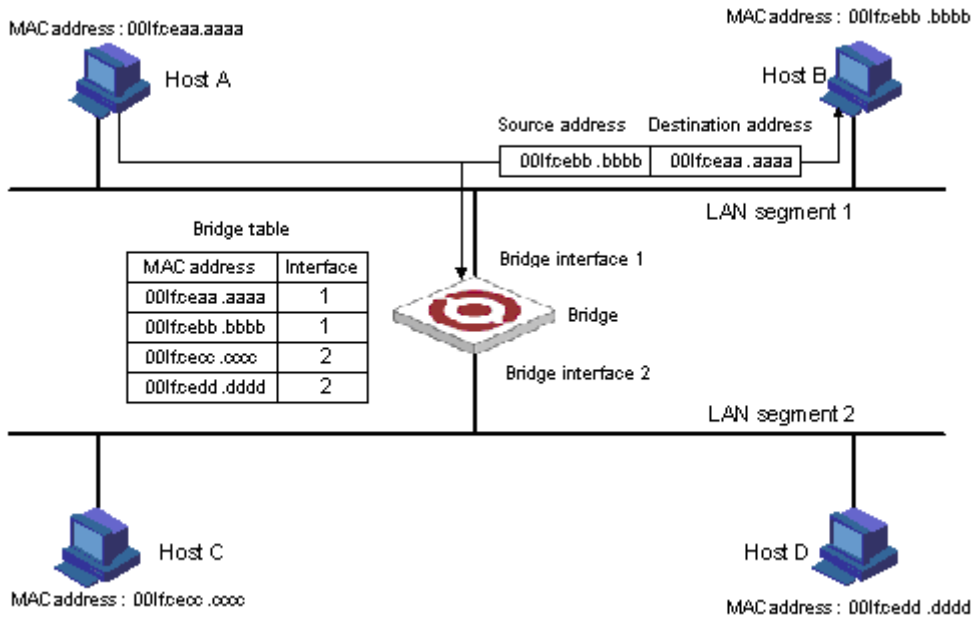


Figure 6 Filtering

When Host A sends an Ethernet frame to Host C, if the bridge does not find a MAC-to-interface mapping about Host C in its bridge table, the bridge forwards the Ethernet frame to all interfaces except the interface on which the frame was received, as shown in Figure 7.

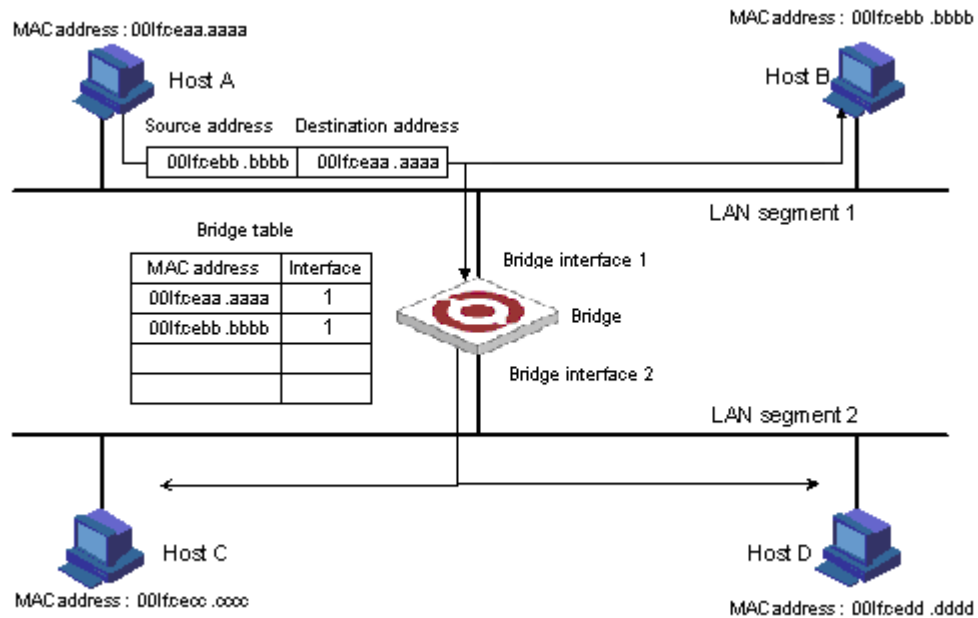


Figure 7 The proper MAC-to-interface mapping is not found in the bridge table

Note :

When a bridge receives a broadcast or multicast frame, it forwards the frame to all interfaces other than the receiving interface.

3.3 Brief introduction of MAC address table management

System maintains a MAC address table which is used to transfer packet. The item of this table contains MAC address, VLAN ID and interface number of packet entering. When a packet entering switch, switch will look up the MAC address table according to destination MAC and VLAN ID of the packet. If it is found out, send packet according to the specified interface in the item of MAC address table, or the packet will be broadcasted in this VLAN. In SVL learning mode, look up the table only according to MAC in packet and neglect VLAN ID.

System possesses MAC address learning. If the source MAC address of the received packet does not exist in MAC address table, system will add source MAC address, VLAN ID and port number of receiving this packet as a new item to MAC address table.

MAC address table can be manual configured. Administrator can configure MAC address table according to the real situation of the network. Added or modified item can be static, permanent, blackhole and dynamic.

System can provide MAC address aging. If a device does not receive any packet in a certain time, system will delete related MAC address table item. MAC address aging is effective on (dynamic) MAC address item which can be aging by learning or user configuration.

3.4 MAC address table management list

MAC address table management

- Configure system MAC address aging time
- Configure MAC address item
- Enable/disable MAC address learning
- Modify MAC address learning mode

3.5 Configure system MAC address aging time

Use `mac-address-table age-time` command in global configuration mode to configure MAC address aging time. Use `no mac-address-table age-time` command to restore it to default time.

mac-address-table age-time { *agetime* | *disable* }

no mac-address-table age-time

Agetime means MAC address aging time which ranges from 1 to 1048575 seconds. Default MAC address aging time is 300 seconds. *Disable* means MAC address not aging. Use `no` command to restore the default MAC address aging time.

For example :

! Configure MAC address aging time to be 3600 seconds

```
QTECH(config)#mac-address-table age-time 3600
```

! Restore MAC address aging time to be 300 seconds

```
QTECH(config)#no mac-address-table age-time
```

Display MAC address aging time

show mac-address-table age-time

Use `show mac-address-table age-time` command to display MAC address aging time.

show mac-address-table age-time

For example :

! Display MAC address aging time.

```
QTECH(config)#show mac-address-table aging-time
```

3.6 Configure MAC address item

3.6.1 Add MAC address

MAC address table can be added manually besides dynamically learning.

mac-address-table { *dynamic* | *permanent* | *static* } *mac* **interface** *interface-num* **vlan** *vlan-id*

Parameter *mac*, *vlan-id* and *interface-num* corresponded to the three attributions of the new MAC address table item.

MAC address attribution can be configured to be *dynamic*, *permanent* and *static*. *Dynamic* MAC address can be aging; *permanent* MAC address will not be aging and this MAC address will exist after rebooting; *static* MAC address will not be aging, but it will be lost after rebooting.

For example :

! Add mac address 00 : 01 : 02 : 03 : 04 : 05 to be static address table.

```
QTECH(config)#mac-address-table static 00 : 01 : 02 : 03 : 04 : 05 interface ethernet 0/0/1
vlan 1
```

3.6.2 Add blackhole MAC address

System can configure MAC address table item to be blackhole item. When the source address or destination address is blackhole MAC address, it will be dropped.

mac-address-table blackhole mac *vlan* *vlan-id*

For example :

! When tagged head of the packet is VLAN 1, forbid packet with its source address or destination address being 00 : 01 : 02 : 03 : 04 : 05 to go through system

QTECH(config)#mac-address-table blackhole 00 : 01 : 02 : 03 : 04 : 05 vlan 1

3.6.3 Delete MAC address item

Use no mac-address-table command to remove mac address table.

no mac-address-table [blackhole | dynamic | permanent | static] *mac* vlan *vlan-id*

no mac-address-table [dynamic | permanent | static] *mac* interface *interface-num* vlan *vlan-id*

no mac-address-table [dynamic | permanent | static] interface *interface-num*

no mac-address-table [blackhole | dynamic | permanent | static] vlan *vlan-id*

no mac-address-table

Vlan means delete MAC address table item according to vlan-id; mac means deleting a specified MAC address table item; interface-num means delete MAC address table item according to interface number; command no mac-address-table means delete all MAC address.

For example :

! Delete all MAC address table item

QTECH(config)#no mac-address-table

3.6.4 Display MAC address table

Use show mac-address command to display MAC address table.

show mac-address-table

show mac-address-table { *interface-num* [**vlan** *vlan-id*] | cpu }

show mac-address-table *mac* [vlan *vlan-id*]

show mac-address-table { blackhole | dynamic | permanent | static } [**vlan** *vlan-id*]

show mac-address-table { blackhole | dynamic | permanent | static } **interface** *interface-num* [**vlan** *vlan-id*]

show mac-address-table vlan *vlan-id*

The parameter meaning is the same as that of add/delete MAC address table item.

3.6.5 Enable/disable MAC address learning

This command is a batch command in global configuration mode to configure all interfaces to be the same; in interface configuration mode, it can configure interface MAC address learning. When MAC address learning is forbidden in an interface, packet with unknown destination address received from other interface will not be transmitted to this interface; and packet from this interface whose source address is not in this interface will not be transmitted. By default, all interface MAC address learning enable.

mac-address-table learning

no mac-address-table learning

For example :

! Enable MAC address learning on interface Ethernet 0/0/7.

QTECH(config-if-ethernet-0/0/7)#no mac-address-table learning

3.6.6 Display MAC address learning

show mac-address learning [interface [*interface-num*]]

Use show mac-address-table learning command to display MAC address learning.

3.6.7 Modify MAC address learning mode

System supports SVL and IVL learning modes. The default one is SVL. User can configure MAC learning mode in global configuration mode. It will be effective after rebooting.

mac-address-table learning mode { svl | ivl }

show mac-address-table learning mode

For example :

! Modify MAC address to be IVL

```
QTECH(config)#mac-address-table learning mode ivl
```

! Display MAC address learning mode.

```
QTECH(config)#show mac-address-table learning mode
```

Chapter 4 Port Configuration

4.1 Port configuration introduction

System can provide 24 10/100Base-T Ethernet interfaces, 2 1000Base-TX(LX/SX) Ethernet interfaces and a Console interface. Ethernet interface can work in half duplex and full duplex mode, and can negotiate other working mode and speed rate with other network devices to option the best working mode and speed rate automatically to predigest system configuration and management.

4.2 Port Configuration

4.2.1 Port related configuration

Configure related feature parameter of ports should enter interface configuration mode first, and then configure.

Interface configuration list is as following :

- Enter interface configuration mode
- Enable /disable specified interface
- Configure duplex mode and speed rate
- Configure interface privilege
- Configure interface limited speed
- Configure type of receiving frame
- Configure interface type
- Configure default VLAN ID of trunk port
- Add access port to specified VLAN
- Display interface information

4.2.2 Enter interface configuration mode

Enter interface configuration mode before configuration.

Configure as following in global configuration mode :

Enter interface configuration mode

interface ethernet *interface-number*

Interface-num is Ethernet interface number which is in the form of device-num/slot-num/port-num, in which device-num is in the range of 0 to 7, slot-num is in the range of 0 to 2, and port-num is in the range of 1 to 48

4.2.3 Enable/disable specified interface

After system booting, all the interfaces are defaulted to be enable, and each interface can be configured according to real situation.

Use following commands to enable/disable an Ethernet port.

shutdown

no shutdown

Shutdown means disable a port, while no shutdown means enable a port.

For example :

! Enable Ethernet interface 1

```
QTECH(config-if-ethernet-0/0/1)#no shutdown
```

! Disable Ethernet interface 25

```
QTECH(config-if-ethernet-0/1/1)#shutdown
```

When interface is shutdown, the physical link is working for diagnosis.

4.2.4 Configure interface duplex mode and speed rate

100 BASE TX supports the speed of 10Mbps and 100Mbps, while 100 BASE FX supports the speed of 100Mbps. 1000 BASE TX supports the speed of 10Mbps, 100Mbps and 1000Mbps, while 1000 BASE FX supports the speed of 1000Mbps. 100 BASE TX and 1000 BASE TX support the duplex mode of half, full duplex and auto-negotiation mode. 100 BASE FX and 1000 Base FX only support the duplex mode of full duplex. By default, 100 Base FX is in the mode of 100M and full duplex, and other interfaces are auto-negotiation. User can configure the working mode by himself. Use speed command to configure the speed and duplex command to configure duplex.

Command form in interface mode

```
speed { 10 | 10auto | 100 | 100 auto | 1000 | 1000 auto | auto }
```

no speed

```
duplex { auto | full | half }
```

no duplex

For example :

! Configure the speed of Ethernet 0/0/1 to 100Mbps and duplex mode to be full duplex

```
QTECH(config-if-ethernet-0/0/1)#speed 100
```

```
QTECH(config-if-ethernet-0/0/1)#duplex full
```

In system, which of the speed or duplex setup to auto, and the another will be setup to auto too.

4.2.5 Interface Priority Configuration

There are 8 priorities from 0 to 7, and the default interface priority is 0. The larger the priority value is, the higher the priority is. And the packet with the higher priority will be quickly handled. If there are too much packet to be handled in some interface or the packet is urgent to be handled, priority of this interface can be configured to be high-priority.

Use following command in interface configuration mode :

Configure priority of Ethernet 0/0/5 to be 1

```
QTECH(config-if-ethernet-0/0/5)#priority 1
```

Restore the default priority of Ethernet 0/0/5

```
QTECH(config-if-ethernet-0/0/5)#no priority
```

4.2.6 Interface description configuration

Use following command to describe interface to distinguish each interface from others. Configure it in interface configuration mode.

description *description-list*

For example :

! Configure description string “red” for the Ethernet 0/0/3

```
QTECH(config-if-ethernet-0/0/3)#description qtech
! Display description of Ethernet 0/0/3
QTECH(config)#show description interface ethernet 0/0/3
```

4.2.7 Ingress/egress bandwidth-control configuration

Egress/ingress bandwidth-control is to restrict the total speed rate of all sending and receiving packets. Use following command to configure egress/ingress bandwidth-control. Configure it in interface configuration mode :
Interface egress/ingress bandwidth-control

bandwidth { **ingress** | **egress** } *target-rate*
Cancel egress/ingress bandwidth

no bandwidth { **ingress** | **egress** }
Detailed description of this command please refer to the corresponding command reference.

4.2.8 Enable/disable VLAN filtration of receiving packet of interface

When enabling VLAN ingress filtration, received 802.1Q packet which doesn't belong to the VLAN where the interface locates will be dropped. The packet will not be dropped if it is disabled.

Use this command in interface configuration mode.

ingress filtering

no ingress filtering

Example :

```
! Enable VLAN ingress filtration of e0/0/5
```

```
QTECH(config-if-ethernet-0/0/5)#ingress filtering
```

```
! Disable VLAN ingress filtration of e0/0/5
```

```
QTECH(config-if-ethernet-0/0/5)#no ingress filtering
```

4.2.9 Interface ingress acceptable-frame configuration

Configure ingress acceptable frame mode to be all types or only tagged.

Use following command in interface configuration mode to configure or cancel the restriction to ingress acceptable-frame :

ingress acceptable-frame { all | tagged }

no ingress acceptable-frame

For example :

```
! Configure Ethernet 0/0/5 only to receive tagged frame
```

```
QTECH(config-if-ethernet-0/0/5)#ingress acceptable-frame tagged
```

4.2.10 Enable/disable interface flow-control

If the port is crowded, it needs controlling to avoid congestion and data loss. Use flow-control command to control the flow. Use following command to enable/disable flow-control on current Ethernet port.

flow-control

no flow-control

For example :

! Enable flow control on Ethernet 0/0/5

```
QTECH(config-if-ethernet-0/0/5)#flow-control
```

! Disable flow control on Ethernet 0/0/5

```
QTECH(config-if-ethernet-0/0/5)#no flow-control
```

Use following command in any configuration mode to display interface flow-control :

show flow-control [*interface-num*]

For example :

! Display flow-control of Ethernet 0/0/5

```
QTECH(config-if-ethernet-0/0/5)#show flow-control ethernet 0/0/5
```

4.2.11 Port mode configuration

Use this command to configure port mode. If a port configures to be a trunk port, the vlan mode changes untagged into tagged, and if a port configures to be an access one, the vlan mode changes tagged into untagged. Configure it in interface configuration mode :

Configure port mode

switchport mode { trunk | access }

Restore default port mode : access port

no switchport mode

For example :

! Configure Ethernet 0/0/1 to be trunk port

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

4.2.12 Trunk allowed VLAN configuration

Use switchport trunk allowed vlan command to add trunk port to specified VLAN. Use no switchport trunk allowed vlan command to remove trunk port from specified vlan.

Add trunk port to specified vlan

switchport trunk allowed vlan { *vlan-list* | all }

Remove trunk port from specified vlan

no switchport trunk allowed vlan { *vlan-list* | all }

For example :

! Add trunk ports Ethernet0/0/1 to VLAN 3, 4, 70 to 150

```
QTECH(config-if-ethernet-0/0/1)# switchport trunk allowed vlan 3, 4, 70- 150
```

4.2.13 The default vlan-id of trunk port configuration

Use switchport trunk native vlan command to configure the default vlan-id (pvid) of trunk port. When receiving untagged packet, it will be transferred to VLAN defaulted VLAN ID. Packet receiving and sending follow IEEE 802.1Q. Configure it in interface configuration :

Configure default VLAN ID of trunk port

switchport trunk native vlan *vlan-id*

Restore default VLAN ID of trunk port

no switchport trunk native

Caution : above configuration is effective to trunk port. By default, default VLAN ID is 1. If this port is not in VLAN 1, configuration fails.

4.2.14 Add access port to specified VLAN

Use switchport access command to add access port to specified VLAN, and the default VLAN-ID is configured to be the specified VLAN. Configure it in interface configuration mode :

Add current port to specified VLAN, and the default VLAN-ID is configured to be the specified VLAN

switchport access vlan *vlan-id*

Remove current port from specified VLAN, if the default vlan-id of the current port is the specified VLAN and this port also belongs to VLAN 1, the default vlan-id of the current port restores to be 1, or the default VLAN ID will not be changed.

no switchport access vlan *vlan-id*

The precondition to use this command is the current port is access port and the VLAN to be added is not default VLAN 1.

4.2.15 Display interface information

Use **show interface** [*interface-num*] to display information of specified interface or all interfaces :

- 1) Interface state (enable/disable)
- 2) Connection
- 3) Working mode (full duplex, half duplex or auto-negotiation)
- 4) Default VLAN ID
- 5) Interface priority
- 6) Port mode (trunk/access port)

If no parameter is input in show interface [interface-num] command, information of all interfaces will be displayed.

4.2.16 Display/ clear interface statistics information

Use **show statistics interface** [*interface-num*] command in any configuration mode to display information of specified interface or all interfaces :

- Byte receiving
- Unicast packet receiving
- Non-unicast packet receiving
- Unicast packet sending
- Non-unicast packet sending

Use **clear interface** [*interface-num* | *slot-num*] command in global configuration mode to clear information of specified interface or all interfaces in specified slot or all interfaces. Use clear interface command in interface configuration mode to clear information of current interface.

4.3 Interface mirror

4.3.1 Brief introduction of interface mirror

System provides mirror based on interface, that is, copy packet in a or more specified interface to monitor interface to analyze and monitor packet. For example, copy packet of Ethernet 0/0/2 to specified monitor interface Ethernet 0/0/3 so that test and keep record by protocols linked by monitor interface Ethernet 0/0/3.

System also provides packet mirror for specified source/destination MAC address. For example, mirror packet from Ethernet 0/0/3 with the destination MAC address of 00 : 1f : ce : 10 : 14 : f1.

System also provides mirror divider, that is, sample packet that can be mirrored and send it to mirror

destination interface to reduce the number of packet to mirror destination interface.

4.3.2 Interface mirror configuration

Interface Mirror configuration command includes :

- Configure mirror destination interface
- Configure mirror source interface
- Display interface mirror

4.3.2.1 Configure mirror interface

Configure mirror destination interface in global configuration mode :

mirror destination-interface *interface-num*

This command will cancel original mirror destination interface.

Remove mirror interface :

no mirror destination-interface *interface-num*

For example :

! Configure Ethernet 0/0/1 to be mirror interface

```
QTECH(config)# mirror destination-interface ethernet 0/0/1
```

4.3.2.2 Configure mirror source interface

Configure mirror source-interface of switch in global configuration mode :

Configure mirror source-interface

mirror source-interface { *interface-list* | cpu } { both | egress | ingress }

interface-list is in the form of *interface-num* [to *interface-num*], which can be repeated for 3 times. *Cpu* interface is in the form of character string "cpu", both means mirroregress and ingress interfaces, egress means mirror interface egress and ingress means mirror interface ingress.

Remove mirror source interface

no mirror source-interface { *interface-list* | cpu }

For example :

! Configure Ethernet 0/0/1 to Ethernet 0/0/12 to be mirror source interfaces

```
QTECH(config)# mirror source-interface ethernet 0/0/1 to ethernet 0/0/12 both
```

! Remove Ethernet 0/0/10 to Ethernet 0/0/12 from mirror source interfaces

```
QTECH(config)#no mirror source-interface ethernet 0/0/10 to ethernet 0/0/12
```

4.3.2.3 Display interface mirror

Use show mirror command to display system configuration of current mirror interface, including monitor port and mirrored port list. Use this command in any configuration mode :

show mirror

For example :

! Display monitor port and mirrored port list

```
QTECH#show mirror
```

4.4 Brief introduction of Port LACP

Port convergence is a channel group formed by many ports convergence to realize flow load sharing for each member. When a link cannot be used, flow of this link will be transferred to another link to guarantee the smoothness of the flow.

4.4.1 LACP

The link aggregation control protocol (LACP) is defined in IEEE 802.3ad. Link aggregation control protocol data unit (LACPDU) is used for exchanging information among LACP-enabled devices.

After LACP is enabled on a port, the port sends LACPDUs to notify the remote system of its system LACP priority, system MAC address, port LACP priority, port number, and operational key. Upon receipt of an LACPDU, the remote system compares the received information with the information received on other ports to determine the ports that can operate as selected ports. This allows the two systems to reach agreement on the states of the related ports

When aggregating ports, link aggregation control automatically assigns each port an operational key based on its rate, duplex mode, and other basic configurations. In an LACP aggregation group, all ports share the same operational key; in a manual or static LACP aggregation, the selected ports share the same operational key.

Basic configurations are :

1. 13 static or dynamic channel groups can be configured and at most 12 interface members can be configured in each group, and at most 8 interfaces can be convergent at the same time in each group which is determined by up/down status, interface number, LACP priority. Each group is defined to be a channel group, and the command line is configured around it.

2. Load balance strategy of each group can be divided into source MAC, destination MAC, source and destination MAC, source IP, destination IP, and source and destination IP. The default strategy is source MAC.

3. System and interface LACP priority can be configured. The default system priority is 32768, and interface priority is 128. To remove system and interface priority is to restore them to default ones.

4. LACP protocol of each interface can be configured. In static mode, interface is static convergent, and LACP protocol does not run; in active mode, interface will initiate LACP negotiation actively; in passive mode, interface only can response LACP negotiation. When interconnecting with other device, static mode only can interconnect with static mode; active can interconnect with active and passive mode, but passive mode only can interconnect with active mode. The default mode of interface is ACTIVE mode.

Each convergent interface need same layer 2 features, so there are following restrictions to interfaces in a channel group :

Static convergent interfaces and dynamic convergent interfaces can not be in a same channel group, but there can be static convergent channel as well as dynamic convergent channel.

Each interface in a same channel group must possess the same features as following : interface speed rate, working mode of full duplex, STP/GVRP/GMRP function, STP cost, STP interface priority, VLAN features (interface mode, PVID, VLAN belonged to, tag vlan list of access interface, allowed vlan list of trunk interface) and layer 2 multicast group belonged to.

If modifying the feature of one interface in the channel group, other interfaces will be modified automatically in the same place. The feature refers to point 2.

After convergence, static hardware item (ARL, MARL, PTABLE, VTABLE) will be modified, but there will be delay.

After convergence, only host interface can send CPU packet. If STP changes status of some interface, the status of the whole channel group will be changed.

After convergence, when transferring layer 2 protocol packet, STP/GARP/GNLINK will not transfer packet to the current channel grou. If transferring to other channel group, only one packet will be transferred.

If there are members in the channel group, this channel group cannot be deleted. Delete interface members first.

Influence on choosing link redundancy caused by LACP system and interface priority. LACP provides link redundancy mechanism which needs to guarantee the redundancy consistency of two interconnected switches and user can configure redundancy link which is realized by system and interface priority. The redundancy choosing follows the following steps :

First, determine which switch is the choosing standard. For LACP packets interaction, each of the two switches knows each other's LACP system priority and system MAC and compares the LACP system priority to choose the smaller one; if the system priority is the same, compare MAC and choose the smaller one.

Then, choose redundancy link according to the interface parameter of the chosen switch. Compare interface LACP priority, and choose the inferior one to be redundant. If the priorities are the same, choose the interface whose interface number is larger to be redundant.

Link aggregation aggregates multiple physical Ethernet ports into one logical link, also called a logical group, to increase reliability and bandwidth.

4.4.2 Manual Link Aggregation

4.4.2.1 Overview

Manual aggregations are created manually. Member ports in a manual aggregation are LACP-disabled.

4.4.2.2 Port states in a manual aggregation

In a manual aggregation group, ports are either selected or unselected. Selected ports can receive and transmit data frames whereas unselected ones cannot. Among all selected ports, the one with the lowest port number is the master port and others are member ports.

When setting the state of ports in a manual aggregation group, the system considers the following :

- Select a port from the ports in up state, if any, in the order of full duplex/high speed, full duplex/low speed, half duplex/high speed, and half duplex/low speed, with the full duplex/high speed being the most preferred. If two ports with the same duplex mode/speed pair are present, the one with the lower port number wins out. Then, place those ports in up state with the same speed/duplex pair, link state and basic configuration in selected state and all others in unselected state.
- When all ports in the group are down, select the port with the lowest port number as the master port and set all ports (including the master) in unselected state.
- Place the ports that cannot aggregate with the master in unselected state, for example, as the result of the cross-board aggregation restriction.

Manual aggregation limits the number of selected ports in an aggregation group. When the limit is exceeded, the system changes the state of selected ports with greater port numbers to unselected until the number of selected ports drops under the limit.

In addition, unless the master port should be selected, a port that joins the group after the limit is reached will not be placed in selected state even if it should be in normal cases. This is to prevent the ongoing service on selected ports from being interrupted. You need to avoid the situation however as the selected/unselected state of a port may become different after a reboot.

4.4.2.3 Port Configuration Considerations in manual aggregation

As mentioned above, in a manual aggregation group, only ports with configurations consistent with those of the master port can become selected. These configurations include port rate, duplex mode, link state and other basic configurations.

You need to maintain the basic configurations of these ports manually to ensure consistency. As one configuration change may involve multiple ports, this can become troublesome if you need to do that port by port. As a solution, you may add the ports into an aggregation port group where you can make configuration for all member ports.

When the configuration of some port in a manual aggregation group changes, the system does not remove the aggregation; instead, it re-sets the selected/unselected state of the member ports and re-selects a master port.

4.4.3 Static LACP link aggregation

4.4.3.1 Overview

Static aggregations are created manually. After you add a port to a static aggregation, LACP is enabled on it automatically.

4.4.3.2 Port states in static aggregation

In a static aggregation group, ports can be selected or unselected, where both can receive and transmit LACPDU's but only selected ports can receive and transmit data frames. The selected port with the lowest port number is the master port and all others are member ports.

All member ports that cannot aggregate with the master are placed in unselected state. These ports include those using the basic configurations different from the master port or those located on a board different from the master port because of the cross-board aggregation restriction.

Member ports in up state can be selected if they have the configuration same as that of the master port. The number of selected ports however, is limited in a static aggregation group. When the limit is exceeded, the local and remote systems negotiate the state of their ports as follows :

1) Compare the actor and partner system IDs that each comprises a system LACP priority plus a system MAC address as follow :

- First compare the system LACP priorities. The system with lower system LACP priority wins out.

- If they are the same, compare the system MAC addresses. The system with the smaller ID has higher priority. (the lower the LACP priority, the smaller the MAC address, and the smaller the device ID)

2) Compare the port IDs that each comprises a port LACP priority and a port number on the system with higher ID as follows :

- Compare the port LACP priorities. The port with lower port LACP priority wins out.
- If two ports with the same port LACP priority are present, compare their port numbers. The state of the ports with lower IDs then change to selected and the state of the ports with higher IDs to unselected, so does the state of their corresponding remote ports. (the lower the LACP priority, the smaller the port number, and the smaller the port ID)

4.4.3.3 Port configuration considerations in static aggregation

Like in a manual aggregation group, in a static LACP aggregation group, only ports with configurations consistent with those of the master port can become selected. You need to maintain the basic configurations of these ports manually to ensure consistency. As one configuration change may involve multiple ports, this can become troublesome if you need to do that port by port. As a solution, you may add the ports into an aggregation port group where you can make configuration for all member ports.

When the configuration of some port in a static aggregation group changes, the system does not remove the aggregation; instead, it re-sets the selected/unselected state of the member ports and re-selects a master port.

4.5 Load-Balance in a Link Aggregation Group

Link aggregation groups fall into load sharing aggregation groups and non-load sharing aggregation groups depending on their support to load sharing.

A load sharing aggregation group can contain at least one selected port but a non-load sharing aggregation group can contain only one.

Link aggregation groups perform load sharing depending on availability of hardware resources. When hardware resources are available, link aggregation groups created containing at least two selected ports perform load sharing, while link aggregation groups created with only one selected port perform load sharing depending on the model of your device. After hardware resources become depleted, link aggregation groups work in non-load sharing mode.

4.6 Aggregation Port Group

As mentioned earlier, in a manual or static aggregation group, a port can be selected only when its configuration is the same as that of the master port in terms of duplex/speed pair, link state, and other basic configurations. Their configuration consistency requires administrative maintenance, which is troublesome after you change some configuration.

To simplify configuration, port-groups are provided allowing you to configure for all ports in individual groups at one time. One example of port-groups is aggregation port group.

Upon creation or removal of a link aggregation group, an aggregation port-group which cannot be administratively created or removed is automatically created or removed. In addition, you can only assign/remove a member port to/from an aggregation port-group by assigning/removing it from the corresponding link aggregation group.

4.7 Link aggregation configuration

Port LACP configuration command includes channel group configuration
Please configure it in global configuration mode :

channel-group *channel-group-number*

Parameter “channel-group-number” is range from 0 to 16.

For example :

! Create a channel group with the group number being 0

QTECH(config)#channel-group 0

Delete channel group

no channel-group *channel-group-number*

Add add port members to the group

channel-group *channel-group-number* **mode** {active | passive | on}

In interface configuration mode, add current interface to channel group and specify the mode of interface. If the channel group doesn't exist, create it.

For example :

! Add Ethernet 0/0/3 to channel-group 3 and specify the port to be active mode

```
QTECH(config-if-ethernet-0/0/3)#channel-group 3 mode active
```

Delete interface member in channel group

no channel-group *channel-group-number*

In interface configuration mode, delete current interface from channel group.

For example :

! Delete interface Ethernet 0/0/3 from channel group 3

```
QTECH(config-if-ethernet-0/0/3)#no channel-group 3
```

Configure load balance of switch

channel-group load-balance {dst-ip|dst-mac|src-dst-ip|src-dst-mac|src-ip|src-mac}

For example :

! Specify load-balance of channel-group 0 is destination mac

```
QTECH(config)#channel-group load-balance dst-mac
```

Configure system LACP priority

lACP system-priority *priority*

For example :

! Configure LACP system priority is 40000

```
QTECH(config)#lACP system-priority 40000
```

Delete system LACP priority

no lACP system-priority

Use this command to restore system default LACP priority to be 32768.

Configure interface LACP priority

lACP port-priority *priority*

Use this command in interface configuration mode to configure LACP priority of the current interface

For example :

! Configure lACP port-priority of Ethernet 0/0/2 to be 12345

```
QTECH(config-if-ethernet-0/0/2)#lACP port-priority 12345
```

Delete interface LACP priority

no lACP port-priority

Use this command to restore interface default LACP priority to be 128.

Display system LACP ID

show lACP sys-id

System id is in the form of 16 characters of system priority and 32 characters of system MAC address.

For example :

! Display lACP system id

```
QTECH(config)#show lACP sys-id
```

Display local information of channel group

show lACP internal [*channel-group-number*]

Use show lACP interval command to display the information of group members, if the there is no keywords, all groups are displayed.

For example : Display the member information of channel group 2.

```
QTECH#show lacp internal 2
```

Display information of neighbour interface of channel group

show lacp neighbor [*channel-group-number*]

Use show lacp neighbor command to display the information of the neighbour port in the group. If there is no keyword, the neighbor ports of all the groups are displayed.

For example : Display the information of the neighbour port of the group 2

```
QTECH#show lacp neighbor 2
```

4.8 Interface BPDU-rate configuration

4.8.1 Brief introduction of interface CAR

Interface CAR is used to restrict the speed rate of BPDU packets with MAC 01 : 80 : C2 : xx : xx : xx impacted CPU of single interface. CPU can make speed rate statistics of each interface. If the speed rate is larger than the configured threshold (it is defaulted to be 300 packet/second), disable this interface and send trap of interface being abnormal. After a certain time (it is defaulted to be 480 seconds), re-enable the interface. If this interface will not be re-disabled by interface CAR in 2 seconds, the storm of impacting CPU by interface is over, and the interface recovers, and sends the trap of interface being normal. Caution : If the re-enabled interface is disable again by impacting CPU packet in 2 seconds, no trap of interface being abnormal is sent.

4.8.2 Port CAR configuration command list

Port CAR configuration command includes :

- Enable/disable interface CAR globally
- Enable/disable interface CAR on a port
- Configure interface CAR re-enable time
- Configure interface CAR
- Display interface CAR status

4.8.3 Enable/disable interface globally

Configure it in global configuration mode
Enable global interface

port-car

Disable global interface

no port-car

By default, port-car globally enables

For example :

! Enable port-car globally

```
QTECH(config)#port-car
```

4.8.4 Enable/disable interface CAR on interface

Please configure it in interface configuration mode :
Enable interface CAR

port-car

Disable interface CAR

no port-car

For example :

! Enable port-car of Ethernet 0/0/8

```
QTECH(config-if-ethernet-0/0/8)#port-car
```

4.8.5 Configure the reopen time of the port shutdown by

port-car

Please configure it in global configuration mode :

Configure the reopen time of the port shutdown by port-car

port-car-open-time *time*

By default, port-car-open-time is 480 seconds

For example :

! Configure port-car-open-time to be 10 seconds

```
QTECH(config)#port-car-open-time 10
```

4.8.6 Configure the port-car-rate

Please configure it in global configuration mode :

Configure the port-car-rate

port-car-rate *rate*

Default port-car-rate is 300 packet/second

For example :

! Configure port-car-rate to be 200 packet/second

```
QTECH(config)#port-car-rate 200
```

4.8.7 Display port-car information

Input following command in any configuration mode to display port-car information :

show port-car

For example :

! Display port-car information

```
QTECH(config)#show port-car
```

4.9 Port Alarm Configuration

4.9.1 Brief introduction of port alarm configuration

System can monitor port packet receiving rate. If the rate of receiving packet is beyond the interface flow exceed threshold, send alarm of large interface flow and the interface is in the status of large interface flow. In this status, if the rate of receiving packet is lower than the interface flow normal threshold, send alarm of normal interface flow. This function can actively report the rate of receiving packet to user.

4.9.2 Port alarm configuration list

Port alarm configuration command includes :

- Enable/disable port alarm globally
- Enable/disable port alarm on the port
- Configure the exceed threshold and normal threshold of port alarm
- Display port alarm

4.9.3 Enable/disable port alarm globally

Please configure it in global configuration mode :
Enable port alarm globally

alarm all-packets

Disable port alarm globally

no alarm all-packets

By default, alarm all-packets enable.

For example :

! Enable global alarm all-packets

QTECH(config)#alarm all-packets

4.9.4 Enable/disable port alarm on the port

Please configure it in interface configuration mode :
Enable port alarm on the port

alarm all-packets

Disable port alarm on the port

no alarm all-packets

For example :

! Enable alarm all-packets of Ethernet 0/0/8

QTECH(config-if-ethernet-0/0/8)# alarm all-packets

4.9.5 Configure the exceed threshold and normal threshold of port alarm

Configure the exceed threshold and normal threshold of port alarm

alarm all-packets threshold [*exceed rate*] [*normal rate*]



Caution : Exceed > normal. By default, 100 BASE exceed threshold is 85, normal threshold is 60

For example :

! Configure alarm all-packets exceed threshold to be 500, and normal threshold to be 300

QTECH(config)#alarm all-packets threshold exceed 500 normal 300

4.9.6 Display port alarm

Input following command in any configuration mode to display global interface alarm :

show alarm all-packets

For example :

! Display global alarm all-packets information

QTECH(config)#show alarm all-packets interface ethernet 0/0/1

Input following command in any configuration mode to display interface alarm on the port :

show alarm all-packets interface [*interface-list*]

Keyword “interface-list” is alternative. If there is no keyword, the alarm all-packets of all the interfaces are displayed, or the information of specified port is displayed.

For example :

! Display the alarm all-packets interface information of Ethernet 0/0/1

```
QTECH(config)#show alarm all-packets interface ethernet 0/0/1
```

```
e0/0/1 port alarm information
```

```
Port alarm status          :  enable
```

```
Port alarm exceed threshold(Mbps)  :  85
```

```
Port alarm normal threshold(Mbps)  :  60
```

```
Total entries :  1.0
```

4.10 Shutdown-control feature

Interface shutdown-control is used to restrict the speed rate of unicast\ multicast\broadcast of single interface. If the rate is beyond the configured restricted value (that can be configured) the interface will be shut down and failure trap will be sent. After a while (it is defaulted to be 480 seconds, which can be configured) it may reopen, or may be reopened by manual. If the interface will not reshutdown-control in 2 seconds, it turns normal and normal trap will be sent. If the interface reshutdown-control in 2 seconds, the failure trap will not be sent.

4.11 Interface shutdown-control configuration list

Interface shutdown-control configuration list is as following :

- Configuration mode and time
- Configuration shutdown-control
- Configure shutdown-control open-time
- Display shutdown-control

4.11.1 Configuration mode and time

Configure it in global configuration mode.

shutdown-control-recover {automatic-open-time *seconds*, mode [automatic, manual]}

seconds mean time for unshutdown interface

Use automatic or manual mode for control port shutdown.

4.11.2 Configuration interface shutdown-control

Configure it in interface configuration mode :

Enable shutdown-control

shutdown-control [broadcast | multicast | unicast] *target-rate*

Disable shutdown-control

no shutdown-control [broadcast | multicast | unicast]

By default, shutdown-control is disabled.

Example :

! Enable shutdown-control of e0/0/8 for broadcast and speed rate is 100pps.

QTECH(config-if-ethernet-0/0/8)#shutdown-control broadcast 100

4.11.3 Display shutdown-control

Configure it in any configuration mode :

show shutdown-control

Example :

! Display interface shutdown-control information

QTECH(config)#show shutdown-control

4.12 Port isolation configuration

Forbid intercommunication of users in different interfaces by port isolation configuration.

There are two kinds of interfaces in port isolation function. One is uplink port, and the other is downlink port. Uplink port can transmit any packet, but downlink port can only transmit the packet whose destination is uplink port. Connect user's computer to downlink port, and advanced devices connect to uplink port to shield intercommunication bwtween users and not influence user accessing exterior network through advanced switching devices.

Use port-isolation command in global configuration mode to add a or a group of descendent isolation port. Use no port-isolation command to remove a or a group of descendent isolation port :

Add port isolation downlink port

port-isolation { *interface-list* }

Delete port isolation downlink port

no port-isolation { *interface-list* | all }

interface-list is the optioned interface list which means one or more Ethernet interfaces. When adding port isolation downlink ports, not all ports can be added to be port isolation downlink ports. Choose all only when delete port isolation downlink ports. Choose "all" to remove all downlink isolation ports. By default, all ports are port isolation uplink ports.

For example :

! Add Ethernet 0/0/1, Ethernet 0/0/3, Ethernet 0/0/4, Ethernet 0/0/5, Ethernet 0/0/8 to be downlink isolation port.

QTECH(config)#port-isolation ethernet 0/0/1 ethernet 0/0/3 to ethernet 0/0/5 ethernet 0/0/8

! Remove ethernet 0/0/3, Ethernet 0/0/4, Ethernet 0/0/5, ethernet 0/0/8 from downlink isolation port.

QTECH(config)#no port-isolation ethernet 0/0/3 to ethernet 0/0/5 ethernet 0/0/8

4.13 Storm control configuration

Restrict the speed rate of port receiving broadcast, known multicast/ unknown unicast packets by storm control configuration.

Use storm-control command in interface configuration mode to configure storm-control. Use show interface command to display storm-control information.

Configure the speed rate of storm control

storm-control rate *target-rate*

Enable storm control

storm-control { broadcast | multicast | dlf }

Disable storm control

no storm-control { broadcast | multicast | dlf }

For example :

! Configure storm control of e0/0/1 with the speed rate being 2Mbps

QTECH(config-if-ethernet-0/0/1)#storm-control rate 2048

! Enable known multicast storm control of e0/0/1

QTECH(config-if-ethernet-0/0/1)#storm-control multicast

! Configure known multicast storm control of e0/0/3 with the speed rate being 5Mbps

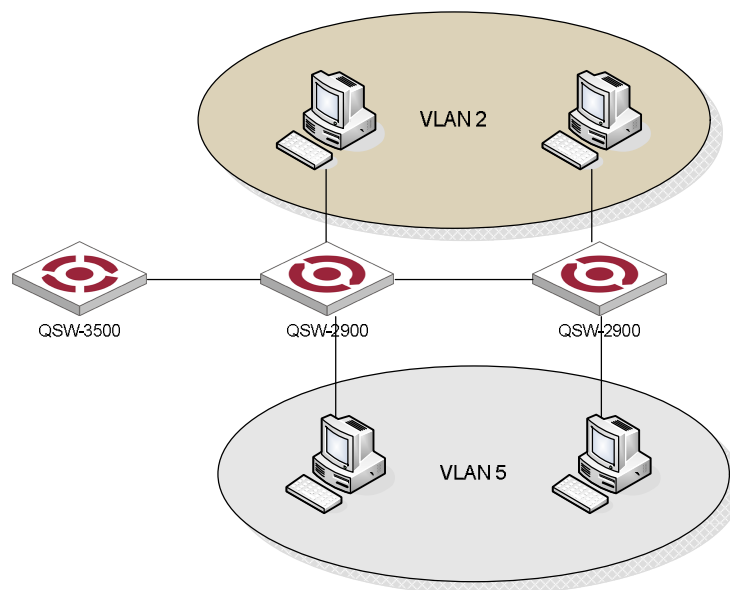
QTECH(config-if-ethernet-0/0/3)#storm-control multicast 5120

Chapter 5 VLAN Configuration

5.1 Introduction to VLAN

5.1.1 VLAN Overview

Ethernet is a network technology based on the Carrier Sense Multiple Access/Collision Detect (CSMA/CD) mechanism. As the medium is shared in an Ethernet, network performance may degrade as the number of hosts on the network is increasing. If the number of the hosts in the network reaches a certain level, problems caused by collisions, broadcasts, and so on emerge, which may cause the network operating improperly. In addition to the function that suppresses collisions (which can also be achieved by interconnecting LANs), virtual LAN (virtual LAN) can also isolate broadcast packets. VLAN divides a LAN into multiple logical LANs with each being a broadcast domain. Hosts in the same VLAN can communicate with each other like in a LAN. However, hosts from different VLANs cannot communicate directly. In this way, broadcast packets are confined to a single VLAN, as illustrated in the following figure.



VLAN diagram

A VLAN is not restricted by physical factors, that is to say, hosts that reside in different network segments may belong to the same VLAN, users in a VLAN can be connected to the same switch, or span across multiple switches or routers.

VLAN technology has the following advantages :

- 1) Broadcast traffic is confined to each VLAN, reducing bandwidth utilization and improving network performance.
- 2) LAN security is improved. Packets in different VLANs cannot communicate with each other directly. That is, users in a VLAN cannot interact directly with users in other VLANs, unless routers or Layer 3 switches are used.
- 3) A more flexible way to establish virtual working groups. With VLAN technology, clients can be allocated to different working groups, and users from the same group do not have to be within the same physical area, making network construction and maintenance much easier and more flexible.

5.1.2 VLAN Fundamental

To enable packets being distinguished by the VLANs they belong to, a field used to identifying VLANs is added to packets. As common switches operate on Layer 2, they only process Layer 2 encapsulation information and the field thus needs to be inserted to the Layer 2 encapsulation information of packets.

The format of the packets carrying the fields identifying VLANs is defined in IEEE 802.1Q, which is issued in 1999.

In the header of a traditional Ethernet packet, the field following the destination MAC address and the source MAC address is protocol type, which indicates the upper layer protocol type. Figure 2 illustrates the format of a traditional Ethernet packet, where DA stands for destination MAC address, SA stands for source MAC address, and Type stands for upper layer protocol type.



Figure 2 The format of a traditional Ethernet packet

IEEE802.1Q defines a four-byte VLAN Tag field between the DA&SA field and the Type field to carry VLAN-related information, as shown in Figure 3.

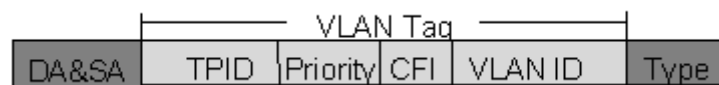


Figure 3 The position and the format of the VLAN Tag field

The VLAN Tag field comprises four sub-fields : the TPID field, the Priority field, the CFI field, and the VLAN ID field.

- The TPID field, 16 bits in length and with a value of 0x8100, indicates that a packet carries a VLAN tag with it.
- The Priority field, three bits in length, indicates the priority of a packet. For information about packet priority, refer to QoS Configuration in QoS Volume.
- The CFI field, one bit in length, specifies whether or not the MAC addresses are encapsulated in standard format when packets are transmitted across different medium. This field is not described here.
- The VLAN ID field, 12 bits in length and with its value ranging from 0 to 4095, identifies the ID of the VLAN a packet belongs to. As VLAN IDs of 0 and 4095 are reserved by the protocol, the actual value of this field ranges from 1 to 4094.

A network device determines the VLAN to which a packet belongs to by the VLAN ID field the packet carries. The VLAN Tag determines the way a packet is processed.

5.1.3 VLAN Classification

Based on different criteria, VLANs can be classified into different categories. The following types are the most commonly used :

- Port-based
- 802.1Q
- Policy-based
- Other types

This chapter will focus on the port-based VLANs and 802.1Q VLANs.

5.1.4 VLAN Interface

VLAN interfaces are virtual interfaces used for communications between different VLANs. Each VLAN can have one VLAN interface. Packets of a VLAN can be forwarded on network layer through the corresponding VLAN interface. As each VLAN forms a broadcast domain, a VLAN can be an IP network segment and the VLAN interface

can be the gateway to enable IP address-based Layer 3 forwarding.

5.1.5 Port-Based and 802.1Q VLAN

This is the simplest yet the most effective way of classifying VLANs. It groups VLAN members by port. After added to a VLAN, a port can forward the packets of the VLAN.

5.1.6 Port link type

Based on the tag handling mode, a port's link type can be one of the following three :

- Trunk port : the port can belong to multiple VLANs, can receive/send packets for multiple VLANs, normally used to connect network devices;
The differences between Access and Trunk port :
- A Access port allows packets of multiple VLANs to be sent with or without the Tag label;
- A Trunk port only allows packets with Tag label.

5.1.7 Default VLAN

You can configure the default VLAN for a port. By default, VLAN 1 is the default VLAN for all ports. However, this can be changed as needed.

- Ports PVID only belongs to one tag of VLAN. Therefore, its default VLAN is the VLAN it resides in and cannot be configured.
- You can configure the default VLAN for the Trunk port or the Access port as they can both belong to multiple VLANs.

5.1.8 Super VLAN

With the development of networks, network address resource has become more and more scarce. The concept of Super VLAN was introduced to save the IP address space. Super VLAN is also named as VLAN aggregation. A super VLAN involves multiple sub-VLANs. It has a VLAN interface with an IP address, but no physical ports can be added to the super VLAN. A sub-VLAN can have physical ports added but has no IP address and VLAN interface. All ports of sub-VLANs use the VLAN interface's IP address of the super VLAN. Packets cannot be forwarded between sub-VLANs at Layer 2.

If Layer 3 communication is needed from a sub-VLAN, it will use the IP address of the super VLAN as the gateway IP address. Thus, multiple sub-VLANs share the same gateway address and thereby save IP address resource.

The local Address Resolution Protocol (ARP) proxy function is used to realize Layer 3 communications between sub-VLANs and between sub-VLANs and other networks. It works as follows : after creating the super VLAN and the VLAN interface, enable the local ARP proxy function to forward ARP response and request packets.

5.1.9 VLAN interface type

System supports IEEE 802.1Q which possesses two types of VLAN interfaces. One is tagged, and the other is untagged.

Tagged interface can add VLAN ID, priority and other VLAN information to the head of the packet which is out of the interface. If the packet has included IEEE 802.1Q information when entering the switch, the mark information will not be changed; if the packet has not included IEEE 802.1Q mark information, system will determine the VLAN it belongs to according to the default VLAN ID of the receiving interface. Network devices

supported IEEE 802.1Q will determine whether or not to transmit this packet by the VLAN information in the mark.

Untagged interface can drop the mark information from all the packets which are out of the interface. When a frame is out of a untagged interface, it will not contain IEEE 802.1Q mark information. The function of dropping the mark makes the packet can be transferred from the network device supported mark to the one which doesn't support it.

Now, only the switch supported IEEE 802.1Q can be recognize IEEE 802.1Q frame so only a port linking to a switch supported IEEE 802.1Q can be configured to be Tagged port.

5.1.10 Default VLAN

There is a default VLAN of production, which possesses following features :

- The name of this VLAN is Default which can be modified.
- It includes all ports which can be added and deleted.
- All the port mode of default VLAN is untagged which can be modified to be tagged.
- VLAN ID of default VLAN is 1 which cannot be deleted.

5.2 VLAN configuration list

Configure VLAN should create VLAN according to the need first, then configure VLAN interface and its parameter.

VLAN configuration list is as following :

- Create/delete VLAN
- Add/delete VLAN interface
- Specify/delete VLAN description
- Configure interface type
- Configure interface default vlan ID
- Configure tag vlan
- Display VLAN information

5.2.1 Create/delete VLAN

Configure it in global configuration mode :

Enter VLAN configuration mode or create VLAN and enter it

vlan *vlan-list*

Delete created VLAN or specified VLAN except VLAN 1

no vlan { *vlan-list* | all }

VLAN-ID allowed to configure by system is in the range of 1 to 4094. *vlan-list* can be in the form of discrete number, a sequence number, or the combination of discrete and sequence number, discrete number of which is separate by comma, and sequence number of which is separate by subtraction sign, such as : 2, 5, 8, 10-20. Use the *vlan* command to enter VLAN configuration mode. If the *vlan* identified by the *vlan-id* keyword exists, enter VLAN configuration mode. If not, this command creates the VLAN and then enters VLAN configuration mode. For example, if VLAN 2 is not existed, system will create VLAN 2 first, then enter VLAN configuration mode; if VLAN 2 has existed, enter VLAN configuration mode.

When deleting VLAN, if the *vlan-list* is specified, delete corresponding VLAN. If choosing all, delete all existed VLAN except default VLAN. If deleting interface in VLAN, and default VLAN id is the same as the VLAN to be deleted, restore interface default VLAN ID to be default VLAN ID.isted VLAN except default VLAN. orresponding VLAN. has existed, enter VLAN configuration mode.. errperpp

If the VLAN to be removed exists in the multicast group, remove the related multicast group first.

5.2.2 Add/delete VLAN interface

Use the switchport command to add a port or multiple ports to current VLAN. Use the no switchport command to remove a port or multiple ports from current VLAN. Use following commands in VLAN configuration mode :

Add interface to specified VLAN

switchport { *interface-list* | all }

Delete some interface from specified VLAN

no switchport { *interface-list* | all }

Interface-list is the optioned interface list which means a or more interfaces. If choose all, add all ports to current VLAN; if choosing all when deleting interface, all ports in current VLAN will be deleted. When deleting interface from VLAN 1, if the PVID of interface is 1, modify the PVID to be other VLAN ID before deleting this interface. When deleting interface in other VLAN ID, port PVID should be the same as the VLAN ID, and the port is also in VLAN 1, delete it. If this port is not in VLAN 1, modify port PVID to be other VLAN ID, delete the port. There are two status of the interface in VLAN, one is tagged and the other is untagged. If the port is access port, add it to VLAN with the status of being untagged. If it is trunk port, change it to be tagged in VLAN.

For example :

! Add Ethernet 1, 3, 4, 5, 8 to current VLAN

```
QTECH(config-if-vlan)#switchport ethernet 0/0/1 ethernet 0/0/3 to ethernet 0/0/5 ethernet 0/0/8
```

! Remove Ethernet 3, 4, 5, 8 from current VLAN

```
QTECH(config-if-vlan)#no switchport ethernet 0/0/3 to ethernet 0/0/5 ethernet 0/0/8
```

Command switchport access vlan and its no command can also add and delete port to or from VLAN. Please refer to interface configuration of chapter 2.

5.2.3 Specify/restore VLAN description

The description string is used to distinguish each VLAN. Please configure it in VLAN configuration mode : Specify a description string to specified VLAN

description *string*

Delete description string of specified VLAN

no description

string : It is in the range of 1 to 32 characters to describe the current VLAN. The characters can be printable, excluding such wildcards as '/', ':', '*', '?', '\\', '<', '>', '|', ''etc.

For example :

! Specify the description string of the current VLAN as “market”

```
QTECH (config-if-vlan)#description market
```

! Delete the description string of VLAN

```
QTECH(config-if-vlan)#no description
```

5.2.4 Configure interface type

Use switchport mode command to configure port type. Please refer to [interface configuration](#).

5.2.5 Configure interface default vlan ID

System supports IEEE 802.1Q. When receiving a untagged packet, system would add a tag to the packet, in which the VLAN ID is determined by the default VLAN ID of the receiving port. The command to configure default VLAN of trunk port is :

switchport trunk native vlan *vlan-id*

For access port, use command to configure default VLAN of specified interface :

switchport access vlan *vlan-id*

The detailed introduction of the corresponding command.

For example :

! Configure default vlan-id of Ethernet interface 1 to be 2

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

```
QTECH(config-if-ethernet-0/0/1)#switchport access vlan 2
```



Caution : To use **switchport trunk native vlan** *vlan-id* must guarantee the specified interface to be trunk, and belongs to specified VLAN, and the VLAN ID is not 1. Use **switchport access vlan** *vlan-id* to configure interface default VLAN and add it to the VLAN. The specified interface is access, and the VLAN is existed and is not the default VLAN.

5.2.6 Configure tag vlan

When the port is access without tag vlan configuration, it can only send untagged packet. If it wants to send tagged packet, use command :

tag vlan *vlan-list*

Use command to disable this function

no tag vlan *vlan-list*

The interface must be access, and configure it in interface configuration mode.

For example :

! Configure Ethernet interface 1 to send IEEE 802.1Q packet with tag VLAN 5, VLAN 7-10

```
QTECH(config-if-ethernet-0/0/1)#tag vlan 5, 7-10
```

5.2.7 Display VLAN information

VLAN information is VLAN description string, vlan-id, VLAN status and interface members in it, tagged interfaces, untagged interfaces and dynamic tagged interfaces. Interface members consist of tagged and untagged members.

show vlan [*vlan-id*]

If the VLAN with specified keyword exists, this command displays the information of the specified VLAN. If no keyword is specified, this command displays the list of all the existing VLANs

For example :

! Display the information of existed VLAN 2.

```
QTECH(config)#show vlan 2
```

5.3 Brief introduction of GVRP

5.3.1 GARP protocol

The **Generic Attribute Registration Protocol (GARP)** was defined by the IEEE to provide a generic framework so bridges (or other devices like switches) could register and de-register attribute values, such as VLAN identifiers and multicast group membership. GARP defines the architecture, rules of operation, state machines and variables for the registration and de-registration of attribute values. GARP is the protocol was used by two applications : GARP VLAN Registration Protocol (GVRP) for registering VLAN trunking between multilayer switches, and by the GARP Multicast Registration Protocol (GMRP). The latter two were both mostly enhancements for VLAN aware switches, which requires IEEE 802.1Q.

Generic Attribute Registration Protocol (GARP) provides a mechanism that allows participants in a GARP application to distribute, propagate, and register with other participants in a bridged LAN the attributes specific to the GARP application, such as the VLAN or multicast address attribute.

GARP itself does not exist on a device as an entity. GARP-compliant application entities are called GARP applications. One example is GVRP. When a GARP application entity is present on a port on your device, this port is regarded a GARP application entity.

5.3.2 Brief introduction of GVRP

Multiple VLAN Registration Protocol (MVRP) formerly known as **GARP VLAN Registration Protocol (GVRP)** is a standards-based Layer 2 network protocol, for automatic configuration of VLAN information on switches. It was defined in the 802.1ak amendment to 802.1Q-2005.

Within a layer 2 network, MVRP provides a method to dynamically share VLAN information and configure the needed VLANs. For example, in order to add a switch port to a VLAN, only the end port, or the VLAN-supporting network device connected to the switchport, need be reconfigured, and all necessary VLAN trunks are dynamically created on the other MVRP-enabled switches. Without MVRP, (or the similar Cisco-proprietary protocol VTP) manual configuration of VLAN trunks is necessary.

It is through MVRP that Dynamic VLAN entries will be updated in the Filtering Database. In short, MVRP helps to maintain VLAN configuration dynamically based on current network configurations.

802.1Q allows for :

1. Dynamic configuration and distribution of VLAN membership information by means of the MVRP
2. Static configuration of VLAN membership information via Management mechanisms, which allow configuration of Static VLAN Registration Entries.
3. Combined static and dynamic configuration, in which some VLANs are configured via Management mechanisms and for other VLANs, MVRP is relied on to establish the configuration.

5.3.3 GARP messages and timers

1) GARP messages

GARP participants exchange attributes primarily by sending the following three types of messages :

- Join to announce the willingness to register some attribute with other participants.
- Leave to announce the willingness to deregister with other participants. Together with Join messages, Leave messages help GARP participants complete attribute reregistration and deregistration.
- LeaveAll to deregister all attributes. A LeaveAll message is sent upon expiration of a LeaveAll timer, which starts upon the startup of a GARP application entity.
- Through message exchange, all attribute information that needs registration propagates to all GARP participants throughout a bridged LAN.

2) GARP timers

GARP sets interval for sending GARP messages by using these four timers :

- Hold timer — When a GARP application entity receives the first registration request, it starts a hold timer and collects succeeding requests. When the timer expires, the entity sends all these requests in one Join message. This can thus help you save bandwidth.
- Join timer — Each GARP application entity sends a Join message twice for reliability sake and uses a join timer to set the sending interval.
- Leave timer — Starts upon receipt of a Leave message sent for deregistering some attribute information. If no Join message is received before this timer expires, the GARP application entity removes the attribute information as requested.
- LeaveAll timer — Starts when a GARP application entity starts. When this timer expires, the entity sends a LeaveAll message so that other entities can re-register its attribute information. Then, a LeaveAll timer starts again.

 Note :

- The settings of GARP timers apply to all GARP applications, such as GVRP, on a LAN.
- Unlike other three timers, which are set on a port basis, the LeaveAll timer is set in system view and takes effect globally.
- A GARP application entity may send LeaveAll messages at the interval set by its LeaveAll timer or the LeaveAll timer on another device on the network, whichever is smaller. This is because each time a device on the network receives a LeaveAll message it resets its LeaveAll timer.

5.3.3.1 Operating mechanism of GARP

The GARP mechanism allows the configuration of a GARP participant to propagate throughout a LAN quickly. In GARP, a GARP participant registers or deregisters its attributes with other participants by making or withdrawing declarations of attributes and at the same time, based on received declarations or withdrawals handles attributes of other participants.

GARP application entities send protocol data units (PDU) with a particular multicast MAC address as destination. Based on this address, a device can identify to which GVRP application, GVRP for example, should a GARP PDU be delivered.

5.3.3.2 GARP message format

The following figure illustrates the GARP message format.

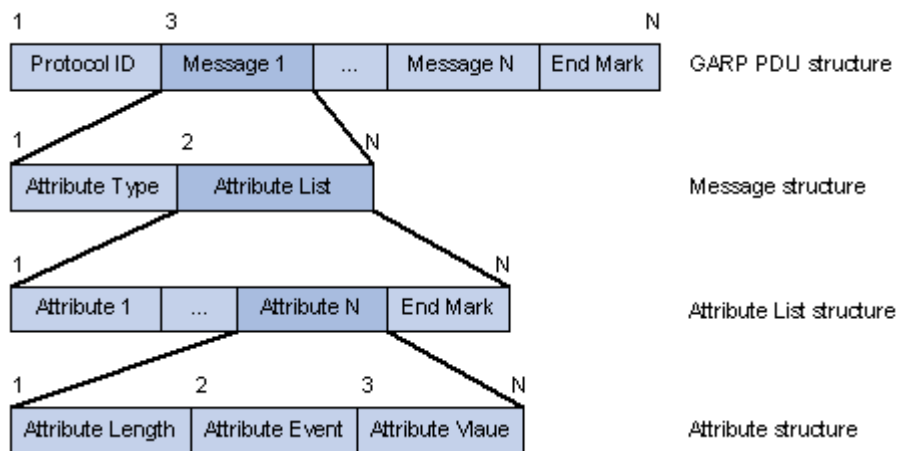


Figure 1 GARP message format

The following table describes the GARP message fields.

Table 1 Description on the GARP message fields

Field	Description	Value
Protocol ID	Protocol identifier for GARP	1
Message	One or multiple messages, each containing an attribute type and an attribute list	—
Attribute Type	Defined by the concerned GARP application	0x01 for GVRP, indicating the VLAN ID attribute
Attribute List	Contains one or multiple attributes	—
Attribute	Consists of an Attribute Length, an Attribute Event, and an Attribute Value	—
Attribute Length	Number of octets occupied by an attribute, inclusive of the attribute length field	2 to 255 (in bytes)
Attribute Event	Event described by the attribute	0 : LeaveAll 1 : JoinEmpty 2 : JoinIn 3 : LeaveEmpty 4 : LeaveIn 5 : Empty
Attribute Value	Attribute value	VLAN ID for GVRP If the Attribute Event is LeaveAll, Attribute Value is omitted.
End Mark	Indicates the end of PDU	—

5.4 GVRP Configuration list

In all configurations, enable global GVRP first before enable GVRP on a port. GVRP must be enabled in the two ends of trunk link which follows IEEE 802.1Q standard.

GVRP Configuration list is as following :

- Enable/disable global GVRP
- Enable/disable GVRP on a port
- Display GVRP
- Add/delete vlan that can be dynamic learnt by GVRP
- Display vlan that can be learnt by GVRP

5.4.1 Enable/disable global GVRP

Please configure it in global configuration mode :

Enable global GVRP

gvrp

Disable global GVRP

no gvrp

By default, GVRP globally disabled.

For example :

! Enable GVRP globally

```
QTECH(config)#gvrp
```

5.4.2 Enable/disable GVRP on a port

Please configure it in interface configuration mode :

Enable GVRP on a port

gvrp

Disable GVRP on a port

no gvrp

For example :

! Enable GVRP on Ethernet port 8

```
QTECH(config-if-ethernet-0/0/8)#gvrp
```



Caution : Enable global GVRP before enable GVRP on a port. By default, global GVRP deisables and GVRP on a port can be enabled in trunk mode interface.

5.4.3 Display GVRP

Use following command in any configuration mode to display global GVRP :

show gvrp

Use following command in any configuration mode to display GVRP on a port :

show gvrp interface [*interface-list*]

Interface-list keyword is optional. If this keyword unspecified, the command displays GVRP information for all the Ethernet ports. If specified, the command displays GVRP information on specified Ethernet port.

For example :

! Display GVRP information on interface Ethernet 0/0/1

```
QTECH(config)#show gvrp interface ethernet 0/0/1
```

5.4.4 Add/delete vlan that can be dynamic learnt by GVRP

Use garp permit vlan command to add configured static vlan to GVRP module for other switches to learn. Configure it in global configuration mode :

garp permit vlan *vlan-list***no garp permit vlan** [*vlan-list*]

For example : ! Add vlan 2, 3, 4 to GVRP

```
QTECH(config)#garp permit vlan 2-4
```

5.4.5 Display vlan that can be learnt by GVRP

Use show garp permit vlan command to display current static vlan permitted learning by GVRP

show garp permit vlan

For example :

Display current static vlan permitted learning by GVRP

```
QTECH(config)#show garp permit vlan
```

5.4.6 Examples for GVRP configuration

! Enable GVRP on Ethernet port 2

```
QTECH(config-if-ethernet-0/0/2)#gvrp
```

! Disable GVRP on Ethernet port 2

```
QTECH(config-if-ethernet-0/0/2)#no gvrp
```

5.5 Brief introduction of QinQ

QinQ is used for the communication between discrete client vlan whose service model is the interconnection of one or more switches supported QinQ by service provider interfaces which are in service provider vlan. The interface linking client vlan is called customer interface. Packet with client vlan tag will add a tag head with the vlan id being service provider vlan when passing through the customer interface. The tag head will be stripped when passing through service provider vlan.

5.5.1 Introduction to QinQ

In the VLAN tag field defined in IEEE 802.1Q, only 12 bits are used for VLAN IDs, so a device can support a maximum of 4,094 VLANs. In actual applications, however, a large number of VLAN are required to isolate users, especially in metropolitan area networks (MANs), and 4,094 VLANs are far from satisfying such requirements.

The port QinQ feature provided by the device enables the encapsulation of double VLAN tags within an Ethernet frame, with the inner VLAN tag being the customer network VLAN tag while the outer one being the VLAN tag assigned by the service provider to the customer. In the backbone network of the service provider (the public network), frames are forwarded based on the outer VLAN tag only, while the customer network VLAN tag is shielded during data transmission.

Figure 1 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 in the MAN.

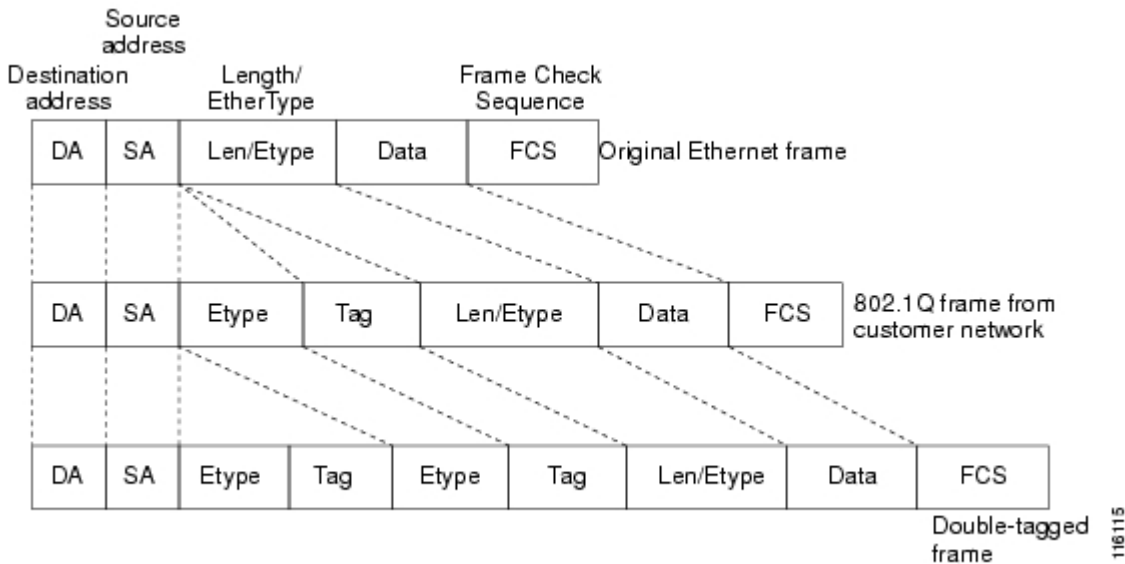


Figure 1 802.1Q-tagged frame structure vs. double-tagged Ethernet frame structure

Advantages of QinQ :

- Addresses the shortage of public VLAN ID resource
- Enables customers to plan their own VLAN IDs, with running into conflicts with public network VLAN IDs.
- Provides a simple Layer 2 VPN solution for small-sized MANs or intranets.

Note : The QinQ feature requires configurations only on the service provider network, and not on the customer network.

5.5.2 Implementations of QinQ

There are two types of QinQ implementations : basic QinQ and selective QinQ.

1) Basic QinQ

Basic QinQ is a port-based feature, which is implemented through VLAN VPN.

With the VLAN VPN feature enabled on a port, when a frame arrives at the port, the port will tag it with the port's default VLAN tag, regardless of whether the frame is tagged or untagged. If the received frame is already tagged, this frame becomes a double-tagged frame; if it is an untagged frame, it is tagged with the port's default VLAN tag.

2) Selective QinQ

Selective QinQ is a more flexible, VLAN-based implementation of QinQ

5.5.3 Adjustable TPID Value of QinQ Frames

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.

Figure 2 shows the 802.1Q-defined tag structure of an Ethernet frame.

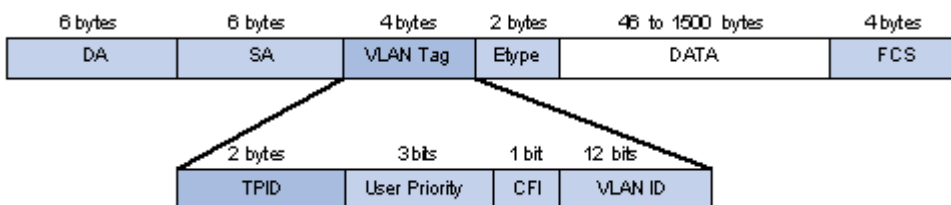


Figure 2 VLAN Tag structure of an Ethernet frame

On devices of different vendors, the TPID of the outer VLAN tag of QinQ frames may have different default values.

You can set and/or modify this TPID value, so that the QinQ frames, when arriving at the public network, carries the TPID value of a specific vendor to allow interoperation with devices of that vendor.

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

Table 1 Reserved 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
Cluster	0x88A7
Reserved	0xFFFD/0xFFFE/0xFFFF

5.6 QinQ configuration list

- Configure global QinQ
- Configure interface QinQ mode
- Configure interface dynamic QinQ
- Enable/disable vlan-swap
- Configure interface switching vlan
- Display dynamic QinQ
- Display switching vlan

5.6.1 Configure global QinQ

QSW-3900 supports two QinQ :

- 1) Static QinQ. Vlan protocol number in this mode can be configured but cannot be configured to ignore tag head of ingress packet. If vlan protocol number is not the same as the port configuration value or the port is configured to ignore tag head, there will be a new tag head between the 12th and 13th bit;
- 2) Flexible QinQ. Configure port vlan protocol number and the ignorance attribution of the tag head of ingress port. Only when vlan protocol number of ingress packet is not the same as the port configuration value and not the default value 8100, a new tag head will be added. If egress is TAG, TPID of TAG head is configured TPID.

! Use dtag command to enable/disable QinQ globally in global configuration mode.

dtag { [flexible-qinq] | outer-tpid *tpid* }

no dtag

For example :

! Configure QinQ global TPID to be non dot1q-in-dot1q
QTECH(config)dtag outer-tpid 9100

5.6.2 Configure QinQ mode of interface

There are two kinds of interface modes : one is service provider port, the other is customer port. The former do not permit ignoring tag head of ingress packet and the latter permits.

! It is in the interface configuration mode.

dtag mode { customer | service-provider }

Example :

Configure interface to be customer

QTECH(config-if-ethernet-0/0/1)#dtag mode customer

5.6.3 Configure interface dynamic QinQ

1. Configure a series vlan to be dynamic QinQ with the start vlan and destination vlan. In the precondition of all vlan tag packets between start vlan are not transparent transmitted, they will transmit in the form of double tag head with destination vlan.

! The command mode is global configuration mode

dtag insert *startvlanid endvlanid targetvlanid*

Example :

Configure all vlan tag packets to add a tag head with destination vlan3 from the start vlan1 to end vlan2

QTECH(config-if-ethernet-0/0/1)#dtag insert 1 2 3

2. Delete a consecutive vlan in configured dynamic QinQ on the form of start vlan and destination vlan, in which the parameter imputed start vlan and the destination vlan must be the same as configuring a vlan series.

! The command mode is global configuration mode

no dtag insert *startvlanid endvlanid*

Example :

Delete all configured vlan tag packets to add a tag head with destination vlan3 from the start vlan1 to end vlan2.

QTECH(config)#no dtag insert 1 2 3

3. Configure a series vlan to be transparent transmitted in dynamic QinQ in the form of start vlan. All vlan tag packets can be transmitted from start vlan without adding new tag head because the priority of transparent transmission id superior than adding tag head, transparent transmission will not be influenced by svlan inset command.

! Command mode is global configuration mode

dtag pass-through *startvlanid endvlanid*

Example :

Configure all vlan tag packet to be transparent transmission from start vlan1 to end vlan2

QTECH(config-if-ethernet-0/0/1)#dtag pass-through 1 2

4. Delete all configured all vlan tag packet to be transparent transmission in the form of start vlan, in which the parameter imputed start vlan must be the same as configuring a vlan series.

! Command mode is global configuration mode

no dtag pass-through *startvlanid endvlanid*

Example :

Delete all configured all vlan tag packet to be transparent transmission from start vlan1 to end vlan2

```
QTECH(config-if-ethernet-0/0/1)#no dtag pass-through 1 2
```

5.6.4 Enable/disable vlan-swap

Configure it in global configuration mode :

Enable vlan-swap

vlan-swap

Disable vlan-swap

no vlan-swap

By default, vlan-swap is disabled.

Example :

! Enable vlan-swap

```
QTECH(config)#vlan-swap
```

5.6.5 Configure global vlan-swap

1. Configure vlan in the tag to be repaced by configured vlan

! Command mode is global configuration mode

vlan-swap [*original vlanID*] [*swap vlan ID*]

Example :

Configure vlan1 in tag head to be replaced by vlan2

```
QTECH(config)#vlan-swap vlan1 vlan2
```

2.Delete configured vlan swap parameter

! Command mode is global configuration mode

no vlan-swap [*original vlanID*] [*swap vlan ID*]

Example :

Delete configured vlan1 in tag to be repaced by vlan2

```
QTECH(config)#no vlan-swap vlan1 vlan2
```

5.6.6 Configure rewrite-outer-vlan

Configure rewrite-outer-vlan. After configuration, all packets from this port without inner vlan ID being specified range and with outer vlan ID being specified one(this condition can be optioned), the outer vlan ID will be modified to be new.

! Command mode is interface configuration mode

rewrite-outer-vlan *start-inner-vid end-inner-vid [outer-vlan outer-vid] new-outer-vlan new-outer-vid*

no rewrite-outer-vlan *start-inner-vid end-inner-vid [outer-vlan outer-vid]*

Example :

Configure rewrite-outer-vlan of e0/0/1 with inner vlan ID being the range of 1~50, outer vlan ID being 3 and new outer vlan ID being 100

```
QTECH(config-if-ethernet-0/0/1)# rewrite-outer-vlan 1 50 outer-vlan 3 new-outer-vlan 100
```

5.6.7 Display dynamic QinQ

1. Display dynamic vlan

! Command mode is global configuration mode

show dtag

Example :

Display QinQ

```
QTECH(config)#show dtag
```

2. Display transparent transmission vlan

! Command mode is global configuration mode

show dtag pass-through

Example :

Display transparent transmission vlan

```
QTECH(config)#show dtag pass-through
```

5.6.8 Display vlan-swap

Display vlan swap status

! Command mode is global configuration mode

show vlan-swap

Example :

Display vlan swap status

```
QTECH(config)#show vlan-swap
```

5.6.9 Display rewrite-outer-vlan

1. Display rewrite-outer-vlan

! Command mode is global configuration mode

show rewrite-outer-vlan

Example :

Display rewrite-outer-vlan

```
QTECH(config)#show rewrite-outer-vlan
```


Chapter 6 Layer 3 Configuration

6.1 Brief Introduction of Layer 3 switching

The major difference between the packet switching operation of a router and that of a Layer 3 switch is the physical implementation. In general-purpose routers, packet switching takes place using a microprocessor, whereas a Layer 3 switch performs this using application-specific integrated circuit (ASIC) hardware.

L3 switching can move traffic at wire speed and also provide layer 3 routing, which can remove the bottleneck from the network routers. This technology is based on the idea of "route once, switch many". L3 switching can make routing/switching decisions based on the following

- MAC source/destination address in a Data Link frame
- IP source/destination address in the Network layer header
- Port source/destination numbers in the Transport layer header

There is no performance difference between a layer 2 and a layer 3 switch because the routing/switching is all hardware based.

QTECH QSW-3900 is a GE Intelligent Routing Switch based on ASIC technology which can support transmission in both layer 2 and layer 3. The interaccessing of hosts in the same VLAN is the transmission in layer 2 and the interaccessing of hosts in the different VLAN is the transmission in layer 3.

6.2 Layer 3 Configuration list

Configuration list is as following :

- [VLAN division and the creation of layer 3 interface](#)
- [Transmission mode configuration](#)
- [Create VLAN interface for normal VLAN](#)
- [Create superVLAN interface and add VLAN to superVLAN](#)
- [Configure IP address for VLAN interface or superVLAN interface](#)
- [ARP proxy configuration](#)
- [Display interface configuration](#)

6.2.1 VLAN division and the creation of layer 3 interface

VLAN division please refers to VLAN configuration chapter.

Layer 3 interface includes normal VLAN interface and superVLAN interface. Normal VLAN interface is the interface in some concrete VLAN; superVLAN interface is created in superVLAN (superVLAN is the VLAN which is not existed and contains no interface) which can contain many subVLANs (subVLAN is the existed concrete VLAN). At most 258 layer 3 interfaces can be created, among which superVLAN can be 128 at most.

The total maximum number of VLAN contained by all layer 3 interfaces is 258. Each VLAN only exists in one layer 3 interface. In superVLAN, interface must be untagged member in only one subVLAN, and tagged in other subVLANs.

6.2.2 Transmission mode configuration

QTECH QSW-3900 supports two types of packet transmission mode : 1) flow transmission ;2) network topology transmission. Searching failed route or host route with the unreached destination in flow transmission mode;

these packet will be dropped in network topology transmission. It is defaulted to be flow transmission mode. Please configure it in global configuration mode :

[no] ip def cpu

6.2.3 Create VLAN interface for normal VLAN

Configure VLAN interface for each VLAN which supports layer 3 transmission or add this VLAN to superVLAN.

Create VLAN interface for VLAN 2 and enter VLAN interface configuration mode :

QTECH(config)#interface vlan-interface 2

6.2.4 Create superVLAN interface and add VLAN to superVLAN

SuperVLAN interface realizes the intercommunication of hosts which belong to different VLAN but the same network interface. superVLAN interface is realized through ARP proxy.

Create superVLAN 1 and add VLAN 3, VLAN 4 to be subVLAN of superVLAN 1.

QTECH(config)#interface supervlan-interface 1

QTECH(config-if-superVLANInterface-1)#subvlan 3

QTECH(config-if-superVLANInterface-1)#subvlan 4

Delete VLAN 3 and VLAN 4 from superVLAN 1.

QTECH(config-if-superVLANInterface-1)#no subvlan 3

QTECH(config-if-superVLANInterface-1)#no subvlan 4

6.2.5 Configure IP address for VLAN interface or superVLAN interface

At most 32 IP address can be configured for each VLAN interface or superVLAN interface the IP address of which cannot be in the same network interface. The IP address firstly configured will be the primary IP address. After deleting primary IP address, there will be another to be the primary IP address automatically and it can also configure an IP address to be the primary one manually. For example, if IP address of VLAN interface 1 is 10.11.0.0/16, other interfaces cannot configure the IP address in the same network interface (10.11.0.0/16), such as 10.11.1.1/24.

Configure IP address of VLAN interface 2 to be 10.11.0.0/16 :

QTECH(config-if-vlanInterface-2)#ip address 10.11.0.1 255.255.0.0

Delete IP address of VLAN interface 2 :

QTECH(config-if-vlanInterface-2)#no ip address

Specify an IP address of specified interface to be the primary IP address :

QTECH(config-if-vlanInterface-2)#ip address primary 10.11.0.1

6.2.6 Configure accessing IP address range of VLAN or superVLAN interface

At most 8 accessing range can be configured for each VLAN or superVLAN interface. After configuring accessing range, ARP must learn in this range to restrict user's accessing. When deleting VLAN or superVLAN interface, related configuration will be deleted.

Use following command in VLAN or superVLAN interface mode :

ip address range *startip endip*

6.2.7 ARP proxy configuration

ARP require packet is broadcasting packet which cannot go through VLAN. If ARP proxy enables, subVLANs of the same superVLAN can ARP exchanges. When ARP proxy disables, subVLANs of the same superVLAN cannot communicate.

Use following command in global configuration command :

arp-proxy

no arp-proxy

It is defaulted to disable ARP proxy.

For example :

!Enable ARP proxy

QTECH(config)#arp-proxy

!Disable ARP proxy

QTECH(config)#no arp-proxy

6.2.8 Display interface configuration

Each created VLAN or superVLAN interface has its own configuration information, including : VLAN number, IP address and netmask. Following command is used to display configuration information of all layer 3 interface, specified normal VLAN or superVLAN interface.

Display all layer 3 interface configuration information :

QTECH(config)#show ip interface

Display VLAN interface 2 configuration information :

QTECH(config)#show ip interface vlan-interface 2

Display superVLAN interface 3 configuration information :

QTECH(config)#show ip interface supervlan-interfac 3

6.3 Brief introduction of static routing

A static route is a special route that is manually configured by the network administrator. If a network's topology is simple, you only need configure static routes for the network to work normally. The proper configuration and usage of static routes can improve a network's performance and ensure bandwidth for important network applications.

The disadvantage of using a static route is that, if a fault or a topological change occurs to the network, the routes will be unavailable and the network breaks. In this case, the network administrator has to modify the static routes manually.

6.3.1 Default Route

A router selects the default route only when it cannot find any matching entry in the routing table.

If the destination address of a packet fails to match any entry in the routing table, the router selects the default route to forward the packet.

If there is no default route and the destination address of the packet fails to match any entry in the routing table, the packet will be discarded and an ICMP packet will be sent to the source to report that the destination or the network is unreachable.

You can create the default route with both destination and mask being 0.0.0.0, and some dynamic routing

protocols, such as OSPF, RIP and IS-IS, can also generate the default route.

6.3.2 Application Environment of Static Routing

Before configuring a static route, you need to know the following concepts :

1) Destination address and mask

In the **ip route** command, an IPv4 address is in dotted decimal format and a mask dotted decimal format.

2) Next hop address

While configuring a static route, you can specify next hop address. The next hop address can not be a local interface IP address; otherwise, the route configuration will not take effect.

In fact, all the route entries must have a next hop address. When forwarding a packet, a router first searches the routing table for the route to the destination address of the packet. The system can find the corresponding link layer address and forward the packet only after the next hop address is specified.

QTECH QSW-3900 is a GE Intelligent Routing Switch based on ASIC technology which maintains a layer 3 transmission routing table to designate the next hop address and related information which can be dynamically learnt and manually configured. Static routing is the route manually designated to some address.

6.4 Static routing configuration list

- [Add/delete static route](#)
- [Display route table information](#)

6.4.1 Add/delete static route

Use this command to add a route table item to designate the next hop transmission address when communication with some address. Destination address, netmask and next hop address must be designated. If the destination address and mask are all 0, the added route is default route.

ip route *ip-dest mask-dest nexthop*

Example :

QTECH(config)#ip route 192.168.0.100 255.255.255.255 10.11.0.254

Add a host route to 192.168.0.100 with the hop address being 10.11.0.254

QTECH(config)#ip route 192.168.0.100 255.255.255.255 10.11.0.254

Delete host route to 192.168.0.100, the next hop address may or may not be inputted, if it is inputted, it must be the same as that in real route table :

QTECH(config)#no ip route 192.168.0.100 255.255.255.255

6.4.2 Display route table information

Use following commands to display existed route table information or specified route.

Display all static route :

show ip route static

Example :

QTECH(config)#show ip route static

Display system core route :

QTECH(config)#show ip route

Display system core route from 192.168.0.1 to 192.168.0.255

QTECH(config)#show ip route 192.168.0.100 255.255.255.0

Chapter 7 VRRP

7.1 Overview

7.1.1 Background

With the development of the Internet, users have higher requirements on network reliability. It is very important to keep contact with other parts on a network especially for end users. Generally speaking, a host communicates with the external networks through the default gateway, as shown in Figure 1.

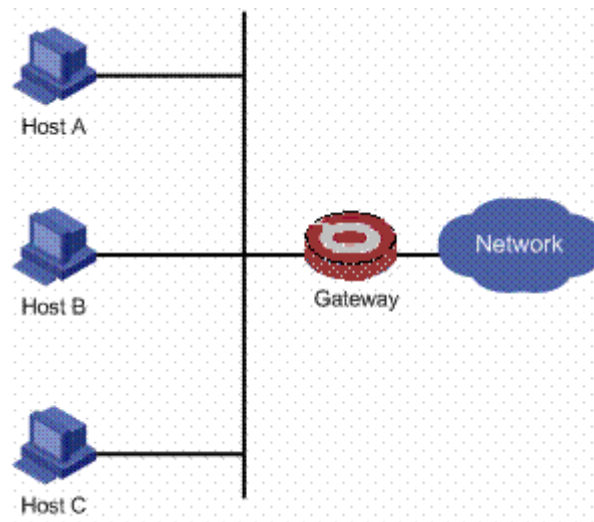


Figure 1 LAN networking

Normally, a host sends the packets to be sent to the external networks to the gateway, which then sends the packets to the external networks. This enables hosts on a network segment to communicate with the external networks. However, when the gateway fails, all the hosts using the gateway as the default next-hop router fail to communicate with the external networks. A common way to improve system reliability is to use more egress gateways. However, most hosts can only be configured with one default gateway. In case that a default gateway fails, you need to manually configure another default gateway for the hosts that originally use the failed gateway, so that the hosts can continue to communicate with the external networks. Some solutions require hosts to run a dynamic routing protocol such as Routing Information Protocol (RIP), or Open Shortest Path First routing protocol (OSPF) to solve the problem. However, these protocols cannot satisfy users' needs due to their complicated configuration or low security guarantee.

VRRP is thus addressed to solve the above-mentioned problems. VRRP does not change the original networking, nor does it require any configuration on the hosts. You only need to configure several commands on the related routers to implement backup of the gateway, without bringing any working load for the hosts. Compared with other methods, VRRP can better satisfy the users' needs.

7.1.2 Benefits

VRRP is an error-tolerant protocol. With VRRP deployed on a network, if the next-hop router of hosts fails,

another router will take it over to ensure continuous and reliable network communication.

VRRP has the following advantages:

- Simplified network management: Deploying VRRP on multicast and broadcast LANs such as Ethernet, you can ensure that the system can still provide highly reliable default links without changing configurations (such as dynamic routing protocols or route discovery protocols) when a device fails, and prevent network interruption due to a single link failure.
- High adaptability: A VRRP packet is encapsulated in an IP packet, and supports different kinds of upper layer protocols.
- Low network overhead: VRRP defines only one packet type, VRRP advertisement, and only the master in a VRRP group can send VRRP advertisements.

7.2 Introduction to VRRP

7.2.1 Concepts

Virtual router: It consists of a master and several backups. Every host on the LAN takes the virtual router as the default gateway.

VRID: Virtual router identifier. A group of routers with the same VRID form a virtual router.

Master: The router that forwards packets in a virtual router.

Backup: The router that can take the responsibility of the master when the master fails.

Virtual IP address: IP address of the virtual router. A virtual router can have one or multiple IP addresses.

IP address owner: The router whose interface IP address is the same as the virtual IP address.

Virtual MAC address: A virtual router has one virtual MAC address. The format of a virtual MAC address is 00-00-5E-00-01- $\{VRID\}$. Generally, a virtual router responds to an ARP request with its virtual MAC address, and only when special configurations are performed on a virtual router does it respond with the real MAC address of the interface.

Priority: VRRP determines the role (master or backup) of each router in a virtual router by priority.

Non-preemptive mode: The backup working in non-preemptive mode remains as a backup as long as the master does not fail. The backup will not become the master even if the former is configured with a higher priority.

Preemptive mode: The backup working in preemptive mode compares the priority in the packet with that of its own when a backup receives a VRRP advertisement. If its priority is higher than that of the master it preempts as the master; otherwise, it remains a backup.

7.2.2 Introduction to Virtual Router

VRRP combines a group of routers (including a master and multiple backups) on a LAN into a VRRP group. The VRRP group functions as a virtual router, and is identified by a virtual router ID. A virtual router (VRRP group) has the following features:

A virtual router has its own virtual IP address and MAC address. Every host on the LAN takes the IP address of the virtual router as its default gateway and communicates with the external networks through the virtual router.

A virtual router consists of multiple physical routers including a master and several backups. When the master works normally, the hosts on the LAN communicate with the external networks through the master; when the master fails, one of the backups becomes the master to forward packets, as shown in Figure 2.

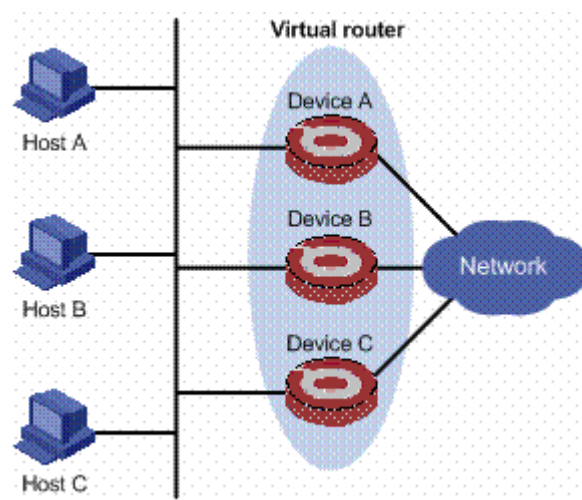


Figure 2 Network diagram for a virtual router

7.2.3 VRRP Working Process

VRRP works as the following:

- The routers in a virtual router elect the master based on their priorities. The master sends a gratuitous ARP packet to notify the devices or hosts connected to it of its virtual MAC address, thus taking on the responsibility to forward packets.
- The master sends VRRP advertisements periodically to advertise its configuration information (for example, its priority) and working status.
- If the master fails, the backups in the virtual router elect a new master based on their priorities.
- When the master in a virtual router is changed, the new master just sends a gratuitous ARP packet carrying the virtual router MAC address and IP address to update the ARP-related information of the hosts or devices connected to it; while the hosts in the network cannot be aware that the master has been changed to another one.

- If the priority of a backup is higher than that of the master, whether a new master needs to be elected depends on the working mode of the backup (preemptive or non-preemptive).

To sum up, to ensure normal working of the master and backups in a virtual router, VRRP needs to implement the following functions:

- Master election
- Master state advertisement
- Authentication to enhance security

7.2.3.1 Master Election

VRRP determines the role (master or backup) of each router in a virtual router by priority. A router with a higher priority has more opportunity to become the master.

A router in a virtual router works as a backup after it is created, and it gets the master priority by receiving VRRP advertisements.

- If the master priority in the VRRP advertisement is higher than the priority of the router, the router remains as a backup.
- If the master priority in the VRRP advertisement is lower than the priority of the router, when the router works in preemptive mode, it becomes the master to periodically send VRRP packets; when the router works in non-preemptive mode, it remains as a backup.
- If the router does not receive a VRRP advertisement in a certain period, it becomes the master.

VRRP priority is in the range of 0 to 255. A bigger number means a higher priority. Priorities 1 to 254 are configurable. Priority 0 is reserved for master that releases its master responsibility, and priority 255 for the IP address owner. When a router acts as the IP address owner, its priority is always 255. That is, if there is an IP address owner in a virtual router, it acts as the master as long as it works properly.

7.2.3.2 Master State Advertisement

The master in a virtual router sends VRRP advertisements periodically to inform the other routers in the virtual router of its configuration information (for example, priority) and working status. The backups judge whether the master works normally according to the advertisements received.

The master can release responsibility of a master by sending a VRRP advertisement with priority being 0 to trigger to trigger an immediate master election among backups. The time used for the election is called Skew time, in seconds, and is calculated as $((256 - \text{Priority})/256)$.

If the master fails and cannot send VRRP advertisements, a backup cannot know the state of the master immediately; it waits for a period of time, and if it still receives no advertisements from the master, it considers that the master fails and assumes itself as the master. If at this time, multiple backups compete for becoming the master, a master election is triggered. The time interval for the backups to declare master down is called `Master_Down_Interval`, in seconds, and is calculated as $(3 \times \text{Advertisement_Interval}) + \text{Skew time}$.

On an unstable network, a backup may fail to receive the packets from the master in `Master_Down_Interval` due to network congestion, thus causing the members in the virtual router to change their states frequently. This problem can be addressed through setting the VRRP preemption delay timer. With the VRRP preemption delay timer set, if a backup receives no advertisement in `Master_Down_Interval` and then the preemption delay, it considers that the master fails. In this case, it assumes itself as the master and sends VRRP advertisements.

7.2.3.3 Authentication Mode

VRRP provides three authentication modes:

- No authentication: No authentication is performed for any VRRP packet, without security guarantee.
- Simple text authentication: You can adopt the simple text authentication mode in a network facing possible security problems. A router sending a VRRP packet fills an authentication key into the packet, and the router receiving the packet compares its local authentication key with that of the received packet. If the two authentication keys are the same, the received VRRP packet is considered valid; otherwise, the received packet is considered an invalid one.
- MD5 authentication: You can adopt MD5 authentication in a network facing severe security problems. The router encrypts a VRRP packet to be sent using the authentication key and MD5 algorithm and saves the encrypted packet in the authentication header. The router receiving the packet uses the authentication key to decrypt the packet and checks whether the validity of the packet.

7.2.4 Backup's Monitoring of the Master State

In normal cases, a backup in a virtual router waits for `Master_Down_Interval` to become the master after the master fails. During this time, the hosts in the LAN cannot communicate as no master can forward packets for them. To solve the problem, VRRP provides the monitoring function for a backup to monitor the master state, making the backup become the new master immediately after the master fails to maintain network communication.

The BFD technology is adopted for a backup to monitor the master state. With this function enabled on a backup, the backup can automatically become the new master as soon as the master fails, with the Skew Time being reduced to milliseconds.

7.3 Application Scenarios

7.3.1 Master/Backup

In master/backup mode, only one router, the master, provides services. When the master fails, a new master is elected from the original backups to take the responsibility of the master, as shown in Figure 3 .

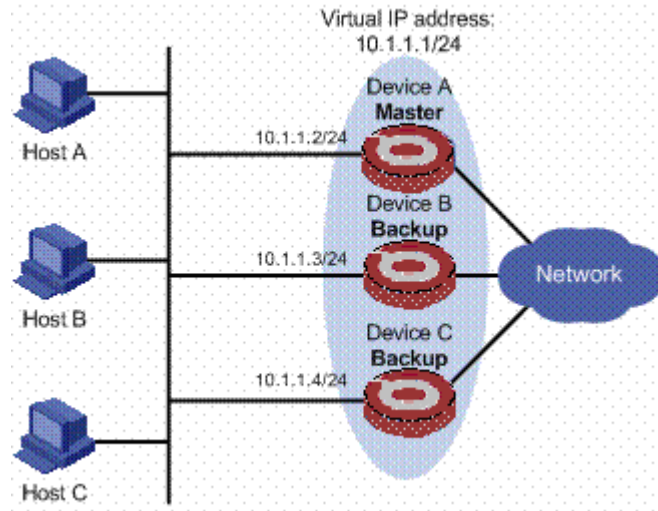


Figure 3 VRRP in master/backup mode

At the beginning, Device A is the master and therefore can forward packets to the external networks, while Device B and Device C are backups and are thus in the state of listening. If Device A fails, Device B and Device C will elect a new master according to their priorities. The new master takes over the forwarding task to provide services to the hosts on the LAN.

7.3.2 4.2 Load Balancing

You can create more than one virtual router on an interface of a router, allowing the router to be the master of one virtual router but a backup of another at the same time.

In load balancing mode, multiple routers provide services at the same time. This mode requires two or more virtual routers, each of which includes a master and one or more backups. The masters of the virtual routers can be assumed by different routers, as shown in Figure 4 .

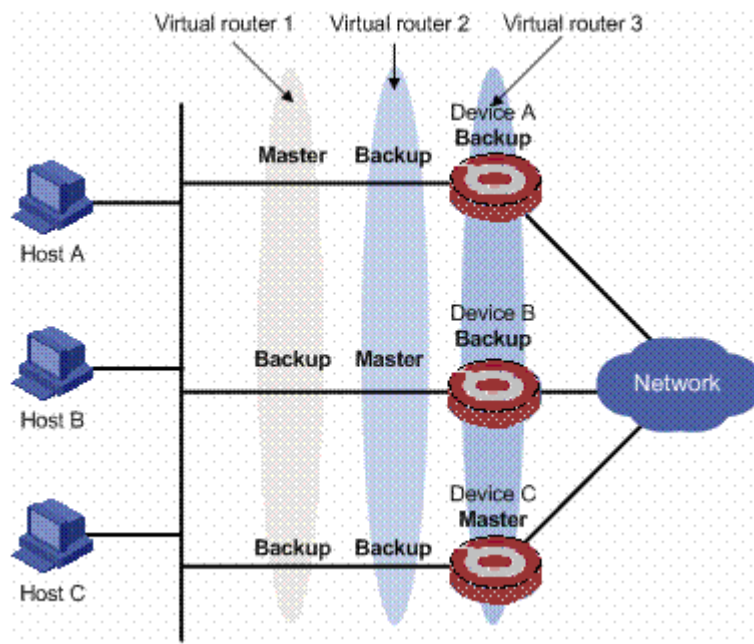


Figure 4 VRRP in load balancing mode

In Figure 4, three virtual routers are present:

- Virtual router 1: Device A is the master; Device B and Device C are the backups.
- Virtual router 2: Device B is the master; Device A and Device C are the backups.
- Virtual router 3: Device C is the master; Device A and Device B are the backups.

For load balancing among Device A, Device B, and Device C, hosts on the LAN need to be configured to use virtual router 1, 2, and 3 as the default gateways respectively. When configuring VRRP priorities, make sure that each router holds such a priority in each virtual router that it will take the expected role in the virtual router.

7.3.3 Master's Monitoring of Uplinks Through BFD/NQA

VRRP monitors the uplinks through BFD or NQA to make the master quickly find network faults and reduce its priority, thus ensuring a backup whose uplink is working normally to assume the responsibility of a master.

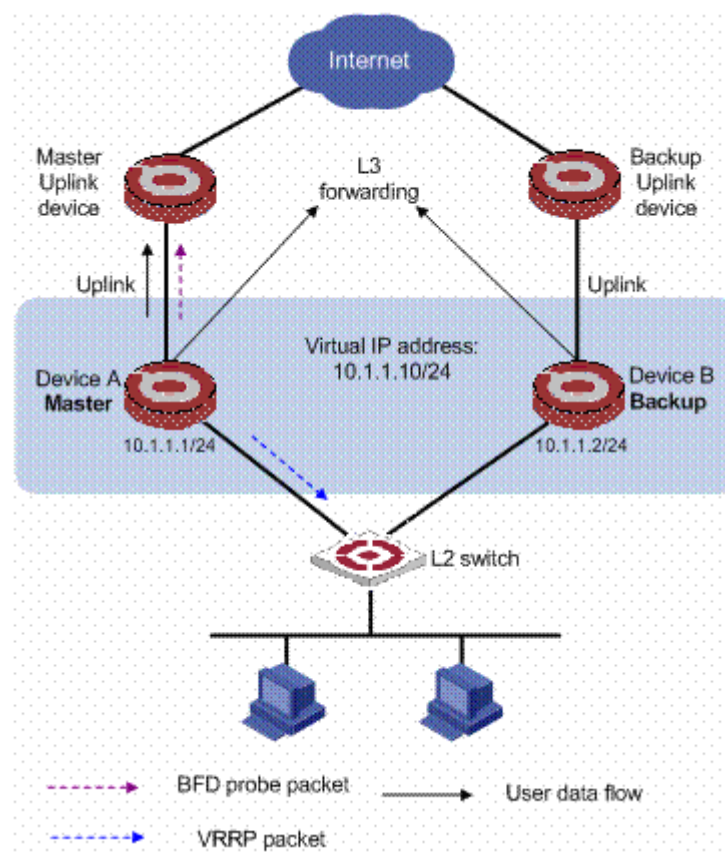


Figure 5 Master monitors the uplinks

As shown in Figure 5, Device A works as the master at first to forward packets; Device B works as the backup and is in the listening state. Device A uses BFD to monitor the state of the uplink to the Internet. If the uplink of Device A fails, Device A can be aware of the network change in milliseconds; then it reduces its priority by a specified value, and sends a VRRP advertisement to Device B; if the priority of Device B is higher than that contained in the VRRP advertisement, Device B will become the master in Skew Time, and then the new master will forward packets for the hosts in the network.

7.3.4 Backup's Monitoring of Master State Using BFD

To ensure transmission stability on a network, BFD can be used on a backup to monitor the master state, thus ensuring that the backup can become the master immediately when the master fails.

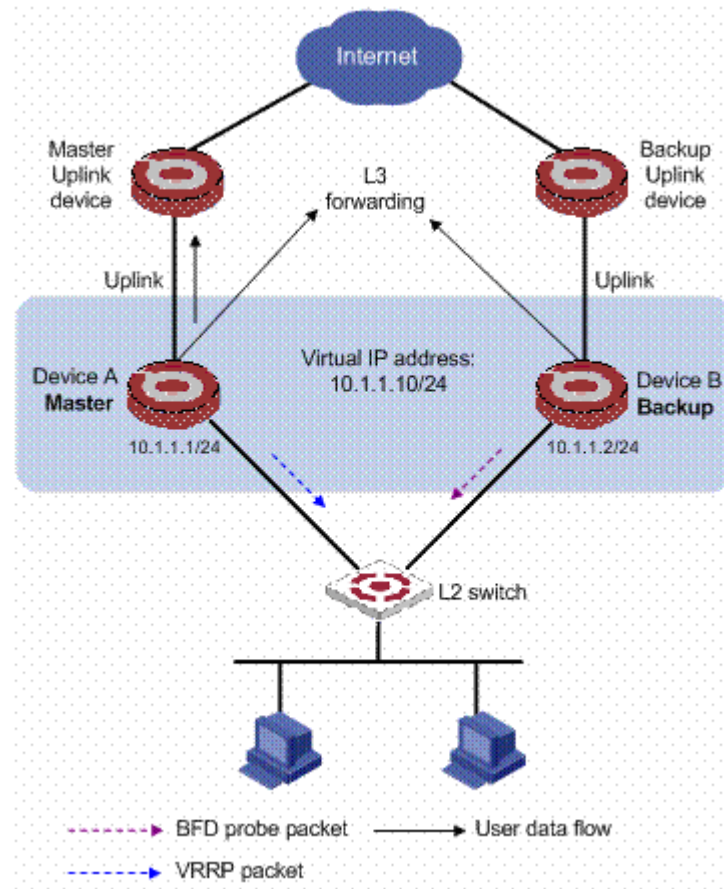


Figure 6 Backup monitors the master state

As shown in Figure 6, Device A works as the master at first to forward packets; Device B works as the backup and is in the listening state. Device B uses BFD to monitor the reachability of the IP address 10.1.1.1 on Device A; if Device A fails, Device B can be aware of the change of Device A through BFD and becomes the new master, and then forwards packets for the hosts in the network.

7.4 References

- RFC 3768: Virtual Router Redundancy Protocol (VRRP)

7.5 VRRP Configuration

VRRP configuration list is as following :

- ❖ Add or delete virtual IP address
- ❖ Configure priority of backup group
- ❖ Configure preemptible way and delay time of backup group
- ❖ Configure timer of backup group
- ❖ Enable ping of virtual IP

7.5.1 Add or delete virtual IP address

Specify IP address of this network interface to a virtual switch (also called a backup group), or remove a virtual IP address from a backup group.

Configure it in VLAN interface configuration mode.

```
ip vrrp vrid vip
no ip vrrp vrid [vip]
```

Backup id is in the range of 1 to 255. Virtual address can be undistributed IP address in the interface where the backup group is in, and also can be IP address of backup group interface. At most 8 backup groups can be configured. If this address is the one the switch has used, it also can be configured. Now, this switch is called an IP Address Owner. When specify the first IP address to a backup group, system will create this backup group, and add virtual IP address to this backup group from that on, system will only add the address to the backup group. At most 8 IP address can be configured to each backup group.

When deleting the last IP address, the backup group will be deleted at the same time, that is, there is no this backup group in this interface and all configurations are not valid.

7.5.2 Configure priority of backup group

In VRRP, determine the position of each switch in backup group according to priority. The one with the superior will be the Master.

The priority value is in the range of 0 to 255 (the larger the number is, the superior the priority level is) and the configured range is 1 to 254. Priority 0 is reserved for special uses and 255 is reserved to IP address owner.

Configure it in VLAN interface configuration mode.

```
vrrp priority vrid priority
no vrrp priority vrid
```

By default, the priority is in the range of 100.



Caution: For IP address owner, the priority cannot be configured. It is 255 all the time.

7.5.3 Configure preemptible way and delay time of backup group

Once there is a Master in the backup group, and there is no failure, and other switch though has configured to possess superior priority, it will not be Master unless the preemption is configured. If the switch is configured to be preempt, once it possesses its priority is superior than the Master, it will be the Master. Accordingly, the original Master will be the backup. The delay time can be configured at the same time as the preemption, which can delay backup being Master. The aim of delay time is: In unstable network, if Backup doesn't receive the packet from Master on time, it will become Master (the reason why Backup cannot receive the packet is because of the congestion of the network, not the abnormal working of Master). So waiting for a certain time, the packet will be received from Master, which avoids frequent changes.

The delay time is in the unit of second which is in the range of 0 to 255.

Configure it in VLAN interface configuration mode.

```
vrrp preempt vrid [ delay delay ]
no vrrp preempt vrid
```

It is defaulted to be preempt with the delay time being 0.



Caution: Cancelling preemption of backup group, the delay time will be 0.

7.5.4 Configure timer of backup group

Master switch in VRRP backup group can timely send VRRP packet (the time interval is `adver_interval`) to inform other switches it works normally. If backup hasn't received VRRP packet from master switch after a certain time (`master_down_interval`), it will think master is abnormal and turn itself to be master.

User can adjust VRRP packet sending time interval `adver_interval` by using configuration command. The time interval of `master_down_interval` is 3 times of `adver_interval`. The large traffic and different timer in switch will cause the abnormal of `master_down_interval` to shift status. For this, you can prolong `adver_interval` and configure delay time. The unit of `adver_interval` is second.

```
vrrp timer vrid adver-interval  
no vrrp timer vrid
```

Configure it in VLAN interface configuration mode.

```
vrrp timer vrid adver-interval  
no vrrp timer vrid
```

By default, `adver_interval` is 1 second.

7.5.5 Enable ping of virtual IP

To test the reachability of main switch, enable ping function of virtual IP, that is, received destination address to be virtual IP and switch is the main switch which can response ping packet, and host can ping virtual gateway.

Configure it in global configuration mode:

```
vrrp ping-enable  
no vrrp ping-enable
```

It is defaulted to disable ping of virtual IP

7.5.6 VRRP monitor and maintenance

User can display VRRP information by following command.

Use it in any configuration mode:

```
show vrrp [ vlan-interface vlan-id [ vrid ] ]
```

Chapter 8 RIP Configuration

8.1 Brief introduction of RIP

RIP is short for Routing Information Protocol. It is a protocol based on D-V (Distance-Vector) algorithm which is widely used in real application. It submits route information through UDP (User Datagram Protocol) and sends upgrade packet every 30 seconds. If local router hasn't received the upgrade packet from opposite end router after 180 seconds, local router will mark all routing information from the opposite end to be unreachable; if some route information hasn't received upgrade packet from the opposite end in 120 seconds after marking to be unreachable, local router will delete it from the route table.

The distance to the destination measured by Hop Count is Routing Metric. In RIP, the hop between router and the straightly connected network is 0, and the hop will be 1 if passing through a network which router can reach, and the rest may be deduced by analogy. To restrict convergence time, RIP prescribe Metric is the intergeral number between 0 to 15. The hop larger or equal to 16 is defined to be infinite, that is, the destination host or network is unreachable.

There are such 2 versions as RIP-1 and RIP-2 (RIP-2 supports plain text authentication) .

To improve capability and prevent routing ring, RIP supports Split Horizon and Poison Reverse.

Each router run RIP manages a routing database which contains all route item to all reachable destination.

These route information includes :

Destination address : IP address of host or network.

Next hop address : the next router address passed when going to the destination.

Output interface : the interface transferring packet.

Metric value : the cost to the destination which is an intergeral number from 0 to 16.

Timer : the time is from the last time the router is modified. Every time when the router is modified, the timer is configured to be 0.

The process of RIP enabling and running is as following :

(1) Enabling RIP, router will send requery packet in the form of broadcast to neighbor routers. After receiving it, neighbor routers (must enable RIP) will send response packet which contains local route table information back.

(2) The router who has sent requery packet modifies local route table after receiving response packet.

(3) At the same time, RIP broadcasts or multicasts local route table every 30 seconds to neighbor routers to maintain local route and choose a best route, and then, broadcast and multicast modify information to neighbor network to make global efficient of upgrading route. At the same time, RIP adopts overtime system to handle overtime route to guarantee real time of route, As internal route protocol, RIP makes router know the route information of the whole network through this system.

RIP has been one of the standard of delivering router and host route. The theory of switch with layer 3 switching IP packet is the same as that of router, so RIP is also adopted by layer 3 switch manufacturer. It can be used in simple structured, strong continuity district network, such as : residential community network. For complicated large network, it is suggested not using RIP.

8.2 RIP Overview

RIP is a simple Interior Gateway Protocol (IGP), mainly used in small-sized networks, such as academic networks and simple structured LANs. RIP is not applicable to complex networks.

RIP is still widely used in practical networking due to easier implementation, configuration and maintenance than OSPF and IS-IS.

8.2.1 RIP Working Mechanism

8.2.1.1 Basic concept of RIP

RIP is a Distance-Vector-based routing protocol, using UDP packets for exchanging information through port 520.

RIP uses a hop count to measure the distance to a destination. The hop count is known as metric. The hop count from a router to a directly connected network is 0. The hop count from one router to a directly connected router is 1. To limit convergence time, the range of RIP metric value is from 0 to 15. A metric value of 16 (or bigger) is considered infinite, which means the destination network is unreachable. That is why RIP is not suitable for large-scaled networks.

RIP prevents routing loops by implementing the split horizon and poison reverse functions.

8.2.1.2 RIP routing table

Each RIP router has a routing table containing routing entries of all reachable destinations, and each routing entry contains :

- Destination address : IP address of a host or a network.
- Next hop : IP address of the adjacent router's interface to reach the destination.
- Egress interface : Packet outgoing interface.
- Metric : Cost from the local router to the destination.
- Route time : Time elapsed since the routing entry was last updated. The time is reset to 0 every time the routing entry is updated.
- Route tag : Identifies a route, used in routing policy to flexibly control routes.

8.2.1.3 RIP initialization and running procedure

The following procedure describes how RIP works.

- After RIP is enabled, the router sends Request messages to neighboring routers. Neighboring routers return Response messages including all information about their routing tables.
- The router updates its local routing table, and broadcasts the triggered update messages to its neighbors. All routers on the network do the same to keep the latest routing information.
- RIP ages out timed out routes by adopting an aging mechanism to keep only valid routes.

8.2.1.4 RIP timers

RIP employs four timers, Update, Timeout, Suppress, and Garbage-Collect.

- The update timer defines the interval between routing updates.
- The timeout timer defines the route aging time. If no update for a route is received after the aging time elapses, the metric of the route is set to 16 in the routing table.
- The suppress timer defines how long a RIP route stays in the suppressed state. When the metric of a route is 16, the route enters the suppressed state. In the suppressed state, only routes which come from the same neighbor and whose metric is less than 16 will be received by the router to replace unreachable routes.
- The garbage-collect timer defines the interval from when the metric of a route becomes 16 to when it is deleted from the routing table. During the Garbage-Collect timer length, RIP advertises the route with the routing metric set to 16. If no update is announced for that route after the Garbage-Collect timer expires, the route will be deleted from the routing table.

8.2.1.5 Routing loops prevention

RIP is a distance-vector (D-V) based routing protocol. Since a RIP router advertises its own routing table to neighbors, routing loops may occur.

RIP uses the following mechanisms to prevent routing loops.

- Counting to infinity. The metric value of 16 is defined as unreachable. When a routing loop occurs, the metric value of the route will increment to 16.
- Split horizon. A router does not send the routing information learned from a neighbor to the neighbor to prevent routing loops and save the bandwidth.
- Poison reverse. A router sets the metric of routes received from a neighbor to 16 and sends back these routes to the neighbor to help delete useless information from the neighbor's routing table.
- Triggered updates. A router advertises updates once the metric of a route is changed rather than after the update period expires to speed up the network convergence.


8.2.2 RIP Version

RIP has two versions, RIPv1 and RIPv2.

RIPv1, a Classful Routing Protocol, supports message advertisement via broadcast only. RIPv1 protocol messages do not carry mask information, which means it can only recognize routing information of natural networks such as Class A, B, and C. That is why RIPv1 does not support discontinuous subnet.

RIPv2 is a Classless Routing Protocol. Compared with RIPv1, RIPv2 has the following advantages.

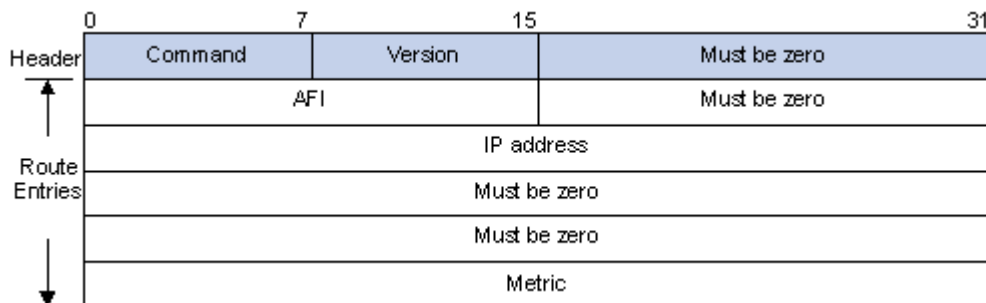
- Supporting route tags. The route tag is used in routing policies to flexibly control routes.
- Supporting masks, route summarization and classless inter-domain routing (CIDR).
- Supporting designated next hop to select the best next hop on broadcast networks.
- Supporting multicast routing update to reduce resource consumption.
- Supporting Plain text authentication and MD5 authentication to enhance security.

 Note : RIPv2 has two types of message transmission : broadcast and multicast. Multicast is the default type using 224.0.0.9 as the multicast address. The interface working in the RIPv2 broadcast mode can also receive RIPv1 messages.

8.2.3 RIP Message Format

8.2.3.1 RIPv1 message format

A RIP message consists of the Header and up to 25 route entries.



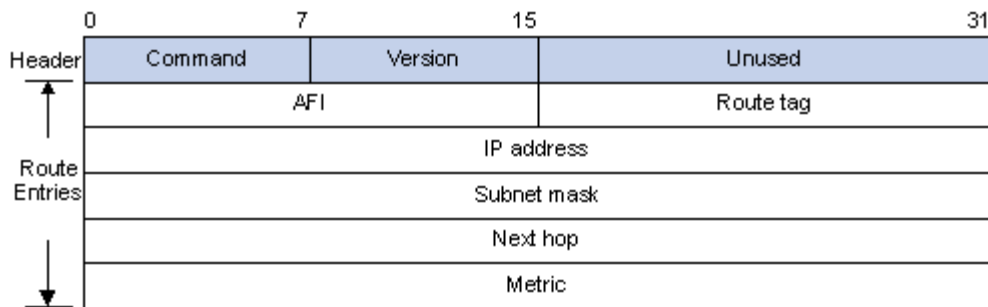
RIPv1 Message Format

- Command : The type of message. 1 indicates Request, 2 indicates Response.
- Version : The version of RIP, 0x01 for RIPv1.
- AFI : Address Family Identifier, 2 for IP.
- IP Address : Destination IP address of the route; can be a natural network, subnet or a host address.

- Metric : Cost of the route.

8.2.3.2 RIPv2 message format

The format of RIPv2 message is similar with RIPv1.



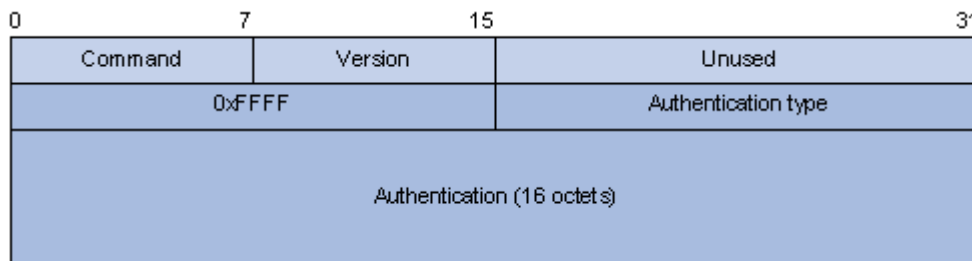
RIPv2 Message Format

The differences from RIPv1 are stated as following.

- Version : Version of RIP. For RIPv2 the value is 0x02.
- Route Tag : Route Tag.
- IP Address : Destination IP address. It could be a natural network address, subnet address or host address.
- Subnet Mask : Mask of the destination address.
- Next Hop : If set to 0.0.0.0, it indicates that the originator of the route is the best next hop; Otherwise it indicates a next hop better than the originator of the route.

8.2.3.3 RIPv2 authentication

RIPv2 sets the AFI field of the first route entry to 0xFFFF to identify authentication information.



RIPv2 Authentication Message

- Authentication Type : 2 represents plain text authentication, while 3 represents MD5.
- Authentication : Authentication data, including password information when plain text authentication is adopted or including key ID, MD5 authentication data length and sequence number when MD5 authentication is adopted.

 Note :

- RFC 1723 only defines plain text authentication. For information about MD5 authentication, refer to RFC2082 “RIPv2 MD5 Authentication”.
- With RIPv1, you can configure the authentication mode in interface view. However, the configuration will not take effect because RIPv1 does not support authentication.

8.2.4 TRIP

Triggered RIP (TRIP), a RIP extension on WAN, is mainly used in dial-up network.

8.2.4.1 Working mechanism

Routing information is sent in triggered updates rather than periodic broadcasts to reduce the routing management cost the WAN.

- Only when data in the routing table changes or the next hop is unreachable, a routing update message is sent.
- Since the periodic update delivery is canceled, an acknowledgement and retransmission mechanism is required to guarantee successful updates transmission on WAN.

8.2.4.2 Message types

RIP use three new types of message which are identified by the value of the Command filed.

- update request (type value 9) : Requests needed routes from the peer.
- update response (type value 10) : Contains the routes requested by the peer.
- Update Acknowledge (type value 11) : Acknowledges received update response messages.

8.2.4.3 TRIP retransmission mechanism

- If receiving no update responses after sending an update request, a router sends the request again after a specified interval. If still receiving no update response after the upper limit for sending requests is reached, the router considers the neighbor unreachable.
- If receiving no Update Acknowledge after sending an update response, a router sends the update response again after a specified interval. If still receiving no Update Acknowledge after the upper limit for sending update responses is reached, the router considers the neighbor unreachable.

8.2.5 Protocols and Standards

RFC 1058 : Routing Information Protocol
RFC 1723 : RIP Version 2 - Carrying Additional Information
RFC 1721 : RIP Version 2 Protocol Analysis
RFC 1722 : RIP Version 2 Protocol Applicability Statement
RFC 1724 : RIP Version 2 MIB Extension
RFC 2082 : IPv2 MD5 Authentication
RFC 2091 : Triggered Extensions to RIP to Support Demand Circuits

8.3 RIP configuration list

In every configuration, enable RIP and RIP network before configuring other functions. Configuring functions which relates to interface is not restricted by RIP enabling. Caution : after disabling RIP, original parameter still exists, and it will be effective when enable RIP next time.

Configuration list is as following :

- [Enable RIP](#)
- [Specify IP network to run RIP protocol](#)
- [RIP working status of specified interface](#)
- [RIP version of specified interface](#)
- [Enable host routing](#)
- [Enable route convergence](#)

- [Configure authentication to RIP packet](#)
- [Configure split](#)
- [Configure metricin](#)
- [Define prefix ACL](#)
- [Configure route redistribute](#)
- [Configure route filtration](#)
- [Display RIP configuration](#)

8.3.1 Enable RIP

By default, RIP is disabled. Enable RIP mode in global configuration mode :
Enable RIP and enter RIP configuration mode

route rip

Disable RIP

no route rip

8.3.2 Specify IP network to run RIP protocol

By default, after RIP enabling, no interface runs RIP protocol, only when administrator specifies some IP network to run RIP protocol, this interface will send and receive RIP packet. Configure it in RIP protocol configuration mode :

Specify to run RIP protocol in IP network

network *ip-address*

Cancel to run RIP protocol in IP network

no network *ip-address*

8.3.3 RIP working status of specified interface

Specify RIP working status in interface configuration mode, such as : run RIP or not in interface, receive and send RIP upgrade packet in interface or not; it can also specify sending (or receiving) RIP upgrade packet.

Configure it in interface configuration mode :

Enable interface to run RIP

ip rip work

Disable interface to run RIP

no ip rip work

After disabling interface running RIP, this interface will not send or receive RIP upgrade packet, but other interface still can send and receive route of this interface.

Permit interface to receive RIP packet

ip rip input

Forbid interface to receive RIP packet

no ip rip input

Permit interface to send RIP packet

ip rip output

Forbid interface to send RIP packet

no ip rip output

8.3.4 RIP version of specified interface

RIP has RIP-1 and RIP-2 two versions which can specify RIP packet version handled by interface.

RIP-1 uses broadcast and RIP-2 supports broadcast and multicast and it is defaulted to use multicast.

Multicast address in RIP-2 is 224.0.0.9.

The advantage of using multicast is that in the same network interface, the host which is not running RIP can avoid receiving RIP broadcast; using multicast can avoid host which runs RIP-1receiving and handling route with subnet mask in RIP-2. When interface running rip-2, it can also receive RIP-1 packet.

Configure it in interface configuration mode :

Specify RIP working version of interface to be RIPV1

ip rip version 1

Specify RIP working version of interface to be RIPV2 multicast

ip rip version 2 mcast

Specify RIP working version of interface to be RIPV2 broadcast

ip rip version 2 bcast

Delete rip version number and configure it to default rip1

no ip rip version

8.3.5 Enable host routing

In some cases, RIP packet received by router contains host route table item which has little to do with searching address but occupies a lot of resources. Configure it to be sure whether the switch receives it.

Configure it in RIP protocol configuration mode :

Permit host route

host-route

Forbid host route

no host-route

8.3.6 Enable route convergence

Route convergence means routes of different subnetwork in the same network convergent to be a route with natural netmask when sending to other networks. Route convergence reduces route information volume and switching information volume.

RIP-1 only sends route with natural netmask, that is, send route out by using route convergence. RIP-2 supports network mask. When sending all routes out in the form of broadcasting, disable route convergence of RIP-2.

Configure it in RIP protocol configuration mode :

Enable RIP-2 route convergence

auto-summary

Disable RIP-2 route convergence

no auto-summary

By default, RIP-2 uses route convergence.

8.3.7 Configure authentication to RIP packet

RIP-1 doesn't support packet authentication. When running RIP-2, it can configure to use packet authentication or not. Authentication is plain text or md5 key used.

Configure it in interface configuration mode :

Configure RIP-2 plain text authentication :

ip rip authentication {simple; md5} {password; key-id number key-string string}

Restore RIP packet authentication

no ip rip authentication

8.3.8 Configure split

Split means not sending route which is learnt by this interface. It can avoid route ring. But in some special cases, split is forbidden to guarantee correct transmission instead of efficiency. By default, interface permits split.

Configure it in interface configuration mode :

Enable split

ip rip split

Disable split

no ip rip split

8.3.9 Configure metricin

Routing Metric is input or output metric added by RIP route. Routing Metric cannot change route metric in route table, but add a specified metric when sending and receiving route.

Configure it in interface configuration mode :

Configure Routing Metric when receiving RIP packet

ip rip metricin value

Disable Routing Metric when receiving RIP packet

no ip rip metricin

Configure Routing Metric when sending RIP packet

ip rip metricout value

Disable Routing Metric when sending RIP packet

no ip rip metricout

By default, RIP Routing Metricis 0 when sending and receiving packet.

8.3.10 Define prefix list

A prefix-list is marked by prefix list name. Each prefix-list can contain many items and each item can specifies a matching range through sequence-number which shows the matching order in prefix-list.

When matching, switch will check each item according to ascending order. It will filtrate the prefix-list when there is one item matches.

Caution : By default, if at least one prefix list is defined, the matching mode of at least one item is permit. Deny mode item can fast filtrate the route information which is not matched. If all item is in deny mode, any route will not pass the filtration. It can define an item of permit 0.0.0.0/0 to permit all route information to pass after many deny mode items.

Above situation can be changed by ip prefix-list default command. Details refer to command line configuration manual.ssss

Configure it in global configuration mode :

Create prefix ACL or adding item

ip prefix-list

Delete prefix list or some item

no ip prefix-list

Configure matching mode when prefix does not exist or there is no matching item

ip prefix-list default

Restore to default matching mode when prefix does not exist or there is no matching item

no ip prefix-list default

8.3.11 Configure redistribution

RIP permits user to introduce other route protocol to RIP.

The route protocol that can be introduced are : connected, static and ospf.

Configure it in RIP protocol configuration :

Introduce other route protocol

redistribute

Cancel introduction of other route protocol

no redistribute

8.3.12 Configure distribute-list

Filtrate route through configuring strategy rules for receiving and sending route by specifying address prefix list. In addition, receive specified switch RIP packet by specifying neighbor switch.

Configure it in RIP protocol configuration :

Configure RIP to filtrate received route

distribute-list prefix-list in

Configure RIP to filtrate sent route

distribute-list prefix-list out

Configure RIP to receive specified route

distribute-list gate-way in

Cancel filtration

no distribute-list

8.3.13 Display RIP configuration

There are 3 commands to display RIP information.

Display RIP statistics information

show ip rip

Display RIP interface configuration, such as version, authentication

show ip rip interface

Display RIP route table

show ip route rip

Chapter 9 BFD

9.1 Overview

9.1.1 Background

To protect key applications, a network is usually designed with redundant backup links. Devices need to quickly detect communication failures and restore communication through backup links as soon as possible. On some links, such as POS links, devices detect link failures by sending hardware detection signals. However, some other links, such as Ethernet links, provide no hardware detection mechanism. In that case, devices can use the hello mechanism of a protocol for failure detection, which has a failure detection rate of more than one second. Such a rate is too slow for some applications. Some routing protocols, such as OSPF and IS-IS, provide a fast hello mechanism for failure detection, but this mechanism has a failure detection rate of at least one second and is protocol-dependent.

9.1.2 Benefits

BFD provides a general-purpose, standard, medium- and protocol-independent fast failure detection mechanism. It has the following benefits:

- Detecting failures on any bidirectional forwarding paths, such as direct physical link, virtual link, tunnel, MPLS LSP, multi-hop path, and unidirectional link, between network devices.
- Providing consistent fast fault detection time for upper-layer applications.
- Providing a failure detection time of less than one second for faster network convergence, short application interruptions, and enhanced network reliability.

9.2 BFD Implementation

9.2.1 Mechanism

BFD establishes a session between two network devices to detect failures on the bidirectional forwarding paths between the devices and provide services for upper-layer protocols. BFD provides no neighbor discovery mechanism. Protocols that BFD services notify BFD of devices to which it needs to establish sessions. After a session is established, if no BFD control packet is received from the peer within the negotiated BFD interval, BFD notifies a failure to the protocol, which then takes appropriate measures. The following section describes the operation of BFD for OSPF.

(1) 1. BFD session establishment

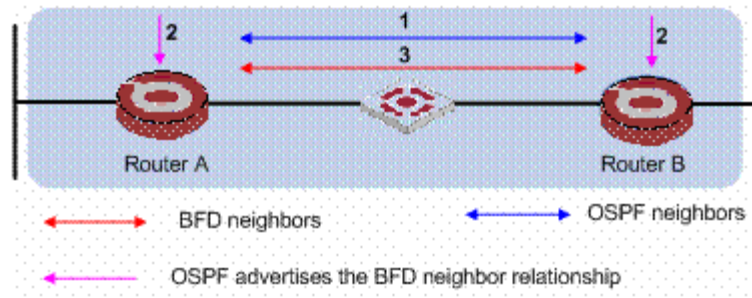


Figure 1 BFD session establishment

- 1) OSPF discovers neighbors by sending Hello packets and establish neighbor relationships.
- 2) After establishing neighbor relationships, OSPF notifies BFD of the neighbor information, including destination and source addresses.
- 3) BFD uses the information to establish BFD sessions.

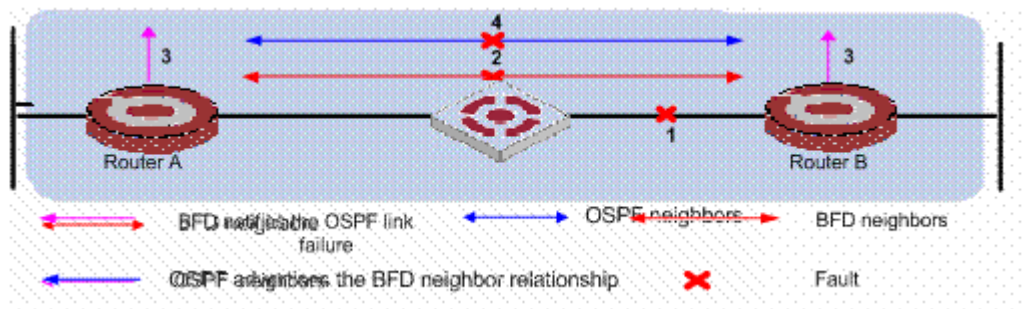


Figure 2 BFD fault detection

- 1) A link failure occurs.
- 2) Upon detection of the link failure, BFD clears the session.
- 3) BFD notifies the neighbor unreachability to OSPF.
- 4) OSPF terminates the neighbor relationship on the link.

There are two BFD operating modes: asynchronous and demand. A device operating in the asynchronous mode periodically sends BFD control packets. It tears down the BFD session if it receives no BFD control packet from the peer within the BFD interval.

A device operating in the echo mode periodically sends BFD echo packets. The peer device returns the received BFD echo packets back without processing them. If the sending device receives no BFD echo packet from the peer within the BFD interval, the session is considered down.

Both ends of a BFD session may be directly (one hop away from each other) or indirectly connected. BFD echo packets can only detect failures for directly connected neighbors. That is, BFD echo packets are sent over a single hop. BFD control packets, however, can detect failures for directly and indirectly connected neighbors. That is, BFD control packets can be sent over one or multiple hops.

9.2.2 BFD Packets

9.2.2.1 BFD Control Packets

A BFD control packet consists of the required fields and the optional authentication fields.

Figure 3 shows the format of the required fields:

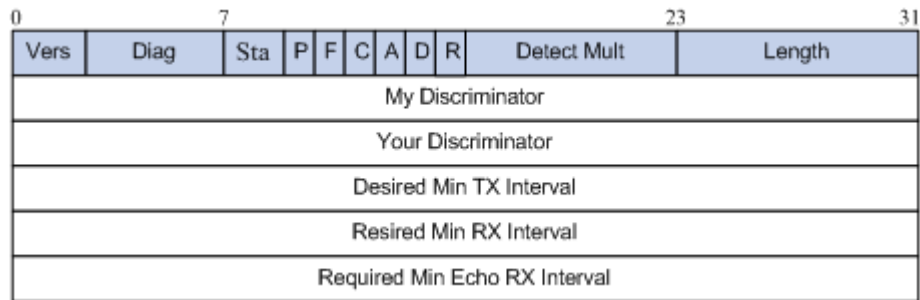


Figure 3 BFD control packet format

Figure 4 shows the format of the optional authentication fields.



Figure 4 BFD control packet (authentication fields)

Table 1 Description of the fields of a BFD control packet

Field	Description
Vers	BFD version. The current version is 1.
Diag	This bit indicates the reason for the last transition of the local protocol from up to some other state.
Sta	Current BFD session state. Its value can be 0 for AdminDown, 1 for Down, 2 for Init, and 3 for Up.
P	If it is set to 1, the transmitting system requests the connection acknowledgement or acknowledges a parameter change.
F	If it is set to 1, the transmitting system responds to a received BFD control packet that has the Poll (P) bit set.
C	If set to 1, it means the BFD implementation for the transmitting system is independent of its control plane.
A	If it is set to 1, the control packet contains the authentication field and the session is authenticated.
D	If set to 1, it means the transmitting system wishes to operate in the demand mode; if set to 0, it means the transmitting system ignores the demand mode or cannot operate in the demand mode.
R	Reserved. It is set to 0 during transmission and ignored during reception.
Detect Mult	Detect time multiplier.
Length	BFD control packet length, in bytes.
My Discriminator	A unique, nonzero discriminator value generated by the transmitting system, used to demultiplex multiple BFD sessions between the same pair of systems.
Your Discriminator	This field reflects back the received value of My Discriminator or is 0 if that value is unknown.

Field	Description
Desired Min TX Interval	Minimum interval at which the local protocol wishes to send BFD control packets, in milliseconds.
Required Min RX Interval	Minimum interval at which the local system can receive BFD control packets, in milliseconds.
Required Min Echo RX Interval	This is the minimum interval, in microseconds, between received BFD Echo packets that this system is capable of supporting. If this value is zero, the transmitting system does not support the receipt of BFD Echo packets.
Auth Type	Authentication type used by BFD control packets.
Auth Len	Authentication field length, including authentication type field and authentication length field, in bytes.

BFD control packets are encapsulated in UDP packets, using destination port 3784 and a source port from 49152 to 65535.

9.2.2.2 BFD Echo Packets

BFD echo packets provide a fault detection mechanism without the use of BFD control packets. One end sends BFD echo packets to the peer, which returns received BFD echo packets back without processing them. Therefore, no BFD echo packet format is defined, as long as the transmitting end can distinguish between sessions through packet contents.

BFD echo packets are encapsulated in UDP packets, using destination port 3785, with the IP address of the transmitting interface as the destination IP address and a configured source IP address, which must not cause ICMP redirection.

9.2.3 BFD Session Establishment

 Note:

The following section describes the process of session establishment and fault detection by sending BFD control packets.

Before a BFD session is established, there are two BFD operation modes: active or passive:

- Active mode: Before a session is established, BFD actively sends BFD control packets regardless of whether any BFD control packet is received from the peer.
- Passive mode: Before a session is established, no BFD control packet is sent until a BFD control packet is received from the peer.

During session initialization, at least one end of the two in communication must operate in the active mode for a session to be established. The following example shows the session establishment process with both ends working in the active mode. The session establishment process with one end in active mode and the other in passive mode is the same.

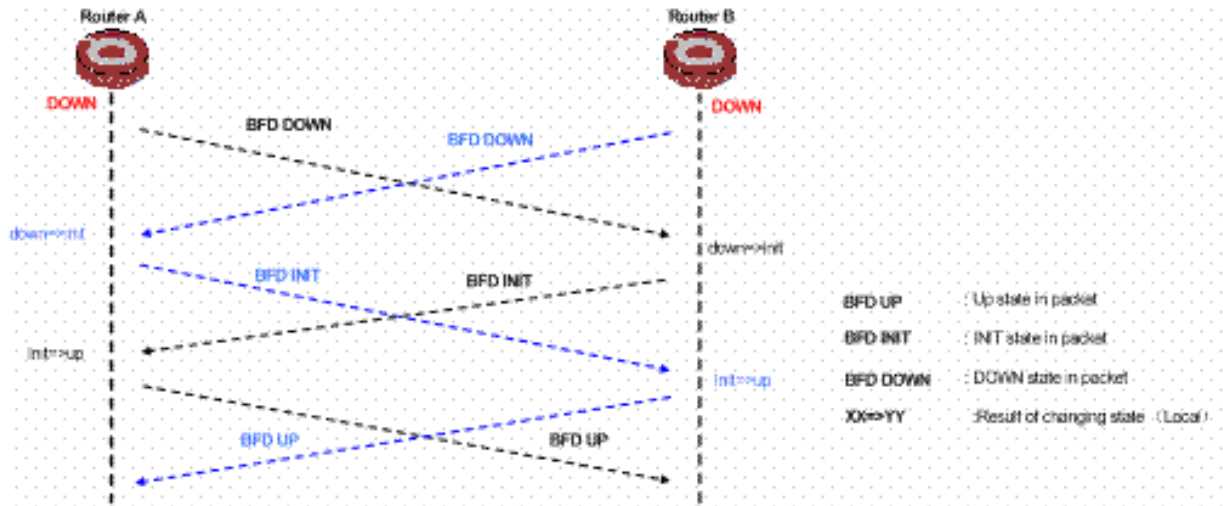


Figure 5 BFD session establishment

BFD uses a three-way handshake mechanism to establish sessions. The transmitting end fills the **Sta** field with its current session state in a transmitted BFD control packet. The receiving end changes its session state based on the **Sta** field value in the received BFD control packet and its own session status to establish a session.

- 1) As shown in the above figure, upon receipt of a notification from an upper-layer application, Routers A and B send a BFD control packet in DOWN state to the peer.
- 2) When Router B receives the BFD control packet in DOWN state, the local session state transits from DOWN to INIT. In the BFD control packet sent subsequently, the Sta field is filled with a value of 2, indicating the session state is INIT. Router A experiences the same BFD state transition as Router B.
- 3) When Router A receives the BFD control packet in INIT state from the peer, the local session state transits from INIT to UP. In the BFD control packet sent subsequently, the Sta field is filled with a value of 3, indicating the session state is UP. Router B experiences the same BFD state transition as Router A.
- 4) Both BFD peers are UP. A session is established successfully and BFD starts to detect link failures.

9.2.4 Timer Negotiation

Before a BFD session is established, BFD control packets are sent every one second to reduce traffic. After a session is established, BFD control packets are sent at the negotiated interval for fast detection, and the BFD control packet transmit interval and detection timer are negotiated. If a BFD session is valid, these timers can be negotiated and modified without affecting session state. The timers for different BFD session directions are negotiated independent of each other and therefore can be different.

The BFD control packet transmit interval is the greater of the Desired Min TX Interval of the local end and the Required Min RX Interval of the peer.

The detection timer is the Detect Mult of the BFD control packets transmitted by the peer times the negotiated BFD control packet transmit interval of the peer.

If the Desired Min TX Interval of the local end increases, the actual BFD control packet transmit interval of the local end cannot be changed until the local end receives a packet with the F bit set from the peer. This ensures that the peer has increased the detection timer before the BFD control packet transmit interval increases on the local end, thus avoiding detection timer timeout errors on the peer.

If the Required Min RX Interval of the local end decreases, the detection timer of the local end cannot be changed until the local end receives a packet with the **F** bit set from the peer. This ensures that the peer has decreased the BFD control packet transmit interval before the local end decreases the detection timer, thus avoiding detection timer timeout errors on the local end.

If the Desired Min TX Interval decreases, so does the BFD control packet transmit interval on the local end immediately. If the Required Min RX Interval increases, so does the detection timer on the local end immediately.

The following describes the timer negotiation process after a parameter change.

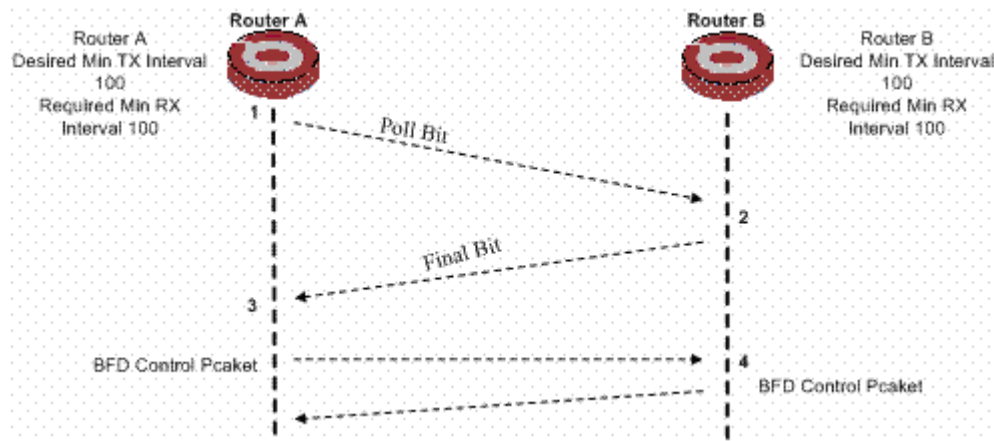


Figure 6 BFD detection timer negotiation

A BFD session is established between Router A and Router B. Both routers have the same Desired Min TX Interval (hereinafter referred to as TX) and Required Min RX Interval (hereinafter referred to as RX) of 100 milliseconds and the same Detect Mult of 3. According to timer negotiation rules, the BFD control packet transmit interval is Router A's TX or Router B's RX, whichever is greater, namely, 100 milliseconds, Router B's transmit interval is also 100 milliseconds, and both routers have a detection timer of 300 milliseconds.

If TX and RX of Router A increase to 150 milliseconds:

- 1) Router A compares its RX (150 milliseconds) with Router B's TX (100 milliseconds) and thus changes the detection timer of the local end to 450 milliseconds. Meanwhile, Router A sends a BFD control packet (with a TX and RX of 150 milliseconds) whose **P** bit is set to the peer.
- 2) Upon receipt of the packet, Router B replies to Router A with a BFD control packet whose **F** bit is set (with a TX and RX of 100 milliseconds). Meanwhile, Router B compares the RX in the received packet with the TX of the local end. As TX is greater, Router B's transmit interval is changed to 150 milliseconds. After comparing the RX of the local end with the TX of the peer, Router B also changes its detection timer to 450 milliseconds.
- 3) Router A receives a BFD control packet with the **F** bit set from the peer. After comparing the RX in the received packet and the TX of the local end, Router A calculates the new transmit interval as 150 milliseconds.
- 4) The timer negotiation is complete. The BFD control packet transmit interval and detection timer for the routers are 150 milliseconds and 450 milliseconds respectively.

9.2.5 Fault Detection

After BFD session establishment and timer negotiation, both ends start to transmit BFD control packets at the negotiated interval. Each time a BFD control packet is received, the detection timer is reset so that the session remains up. If no BFD control packet is received within the detection timer, the BFD session state transits to DOWN and BFD notifies the failure to the upper-layer application it services. The upper-layer application then takes proper measures. When the BFD session on the local end is DOWN, the **Sta** field of the BFD control packet transmitted to the peer is filled with a value of 1 to notify the peer that the session is DOWN. Then, the BFD session state of the peer also transits to DOWN.

9.3 Application Scenarios

9.3.1 Configuring BFD for Routing Protocols

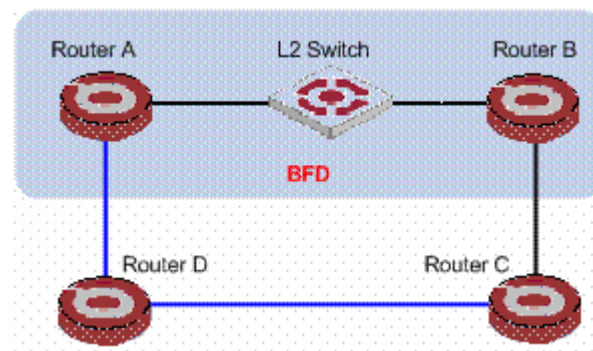


Figure 7 Application scenario where BFD works with routing protocols

Router A and Router B are interconnected through a Layer-2 switch. Both routers run a routing protocol.

As Router A and Router B are interconnected through a Layer-2 switch, a link failure between the routers may not cause an interface to be DOWN, and the link failure can be detected only through protocol handshake. After BFD is configured between Router A and Router B, a link failure between the routers can be detected quickly. Upon receipt of the link failure notification from BFD, the routing protocol recalculates routes for fast convergence.

9.3.2 Configuring BFD for Fast Reroute

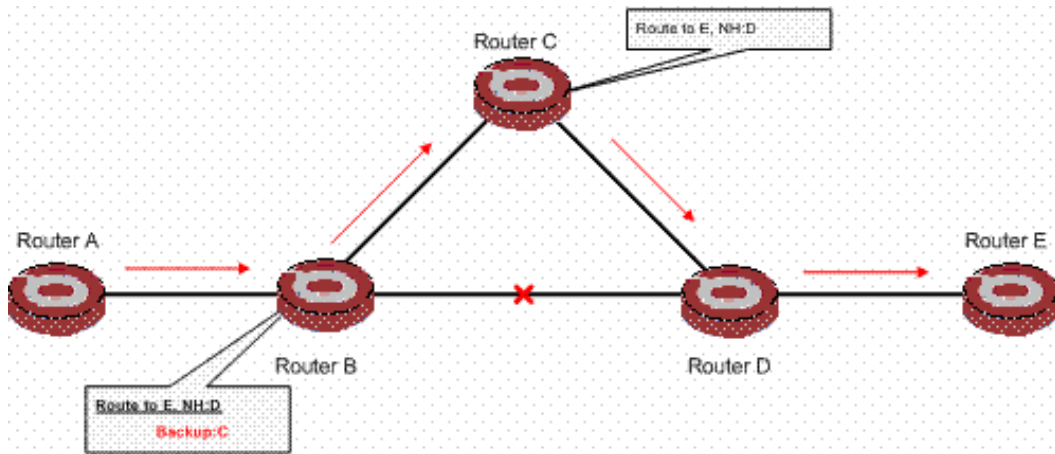


Figure 8 Application scenario where BFD is configured for fast reroute

Many delay-sensitive services on the Internet, such as audio and video services, require fast route convergence. Configuring BFD for routing protocols or using the fast route convergence technologies can greatly speed up convergence but cannot fully meet the failover requirements of audio and video services.

Configuring BFD for fast reroute can satisfy such requirements. Backup paths are calculated in advance and master path failures are detected quickly. When the master path fails, the traffic is directly switched to a backup path at the forwarding plane rather than the control plane, thus greatly shortening service interruptions.

9.3.3 Configuring BFD for VRRP

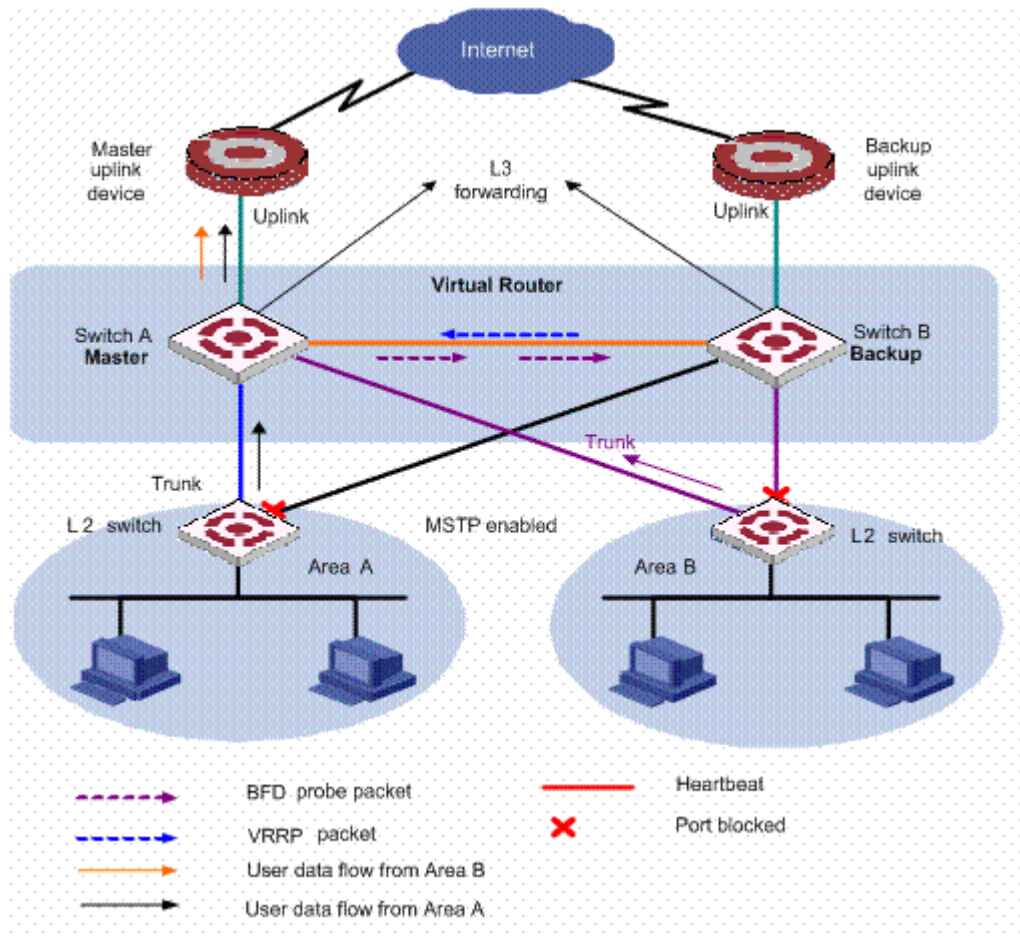


Figure 9 Application scenario where BFD is configured for VRRP

With the Virtual Router Redundancy Protocol (VRRP) enabled, when the master device fails, the backup device can quickly take over the forwarding task from the master, thus minimizing user data flow interruptions. When the master fails, if the backup receives no packet from the master within the preemption delay timer, the backup becomes the new master, with a switchover time of over one second. After BFD is configured for the backup to monitor the master, failures of the master can be detected more quickly to shorten user data flow interruptions.

VRRP also monitors the status of the master's uplink. When the master works normally but its uplink fails, user packets cannot be forwarded normally. VRRP determines whether an uplink is normal by monitoring the uplink interface status. When the monitored interface is down, the master lowers its priority to initiate a switchover. However, this mechanism depends on the protocol status; if the uplink fails but the protocol status of the interface remains up, the failure cannot be detected through VRRP. Configuring BFD for the VRRP master to monitor its uplink can solve this problem.

9.4 BFD Configuration

9.4.1 BFD Configuration list

BFD Configuration list is as following :

- Enable ospf bfd
- Configure BFD min-transmit-interval

- Configure BFD min-receive-interval
- Configure BFD detect multiplier
- Configure BFD session initial mode
- Configure BFD session demand
- Clear statistics of all session sending/receiving packet in current interface
- Show all BFD session
- Show BFD configuration in all interface

9.4.2 Enable ospf bfd

Configure it in vlan interface mode :

- Enable ospf bfd

ip ospf bfd

- Disable ospf bfd

no ip ospf bfd

By default, ospf bfd is disabled (only support detecting ospf)

For example :

!Enable ospf bfd in vlan interface 1

QTECH(config-if-vlanInterface-1)#ip ospf bfd

9.4.3 Configure BFD min-transmit-interval

Configure it in vlan interface mode :

- Configure BFD min-transmit-interval

bfd min-transmit-interval XX

- Restore to default BFD min-transmit-interval

no bfd min-transmit-interval

By default, BFD min-transmit-interval is 400ms

For example :

Configure BFD min-transmit-interval in vlan interface 1 to be 800ms

QTECH(config-if-vlanInterface-1)#bfd min-transmit-interval 800

9.4.4 Configure BFD min-receive-interval

Configure it in vlan interface mode :

- Configure BFD min-receive-interval

bfd min-receive-interval XX

- Restore to default BFD min-receive-interval

no bfd min-receive-interval

By default, BFD min-receive-interval is 400ms

For example :

Configure BFD min-receive-interval in vlan interface 1 to be 800ms

QTECH(config-if-vlanInterface-1)#bfd min-receive-interval 800

9.4.5 Configure BFD detect multiplier

Configure it in vlan interface mode :

- Configure BFD detect multiplier

bfd detect-multiplier XX

- Restore to default BFD detect multiplier
no bfd detect-multiplier

By default, BFD detect multiplier is 5

For example :

Configure BFD detect multiplier in vlan interface 1 to be 10
QTECH(config-if-vlanInterface-1)#bfd detect-multiplier 10

9.4.6 Configure BFD session demand

Configure it in vlan interface mode :

- Configure BFD session demand
bfd demand { on | off }
- Restore to default BFD session demand
no bfd demand

By default, BFD session demand is off.

For example :

Configure BFD session demand in vlan interface 1
QTECH(config-if-vlanInterface-1)#bfd demand on

9.4.7 Configure BFD session initial mode

Configure it in vlan interface mode :

- Configure BFD session initial mode
bfd session init-mode { passive | active }
- Restore to default BFD session initial mode
no bfd session init-mode

By default, BFD session initial mode is active.

For example :

Configure BFD session initial mode in vlan interface 1 to be passive
QTECH(config-if-vlanInterface-1)#bfd session init-mode passive

9.4.8 Clear statistics of all session sending/receiving packet in current interface

Configure it in vlan interface mode :

- Clear statistics of all session sending/receiving packet in current interface
clear bfd statistics

For example :

Clear statistics of all session sending/receiving packet in vlan interface 1
QTECH(config-if-vlanInterface-1)#clear bfd statistics

9.4.9 Show all BFD session

- Show all BFD session
show bfd session [verbose]

Show all BFD session in any configuration mode.

For example :

Show all BFD session

QTECH(config)# show bfd session verbose

9.4.10 Show BFD configuration in all interface

- Show BFD configuration in all interface
show bfd interface [verbose]

Show BFD configuration in all interface in any configuration mode.

For example :

Show BFD configuration in all interface.

QTECH(config)# show bfd interface verbose

Chapter 10 OSPF Configuration

10.1 Brief introduction of OSPF

OSPF is short for Open Shortest Path First which is an internal route protocol based on link status and the shortest path precedence. In IP network, it searches and transmits route dynamically through collecting and delivering link status of autonomy system; OSPF protocol supports packet authentication based on interface to guarantee the safety of route calculating; OSPF protocol sends and receives packets in the form of IP multicast.

Each router supported OSPF protocol maintains a database which describes the topology of the whole autonomy. This database collects the link states advertise (LSA). Each router broadcasts information describing local states to the whole autonomy. In each multiple accessing network, if there are two or more routers, designated router (DR) and backup designated router (BDR) are selected. Designated router broadcasts network link states advertise out. Introducing this concept can reduce the number of neighborhood between each router in multiple accessing network. OSPF protocol permits autonomy system dividing into areas to be managed. Routing information transmitted between areas will be further abstracted to reduce bandwidth occupation.

OSPF uses 4 types of different routing, according to the precedence are :

- Inter Area Routing
- Area Border Routing
- The first type external routing
- The second type external routing

Inter Area Routing and Area Border Routing describe internal network structure of autonomy system; external routing describes how to select route to the destination out of autonomy system. Generally, the first type routing corresponds to information introduced by other internal routing protocol, the cost of which can be comparable with that of the OSPF itself; the second type of routing corresponds to the information introduced by external routing protocol, the cost of which is far beyond that of OSPF itself. So when calculating, only external cost is considered.

According to link state database, each router establishes a shortest path tree with the root of itself which can give out the routing to each node in autonomy system. External routing information appears in leaf node and it can broadcast its router to mark to keep record the extra information about autonomy system.

Areas of OSPF are connected by Backbone which with the mark of 0.0.0.0. All areas must be continuous logically. Backbone specially introduces virtual connection to guarantee the logical connection when the area is physically divided.

All the routers in the same area must be consensus the parameter configuration of this area. Therefore, when configuring routers in the same area, most configuration data must be considered based on area and error configuration may cause the non-communication of neighbour routers or routing information congestion and self-ring.

OSPF has the following features :

- Wide scope : Supports networks of various sizes and up to several hundred routers in an OSPF routing domain.
- Fast convergence : Transmits updates instantly after network topology changes for routing information synchronization in the AS.
- Loop-free : Computes routes with the shortest path first (SPF) algorithm according to the collected link states, so no route loops are generated.
- Area partition : Allows an AS to be split into different areas for ease of management and the routing information transmitted between areas is summarized to reduce network bandwidth consumption.
- Equal-cost multi-route : Supports multiple equal-cost routes to a destination.
- Routing hierarchy : Supports a four-level routing hierarchy that prioritizes the routes into intra-area, inter-area, external Type-1, and external Type-2 routes.
- Authentication : Supports interface-based packet authentication to guarantee the security of packet exchange.

- Multicast : Supports packet multicasting on some types of links.

10.1.1 Basic Concepts

10.1.1.1 Autonomous System

A set of routers using the same routing protocol to exchange routing information constitute an Autonomous System (AS).

10.1.1.2 OSPF route computation

OSPF route computation is described as follows :

- Based on the network topology around itself, each router generates Link State Advertisements (LSA) and sends them to other routers in update packets.
- Each OSPF router collects LSAs from other routers to compose a LSDB (Link State Database). An LSA describes the network topology around a router, so the LSDB describes the entire network topology of the AS.
- Each router transforms the LSDB to a weighted directed graph, which actually reflects the topology architecture of the entire network. All the routers have the same graph.
- Each router uses the SPF algorithm to compute a Shortest Path Tree that shows the routes to the nodes in the autonomous system. The router itself is the root of the tree.

10.1.1.3 Router ID

To run OSPF, a router must have a Router ID, which is a 32-bit unsigned integer, the unique identifier of the router in the AS.

You may assign a Router ID to an OSPF router manually. If no Router ID is specified, the system automatically selects one for the router as follows :

- If the loopback interfaces are configured, select the highest IP address among them.
- If no loopback interface is configured, select the highest IP address among addresses of active interfaces on the router.

10.1.1.4 OSPF packets

OSPF uses five types of packets :

- Hello packet : Periodically sent to find and maintain neighbors, containing the values of some timers, information about the DR, BDR and known neighbors.
- DD packet (database description packet) : Describes the digest of each LSA in the LSDB, exchanged between two routers for data synchronization.
- LSR (link state request) packet : Requests needed LSAs from the neighbor. After exchanging the DD packets, the two routers know which LSAs of the neighbor are missing from the local LSDBs. In this case, they send an LSR packet to each other, requesting the missing LSAs. The LSA packet contains the digest of the missing LSAs.
- LSU (link state update) packet : Transmits the needed LSAs to the neighbor.
- LSack (link state acknowledgment) packet : Acknowledges received LSU packets. It contains the headers of received LSAs (a packet can acknowledge multiple LSAs).

10.1.1.5 LSA types

OSPF sends routing information in LSAs, which, as defined in RFC 2328, have the following types :

- Router LSA : Type-1 LSA, originated by all routers, flooded throughout a single area only. This LSA

describes the collected states of the router's interfaces to an area.

- Network LSA : Type-2 LSA, originated for broadcast and NBMA networks by the designated router, flooded throughout a single area only. This LSA contains the list of routers connected to the network.
- Network Summary LSA : Type-3 LSA, originated by ABRs (Area Border Routers), and flooded throughout the LSA's associated area. Each summary-LSA describes a route to a destination outside the area, yet still inside the AS (an inter-area route).
- ASBR Summary LSA : Type-4 LSA, originated by ABRs and flooded throughout the LSA's associated area. Type 4 summary-LSAs describe routes to ASBR (Autonomous System Boundary Router).
- AS External LSA : Type-5 LSA, originated by ASBRs, and flooded throughout the AS (except stub and NSSA areas). Each AS-external-LSA describes a route to another AS.
- NSSA LSA : Type-7 LSA, as defined in RFC 1587, originated by ASBRs in NSSAs (Not-So-Stubby Areas) and flooded throughout a single NSSA. NSSA LSAs describe routes to other ASs.
- Opaque LSA : A proposed type of LSA, the format of which consists of a standard LSA header and application specific information. Opaque LSAs are used by the OSPF protocol or by some application to distribute information into the OSPF routing domain. The opaque LSA includes three types, Type 9, Type 10 and Type 11, which are used to flood into different areas. The Type 9 opaque LSA is flooded into the local subnet, the Type 10 is flooded into the local area, and the Type 11 is flooded throughout the whole AS.

10.1.1.6 Neighbor and Adjacency

In OSPF, the “Neighbor” and ”Adjacency” are two different concepts.

Neighbor : Two routers that have interfaces to a common network. Neighbor relationships are maintained by, and usually dynamically discovered by, OSPF's hello packets. When a router starts, it sends a hello packet via the OSPF interface, and the router that receives the hello packet checks parameters carried in the packet. If parameters of the two routers match, they become neighbors.

Adjacency : A relationship formed between selected neighboring routers for the purpose of exchanging routing information. Not every pair of neighboring routers become adjacent, which depends on network types. Only by synchronizing the LSDB via exchanging DD packets and LSAs can two routers become adjacent.

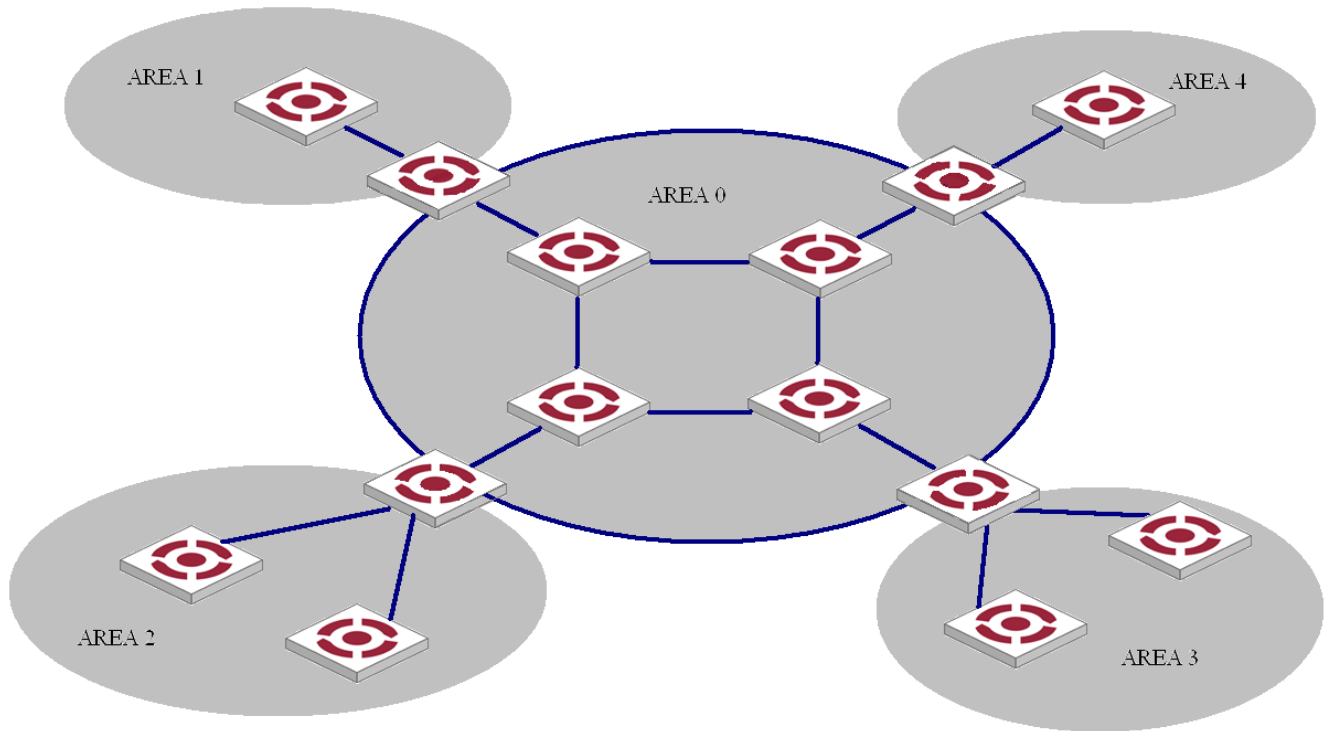
10.1.2 OSPF Area Partition and Route Summarization

10.1.2.1 Area partition

When a large number of OSPF routers are present on a network, LSDBs may become so large that a great amount of storage space is occupied and CPU resources are exhausted by performing SPF computation.

In addition, as the topology of a large network is prone to changes, enormous OSPF packets may be created, reducing bandwidth utilization. Each topology change makes all routers perform route calculation.

To solve this problem, OSPF splits an AS into multiple areas, which are identified by area ID. The boundaries between areas are routers rather than links. A network segment (or a link) can only reside in one area, in other words, an OSPF interface must be specified to belong to its attached area, as shown in the figure below.



OSPF area partition

After area partition, area border routers perform route summarization to reduce the number of LSAs advertised to other areas and minimize the effect of topology changes.

10.1.2.2 Classification of Routers

The OSPF routers fall into four types according to the position in the AS :

1) Internal Router

All interfaces on an internal router belong to one OSPF area.

2) Area Border Router (ABR)

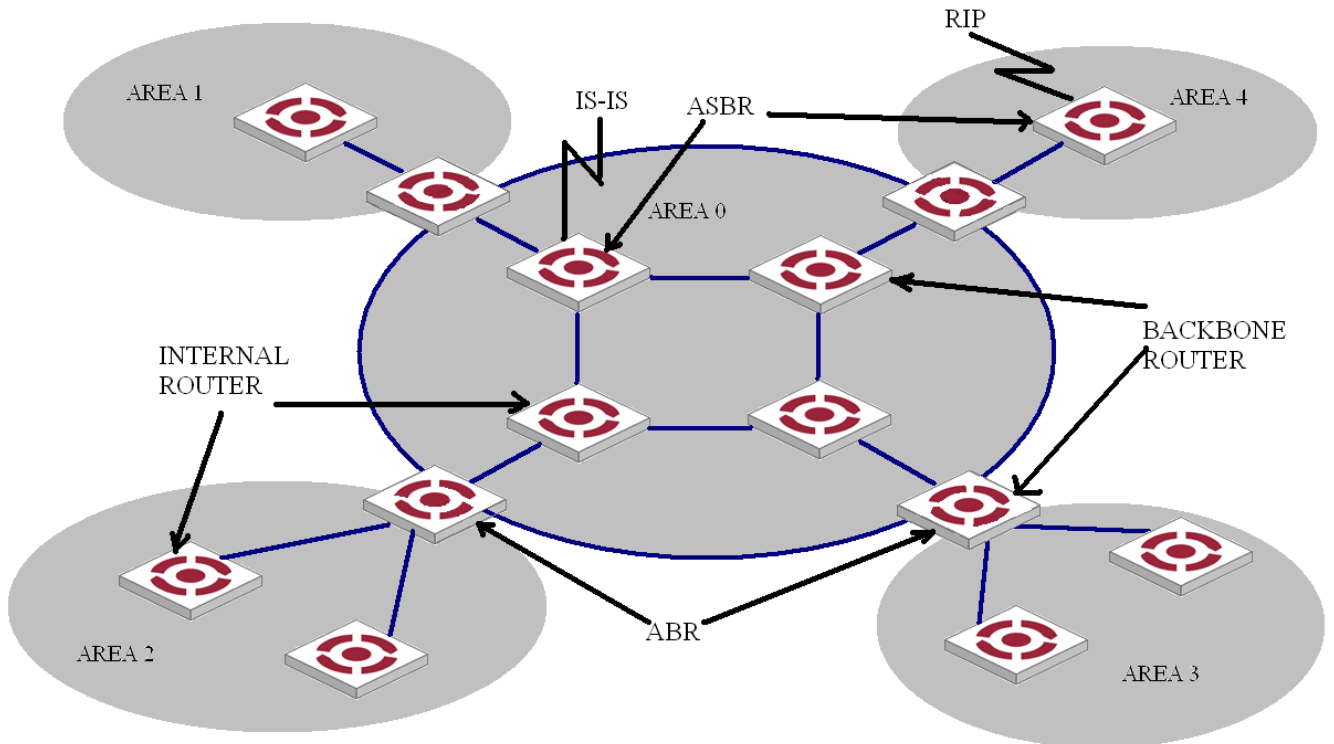
An area border router belongs to more than two areas, one of which must be the backbone area. It connects the backbone area to a non-backbone area. The connection between an area border router and the backbone area can be physical or logical.

3) Backbone Router

At least one interface of a backbone router must be attached to the backbone area. Therefore, all ABRs and internal routers in area 0 are backbone routers.

4) Autonomous System Border Router (ASBR)

The router exchanging routing information with another AS is an ASBR, which may not reside on the boundary of the AS. It can be an internal router or area border router.



OSPF router types

10.1.2.3 Backbone area and virtual links

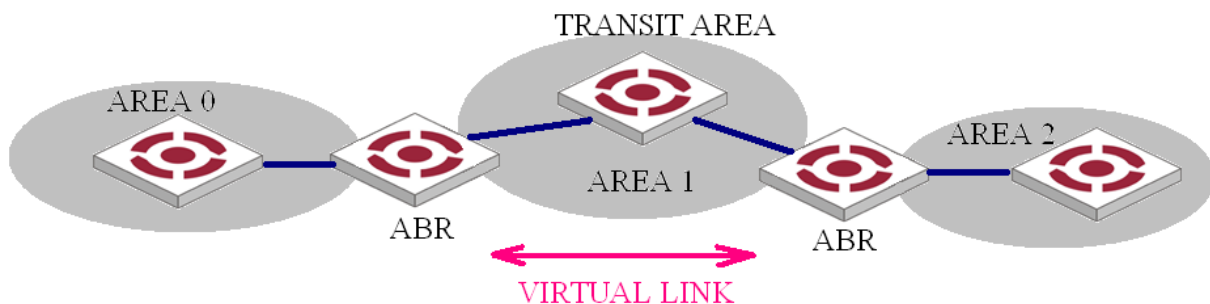
Each AS has a backbone area, which is responsible for distributing routing information between non-backbone areas. Routing information between non-backbone areas must be forwarded by the backbone area. Therefore, OSPF requires that :

- All non-backbone areas must maintain connectivity to the backbone area.
- The backbone area itself must maintain connectivity.

In practice, due to physical limitations, the requirements may not be satisfied. In this case, configuring OSPF virtual links is a solution.

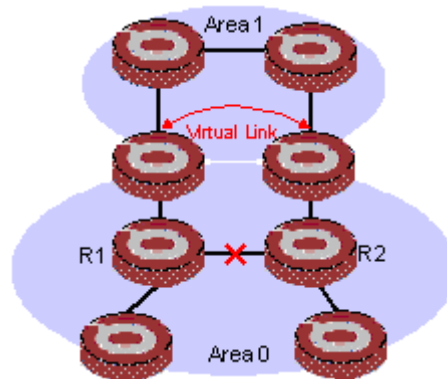
A virtual link is established between two area border routers via a non-backbone area and is configured on both ABRs to take effect. The area that provides the non-backbone area internal route for the virtual link is a “transit area”.

In the following figure, Area 2 has no direct physical link to the backbone area 0. Configuring a virtual link between ABRs can connect Area 2 to the backbone area.



Virtual link application 1

Another application of virtual links is to provide redundant links. If the backbone area cannot maintain internal connectivity due to a physical link failure, configuring a virtual link can guarantee logical connectivity in the backbone area, as shown below.



Virtual link application 2

The virtual link between the two ABRs acts as a point-to-point connection. Therefore, you can configure interface parameters such as hello packet interval on the virtual link as they are configured on physical interfaces.

The two ABRs on the virtual link exchange OSPF packets with each other directly, and the OSPF routers in between simply convey these OSPF packets as normal IP packets.

10.1.2.4 (Totally) Stub area

The ABR in a stub area does not distribute Type-5 LSAs into the area, so the routing table size and amount of routing information in this area are reduced significantly.

You can configure the stub area as a totally stub area, where the ABR advertises neither the destinations in other areas nor the external routes.

Stub area configuration is optional, and not every area is eligible to be a stub area. In general, a stub area resides on the border of the AS.

The ABR in a stub area generates a default route into the area.

Note the following when configuring a (totally) stub area :

- The backbone area cannot be a (totally) stub area.
- The stub command must be configured on routers in a (totally) stub area.
- A (totally) stub area cannot have an ASBR because AS external routes cannot be distributed into the stub area.
- Virtual links cannot transit (totally) stub areas.

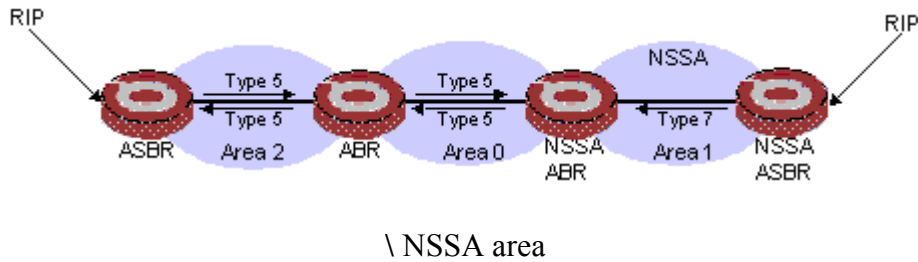
10.1.2.5 NSSA area

Similar to a stub area, an NSSA area imports no AS external LSA (Type-5 LSA) but can import Type-7 LSAs that are generated by the ASBR and distributed throughout the NSSA area. When traveling to the NSSA ABR, Type-7 LSAs are translated into Type-5 LSAs by the ABR for advertisement to other areas.

In the following figure, the OSPF AS contains three areas : Area 1, Area 2 and Area 0. The other two ASs employ the RIP protocol. Area 1 is an NSSA area, and the ASBR in it translates RIP routes into Type-7 LSAs and advertises them throughout Area 1. When these LSAs travel to the NSSA ABR, the ABR translates Type-7 LSAs to Type-5 LSAs for advertisement to Area 0 and Area 2.

On the left of the figure, RIP routes are translated into Type-5 LSAs by the ASBR of Area 2 and distributed into the OSPF AS. However, Area 1 is an NSSA area, so these Type-5 LSAs cannot travel to Area 1.

Like stub areas, virtual links cannot transit NSSA areas.

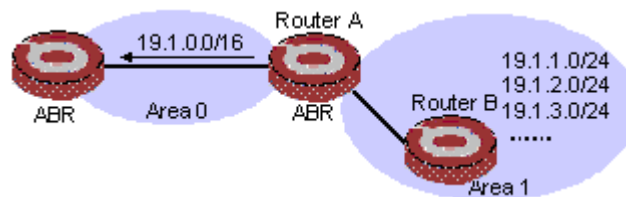


10.1.2.6 Route summarization

Route summarization : An ABR or ASBR summarizes routes with the same prefix with a single route and distribute it to other areas.

Via route summarization, routing information across areas and the size of routing tables on routers will be reduced, improving calculation speed of routers.

For example, as shown in the following figure, in Area 1 are three internal routes 19.1.1.0/24, 19.1.2.0/24, and 19.1.3.0/24. By configuring route summarization on Router A, the three routes are summarized with the route 19.1.0.0/16 that is advertised into Area 0.



Route summarization

OSPF has two types of route summarization :

1) ABR route summarization

To distribute routing information to other areas, an ABR generates Type-3 LSAs on a per network segment basis for an attached non-backbone area. If contiguous network segments are available in the area, you can summarize them with a single network segment. The ABR in the area distributes only the summary LSA to reduce the scale of LSDBs on routers in other areas.

2) ASBR route summarization

If summarization for redistributed routes is configured on an ASBR, it will summarize redistributed Type-5 LSAs that fall into the specified address range. If in an NSSA area, it also summarizes Type-7 LSAs that fall into the specified address range.

If this feature is configured on an ABR, the ABR will summarize Type-5 LSAs translated from Type-7 LSAs.

10.1.2.7 Route types

OSPF prioritize routes into four levels :

- Intra-area route
- Inter-area route
- Type-1 external route
- Type-2 external route

The intra-area and inter-area routes describe the network topology of the AS, while external routes describe routes to destinations outside the AS.

OSPF classifies external routes into two types : Type-1 and Type-2. A Type-1 external route is an IGP route, such as a RIP or static route, which has high credibility and whose cost is comparable with the cost of an OSPF internal route. The cost from a router to the destination of the Type-1 external route= the cost from the router to the corresponding ASBR+ the cost from the ASBR to the destination of the external route.

A Type-2 external route is an EGP route, which has low credibility, so OSPF considers the cost from the ASBR to the destination of the Type-2 external route is much bigger than the cost from the ASBR to an OSPF internal router. Therefore, the cost from the internal router to the destination of the Type-2 external route= the cost from the ASBR to the destination of the Type-2 external route. If two routes to the same destination have the same cost, then take the cost from the router to the ASBR into consideration.

10.1.3 Classification of OSPF Networks

10.1.3.1 OSPF network types

OSPF classifies networks into four types upon the link layer protocol :

- Broadcast : When the link layer protocol is Ethernet or FDDI, OSPF considers the network type broadcast by default. On Broadcast networks, packets are sent to multicast addresses (such as 224.0.0.5 and 224.0.0.6).
- NBMA (Non-Broadcast Multi-Access) : When the link layer protocol is Frame Relay, ATM or X.25, OSPF considers the network type as NBMA by default. Packets on these networks are sent to unicast addresses.
- P2MP (point-to-multipoint) : By default, OSPF considers no link layer protocol as P2MP, which is a conversion from other network types such as NBMA in general. On P2MP networks, packets are sent to multicast addresses (224.0.0.5).
- P2P (point-to-point) : When the link layer protocol is PPP or HDLC, OSPF considers the network type as P2P. On P2P networks, packets are sent to multicast addresses (224.0.0.5).

10.1.3.2 II. NBMA network configuration principle

Typical NBMA networks are ATM and Frame Relay networks.

You need to perform some special configuration on NBMA interfaces. Since these interfaces cannot broadcast hello packets for neighbor location, you need to specify neighbors manually and configure whether the neighbors have the DR election right.

An NBMA network is fully meshed, which means any two routers in the NBMA network have a direct virtual link for communication. If direct connections are not available between some routers, the type of interfaces associated should be configured as P2MP, or as P2P for interfaces with only one neighbor.

Differences between NBMA and P2MP networks :

- NBMA networks are fully meshed, non-broadcast and multi access. P2MP networks are not required to be fully meshed.
- It is required to elect the DR and BDR on NBMA networks, while DR and BDR are not available on P2MP networks.
- NBMA is the default network type, while P2MP is a conversion from other network types, such as NBMA in general.
- On NBMA networks, packets are unicast, and neighbors are configured manually on routers. On P2MP

networks, packets are multicast.

10.1.4 DR and BDR

10.1.4.1 DR/BDR introduction

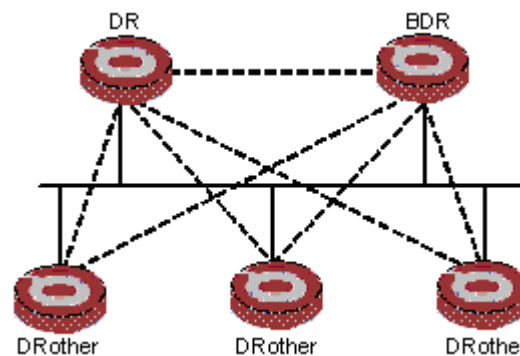
On broadcast or NBMA networks, any two routers exchange routing information with each other. If n routers are present on a network, $n(n-1)/2$ adjacencies are required. Any change on a router in the network generates traffic for routing information synchronization, consuming network resources. The Designated Router is defined to solve the problem. All other routers on the network send routing information to the DR, which is responsible for advertising link state information.

If the DR fails to work, routers on the network have to elect another DR and synchronize information with the new DR. It is time-consuming and prone to routing calculation errors. The Backup Designated Router (BDR) is introduced to reduce the synchronization period.

The BDR is elected along with the DR and establishes adjacencies for routing information exchange with all other routers. When the DR fails, the BDR will become the new DR in a very short period by avoiding adjacency establishment and DR reelection. Meanwhile, other routers elect another BDR, which requires a relatively long period but has no influence on routing calculation.

Other routers, also known as DRothers, establish no adjacency and exchange no routing information with each other, thus reducing the number of adjacencies on broadcast and NBMA networks.

In the following figure, real lines are Ethernet physical links, and dashed lines represent adjacencies. With the DR and BDR in the network, only seven adjacencies are enough.



DR and BDR in a network

10.1.4.2 DR/BDR election

The DR and BDR in a network are elected by all routers rather than configured manually. The DR priority of an interface determines its qualification for DR/BDR election. Interfaces attached to the network and having priorities higher than '0' are election candidates.

The election votes are hello packets. Each router sends the DR elected by itself in a hello packet to all the other routers. If two routers on the network declare themselves as the DR, the router with the higher DR priority wins. If DR priorities are the same, the router with the higher router ID wins. In addition, a router with the priority 0 cannot become the DR/BDR.



Note that :

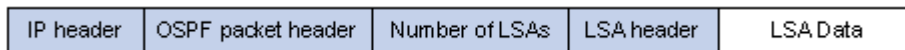
- The DR election is available on broadcast, NBMA interfaces rather than P2P, or P2MP

interfaces.

- A DR is an interface of a router and belongs to a single network segment. The router's other interfaces may be a BDR or DRother.
- After DR/BDR election and then a new router joins, it cannot become the DR immediately even if it has the highest priority on the network.
- The DR may not be the router with the highest priority in a network, and the BDR may not be the router with the second highest priority.

10.1.5 OSPF Packet Formats

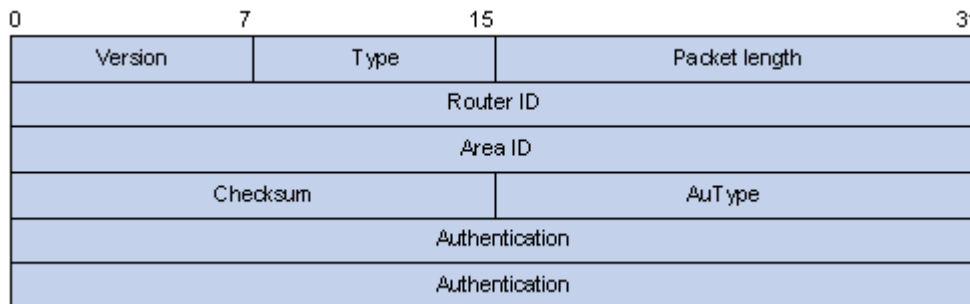
OSPF packets are directly encapsulated into IP packets. OSPF has the IP protocol number 89. The OSPF packet format is shown below (taking a LSU packet as an example).



OSPF packet format


10.1.5.1 OSPF packet header

OSPF packets are classified into five types that have the same packet header, as shown below.



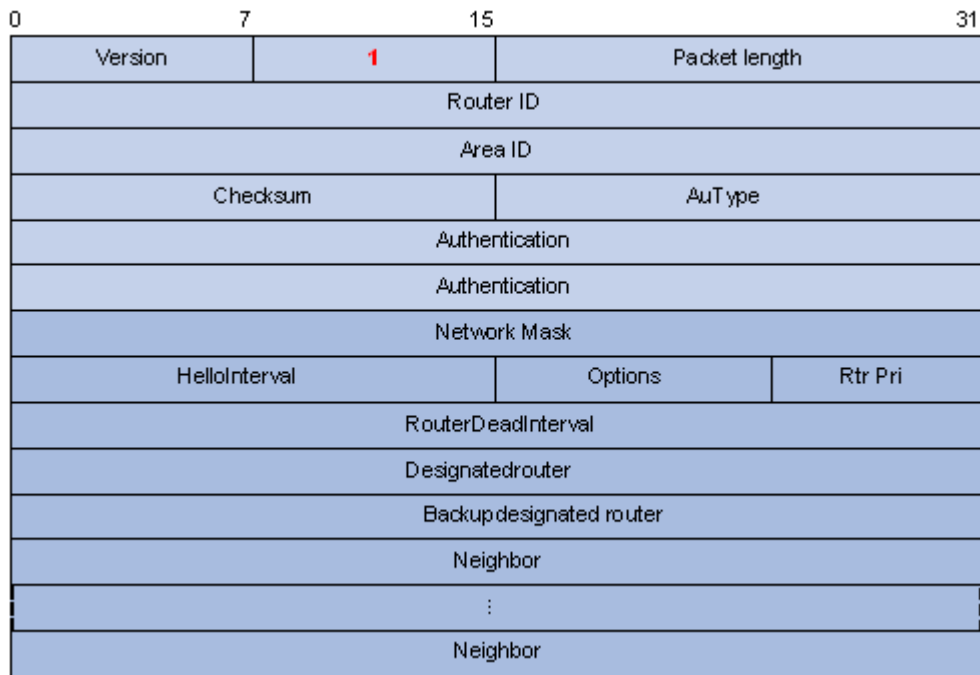
OSPF packet header

- Version : OSPF version number, which is 2 for OSPFv2.
- Type : OSPF packet type from 1 to 5, corresponding with hello, DD, LSR, LSU and LSAck respectively.
- Packet length : Total length of the OSPF packet in bytes, including the header.
- Router ID : ID of the advertising router.
- Area ID : ID of the area where the advertising router resides.
- Checksum : Checksum of the message.
- Autype : Authentication type from 0 to 2, corresponding with non-authentication, simple (plaintext) authentication and MD5 authentication respectively.
- Authentication : Information determined by authentication type. It is not defined for authentication type 0. It is defined as password information for authentication type 1, and defined as Key ID, MD5 authentication data length and sequence number for authentication type 2.

 Note : MD5 authentication data is added following an OSPF packet rather than contained in the Authentication field.

10.1.5.2 Hello packet

A router sends hello packets periodically to neighbors to find and maintain neighbor relationships and to elect the DR/BDR, including information about values of timers, DR, BDR and neighbors already known. The format is shown below :



Hello packet format

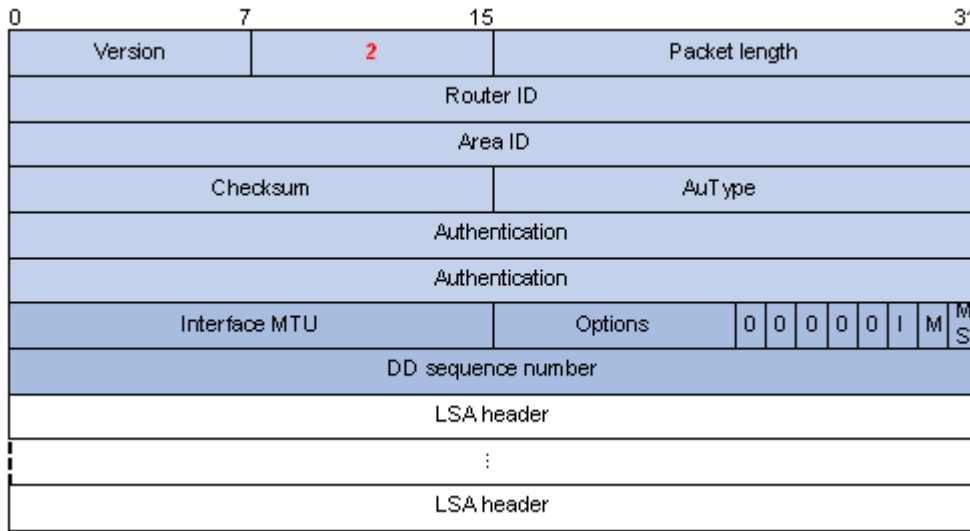
Major fields :

- Network Mask : Network mask associated with the router's sending interface. If two routers have different network masks, they cannot become neighbors.
- HelloInterval : Interval for sending hello packets. If two routers have different intervals, they cannot become neighbors.
- Rtr Pri : Router priority. A value of 0 means the router cannot become the DR/BDR.
- RouterDeadInterval : Time before declaring a silent router down. If two routers have different time values, they cannot become neighbors.
- Designated Router : IP address of the DR interface.
- Backup Designated Router : IP address of the BDR interface
- Neighbor : Router ID of the neighbor router.

10.1.5.3 DD packet

Two routers exchange database description (DD) packets describing their LSDBs for database synchronization, contents in DD packets including the header of each LSA (uniquely representing a LSA). The LSA header occupies small part of an LSA to reduce traffic between routers. The recipient checks whether the LSA is available using the LSA header.

The DD packet format :



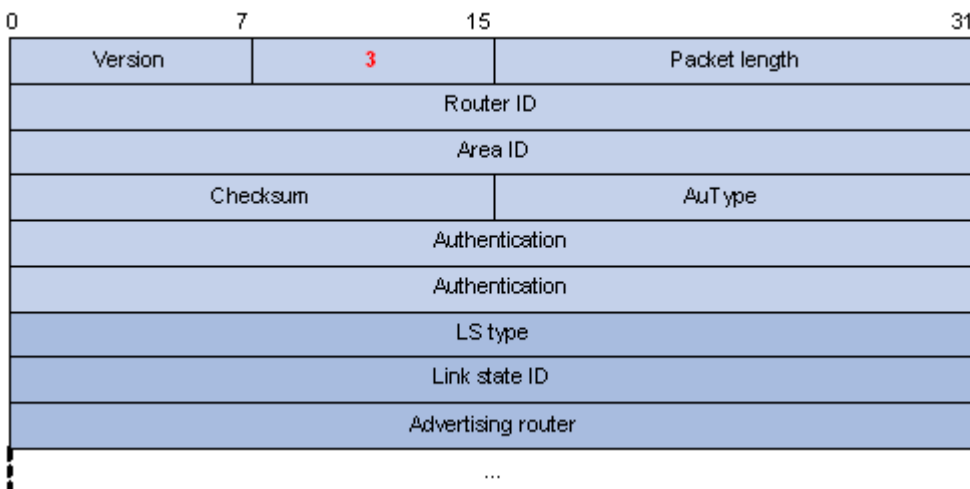
DD packet format

Major fields :

- Interface MTU : Size in bytes of the largest IP datagram that can be sent out the associated interface, without fragmentation.
- I (Initial) The Init bit, which is set to 1 if the packet is the first packet of database description packets, and set to 0 if not.
- M (More) : The More bit, which is set to 0 if the packet is the last packet of DD packets, and set to 1 if more DD Packets are to follow.
- MS (Master/Slave) : The Master/Slave bit. When set to 1, it indicates that the router is the master during the database exchange process. Otherwise, the router is the slave.
- DD Sequence Number : Used to sequence the collection of database description packets for ensuring reliability and intactness of DD packets between the master and slave. The initial value is set by the master. The DD sequence number then increments until the complete database description has been sent.

10.1.5.4 LSR packet

After exchanging DD packets, any two routers know which LSAs of the peer routers are missing from the local LSDBs. In this case, they send LSR (link state request) packets, requesting the missing LSAs. The packets contain the digests of the missing LSAs. The following figure shows the LSR packet format.



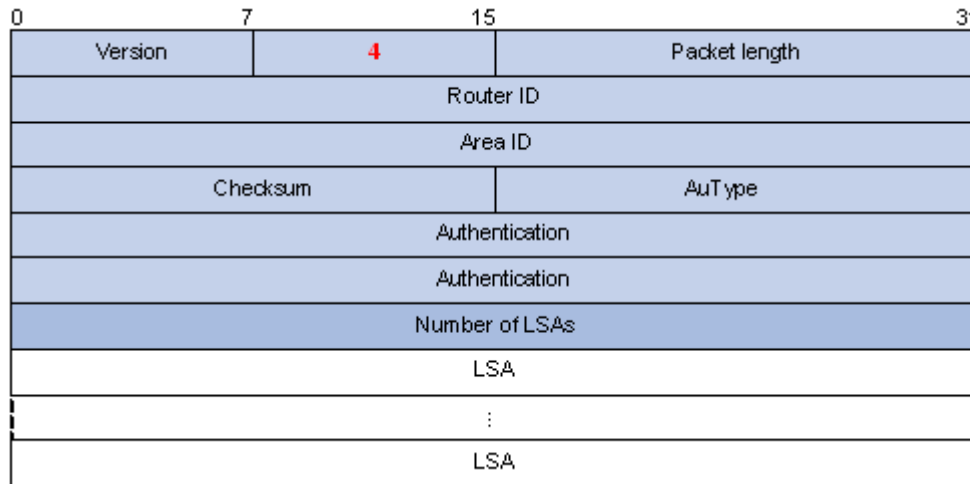
LSR packet format

Major fields :

- LS type : Type number of the LSA to be requested. Type 1 for example indicates the Router LSA.
- Link State ID : Determined by LSA type.
- Advertising Router : ID of the router that sent the LSA.

10.1.5.5 LSU packet

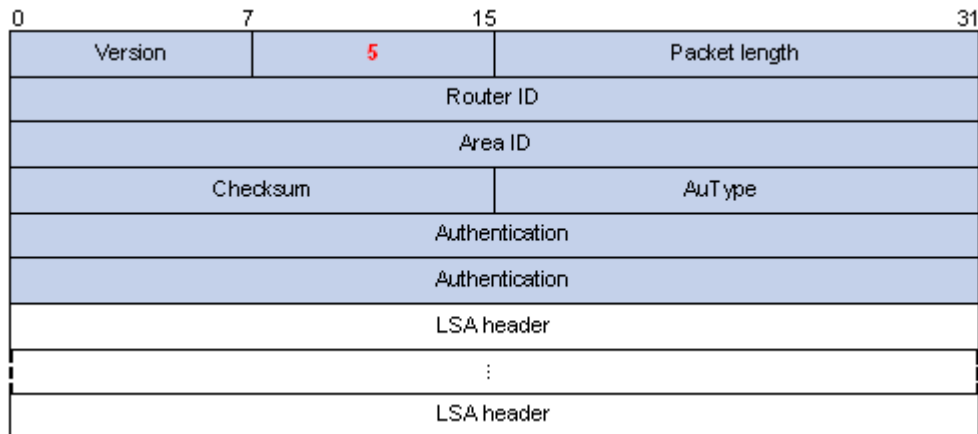
LSU (Link State Update) packets are used to send the requested LSAs to peers, and each packet carries a collection of LSAs. The LSU packet format is shown below.



LSU packet format

10.1.5.6 LSAck packet

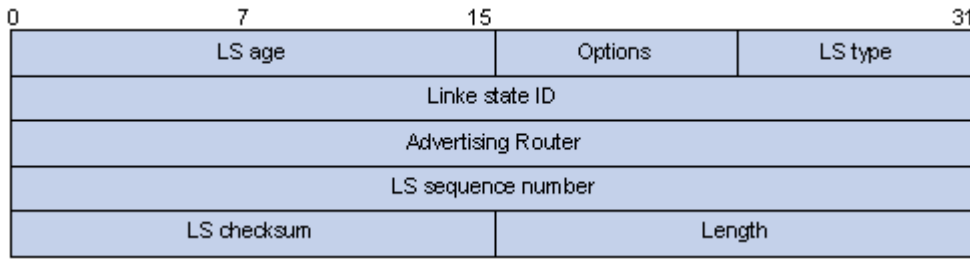
LSAck (Link State Acknowledgment) packets are used to acknowledge received LSU packets, contents including LSA headers to describe the corresponding LSAs. Multiple LSAs can be acknowledged in a single Link State Acknowledgment packet. The following figure gives its format.



LSAck packet format

10.1.5.7 LSA header format

All LSAs have the same header, as shown in the following figure.



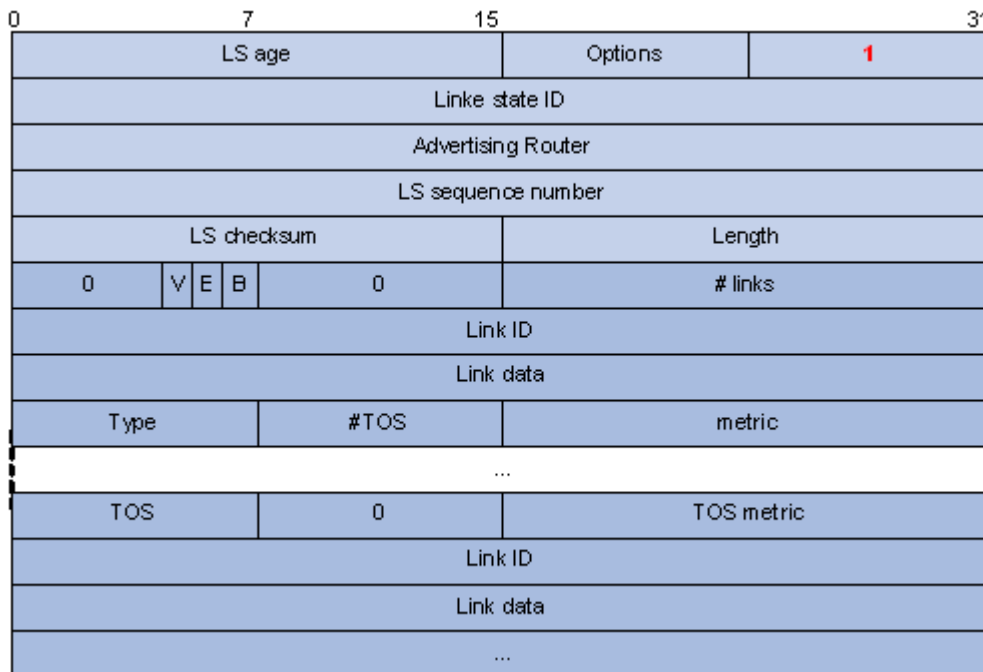
LSA header format

Major fields :

- LS age : Time in seconds elapsed since the LSA was originated. A LSA ages in the LSDB (added by 1 per second), but does not in transmission.
- LS type : Type of the LSA.
- Link State ID : The contents of this field depend on the LSA's type
- LS sequence number : Used by other routers to judge new and old LSAs.
- LS checksum : Checksum of the LSA except the LS age field.
- Length : Length in bytes of the LSA, including the LSA header.

10.1.5.8 Formats of LSAs

1) Router LSA



Router LSA format

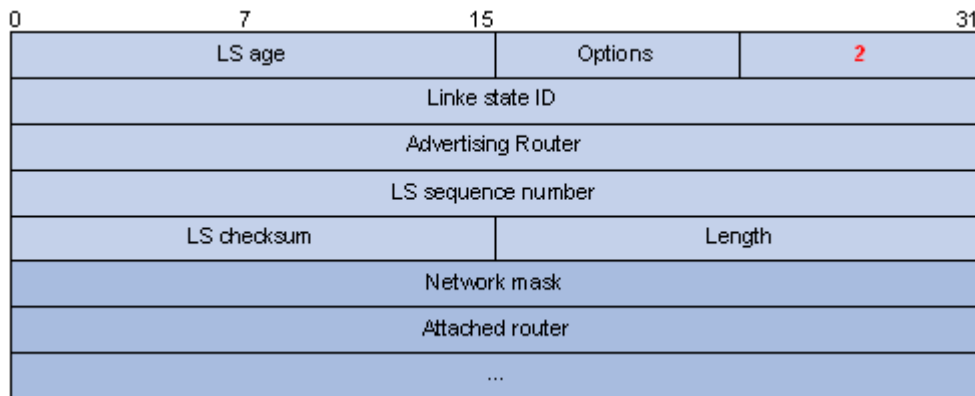
Major fields :

- Link State ID : ID of the router that originated the LSA.
- V (Virtual Link) : Set to 1 if the router that originated the LSA is a virtual link endpoint.
- E (External) : Set to 1 if the router that originated the LSA is an ASBR.
- B (Border) : Set to 1 if the router that originated the LSA is an ABR.
- # links : Number of router links (interfaces) to the area, described in the LSA.
- Link ID : Determined by Link type.
- Link Data : Determined by Link type.

- Type : Link type. A value of 1 indicates a point-to-point link to a remote router; a value of 2 indicates a link to a transit network; a value of 3 indicates a link to a stub network; a value of 4 indicates a virtual link.
- #TOS : Number of different TOS metrics given for this link.
- metric : Cost of using this router link.
- TOS : IP Type of Service that this metric refers to.
- TOS metric : TOS-specific metric information.

2) Network LSA

A Network LSA is originated by the DR on a broadcast or NBMA network. The LSA describes all routers attached to the network.



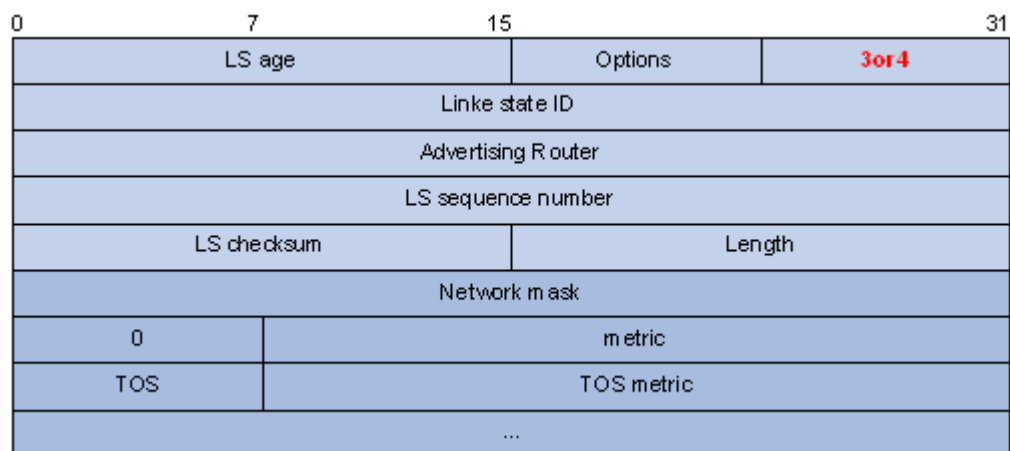
Network LSA format

Major fields :

- Link State ID : The interface address of the DR
- Network Mask : The mask of the network (a broadcast or NBMA network)
- Attached Router : The IDs of the routers, which are adjacent to the DR, including the DR itself

3) Summary LSA


Network summary LSAs (Type-3 LSAs) and ASBR summary LSAs (Type-4 LSAs) are originated by ABRs. Other than the difference in the Link State ID field, the format of type 3 and 4 summary-LSAs is identical.



Summary LSA format

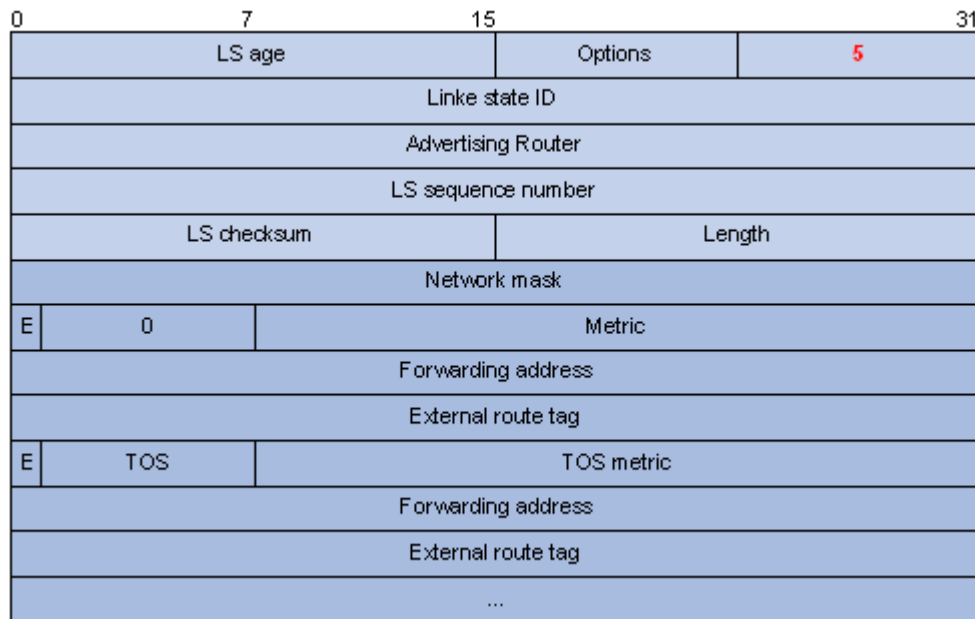
Major fields :

- Link State ID : For a Type-3 LSA, it is an IP address outside the area; for a type 4 LSA, it is the router ID of an ASBR outside the area.
- Network Mask : The network mask for the type 3 LSA; set to 0.0.0.0 for the Type-4 LSA
- metric : The metric to the destination

 Note : A Type-3 LSA can be used to advertise a default route, having the Link State ID and Network Mask set to 0.0.0.0.

4) AS external LSA

An AS external LSA originates from an ASBR, describing routing information to a destination outside the AS.



AS external LSA format

Major fields :

- Link State ID : The IP address of another AS to be advertised. When describing a default route, the Link State ID is always set to Default Destination (0.0.0.0) and the Network Mask is set to 0.0.0.0
- Network Mask : The IP address mask for the advertised destination
- E (External Metric) : The type of the external metric value, which is set to 1 for type 2 external routes, and set to 0 for type 1 external routes. Refer to Route types for description about external route types
- metric : The metric to the destination
- Forwarding Address : Data traffic for the advertised destination will be forwarded to this address
- External Route Tag : A tag attached to each external route. This is not used by the OSPF protocol itself. It may be used to manage external routes.

5) NSSA external LSA

An NSSA external LSA originates from the ASBR in a NSSA and is flooded in the NSSA area only. It has the same format as the AS external LSA.

0	7	15	31
LS age		Options	7
Link state ID			
Advertising Router			
LS sequence number			
LS checksum		Length	
Network mask			
E	TOS	Metric	
Forwarding address			
External route tag			
...			

NSSA external LSA format

10.1.6 Supported OSPF Features

10.1.6.1 Multi-process

With multi-process support, multiple OSPF processes can run on a router simultaneously and independently. Routing information interactions between different processes seem like interactions between different routing protocols. Multiple OSPF processes can use the same RID.

An interface of a router can only belong to a single OSPF process.

10.1.6.2 Authentication

OSPF supports authentication on packets. Only packets that pass the authentication are received. If hello packets cannot pass authentication, no neighbor relationship can be established.

The authentication type for interfaces attached to a single area must be identical. Authentication types include non-authentication, plaintext authentication and MD5 ciphertext authentication. The authentication password for interfaces attached to a network segment must be identical.

10.1.6.3 Hot Standby and GR

Distributed routers support OSPF Hot Standby (HSB). OSPF backs up necessary information of the Active Main Board (AMB) into the Standby Main Board. Once the AMB fails, the SMB begins to work to ensure the normal operation of OSPF.

OSPF supports to backup :

- All OSPF data to the SMB to make sure OSPF recovers normal operation immediately upon the AMB failure.
- Only the OSPF configuration information to the SMB. Once the AMB fails, OSPF will perform Graceful Restart (GR), obtaining adjacencies from and synchronizing the Link State Database with neighbors.

The Graceful Restart of the router is mainly used for High Availability (HA) and will not interfere with any other routers.

When a router shuts down, its neighbors will delete it from their neighbor tables and inform other routers, resulting in SPF recalculation. If the router restarts in several seconds, it is unnecessary to perform SPF recalculation, and reestablish adjacencies.

To avoid unnecessary SPF calculation, when a router restarts, it will inform neighboring routers the shutdown is temporary. Then these routers will not delete the router from their neighbor tables, and other routers have no idea about this restart.

After recovering to normal, the router obtains the Link State Database from neighboring routers via the GR related synchronization mechanism.

10.1.6.4 OSPF Graceful Restart

After an OSPF GR Restarter restarts OSPF, it needs to perform the following two tasks in order to re-synchronize its LSDB with its neighbors.

- To obtain once again effective OSPF neighbor information, supposing the adjacencies are not changed.
- To obtain once again LSDB contents.

Before the restart, the GR Restarter originates Grace-LSAs to negotiate the GR capability. During the restart, the GR Helpers continue to advertise their adjacencies with the GR Restarter.

After the restart, the GR Restarter will send an OSPF GR signal to its neighbors that will not reset their adjacencies with it. In this way, the GR Restarter can restore the neighbor table upon receiving the responses from neighbors.

After reestablishing the neighbor relationship, the GR Restarter will synchronize the LSDB and exchange routing information with all adjacent GR-capable neighbors. After that, the GR Restarter will update its own routing table and forwarding table based on the new routing information and remove the stale routes. In this way, the OSPF routing convergence is complete.

10.1.6.5 TE and DS-TETE

OSPF Traffic Engineering (TE) provides for the establishment and maintenance of Label Switch Paths (LSPs) of TE.

When establishing Constraint-based Routed LSPs (CR LSPs), MPLS obtains the TE information of links in the area via OSPF.

OSPF has a new LSA, Opaque LSA, which can be used for carrying TE information.

DiffServ Aware TE (DS-TE) provides for network resource optimization and allocation, flow classification, and indication of network bandwidth consumption of each flow in a link. TE is implemented on the classified type (thin granularity summarization type) rather than the summarized type (thick granularity summarization type) to improve performance and bandwidth utilization.

To support DS-TE application in MPLS, OSPF supports Local Overbooking Multiplier TLV and Bandwidth Constraint (BC) TLV.

10.1.6.6 IGP Shortcut and Forwarding Adjacency

IGP Shortcut and Forwarding Adjacency enable OSPF to use an LSP as the outbound interface for a destination. Without them, OSPF cannot use the LSP as the outbound interface.

Differences between IGP Shortcut and Forwarding Adjacency :

- If Forwarding Adjacency is enabled only, OSPF can also use an LSP as the outbound interface for a destination
- If IGP Shortcut is enabled only, only the router enabled with it can use LSPs for routing.

10.1.6.7 VPN

OSPF supports multi-instance, which can run on PEs in VPN networks.

In BGP MPLS VPN networks, multiple sites in the same VPN can use OSPF as the internal routing protocol, but they are treated as different ASs. An OSPF route learned by a site will be forwarded to another site as an external route, which leads to heavy OSPF routing traffic and management issues.

Configuring area IDs on PEs can differentiate VPNs. Sites in the same VPN are considered as directly connected. PE routers then exchange OSPF routing information like on a dedicated line; thus network management and OSPF operation efficiency are improved.

10.1.6.8 OSPF sham link

An OSPF sham link is a point-to-point link between two PE routers on the MPLS VPN backbone.

In general, BGP peers exchange routing information on the MPLS VPN backbone using the BGP extended community attribute. OSPF running on a PE at the other end utilizes this information to originate a Type-3 summary LSA as an inter-area route between the PE and CE.

If a router connects to a PE router in the same area and establishes an internal route (backdoor route) to a destination, in this case, since an OSPF intra-area route has a higher priority than a backbone route, VPN traffic will always travel on the backdoor route rather than the backbone route. To avoid this, an unnumbered sham link can be configured between PE routers, connecting the router to another PE router via an intra-area route with a lower cost.

10.1.7 Protocols and Standards

- RFC 1765 : OSPF Database Overflow
- RFC 2328 : OSPF Version 2
- RFC 3101 : OSPF Not-So-Stubby Area (NSSA) Option
- RFC 3137 : OSPF Stub Router Advertisement
- RFC 3630 : Traffic Engineering Extensions to OSPF Version 2

10.2 OSPF Configuration list

OSPF Configuration list is as following :

- Enable/disable OSPF
- Configure router ID
- Specify interface and area id
- Configure area authentication type
- Configure interface type
- Configure interface cost
- Configure priority when selecting DR
- Configure Hello time interval
- Configure interface invalid time of neighbour routers
- Configure retransmission LSA time interval of neighbor router
- Configure time needed when interface sending link state update packet
- Configure packet authentication key
- Configure STUB area of OSPF
- Configure route convergence in OSPF
- Configure OSPF virtual connection
- Configure route introduced by OSPF other route protocol
- Configure OSPF introduced default route
- Configure external route parameter received by OSPF

 OSPF monitor and maintain

10.2.1 Enable/disable OSPF

Configure it in global configuration mode :

Enable OSPF protocol

router ospf

Disable OSPF protocol

no router ospf

It is defaulted to disable OSPF.

10.2.2 Configure router ID

Router ID is a 32 byte intergeral number without symbols wglich is the unique sign of a router in autonomy system and user must configure it. Configuring router ID manually must guarantee the router ID of any two routers are different. Generally, configure router ID to be the same as the IP address of some interface of router.

Configure it in global configuration mode :

Configure router ID

router id *router-id*

Cancel router ID

no router id

To guarantee the stability of running OSPF, when programming network, be sure the division of router ID and configure manually.

By default, choose the smallest IP address from interface IP to be router ID.

Example :

!Configure router ID when switch running OSPF

QTECH(config)#router id 10.11.5.2

10.2.3 Specify interface and area id

OSPF protocol divided autonomy into different area which means dividing router to be different group. Some router will belong to different area (this kind of router is called Area Border Router ABR), and a neywork interface delongs to an area or every interface running OSPF protocol must use area ID to demonstrate which area belonged to. Different area uses ABR to transmit routing information.

In addition, all routers in the same area must be consensus the parameter configuration. Therefore, when configuring routers in the same area, most configuration data must be considered based on area and error configuration may cause the non-communication of neighbour routers or routing information congestion and self-ring. Configure it in OSPF protocol configuration mode :

Specify interface and area number

network *address wildcard-mask area area-id*

Cancel interface to run OSPF protocol

no network *address wildcard-mask area area-id*

After enabling OSPF, it should specify to be applied in which network interface and configure the area it belonged to. wildcard-mask can be IP address mask or the wildcard after NON the mask.

Example :

!Specify running OSPF in IP address 10.11.5.2

QTECH(config-router-ospf)#network 10.11.5.2 255.255.255.0 area 0.0.0.0

10.2.4 Configure area authentication type

Authentication type of all routers in an area must be the same(support plain text authentication, MD5 encrypt

authentication, not authentication)

Configure area authentication type

area *area-id* **authentication** [**message-digest**]

Restore interface authentication type to be non-authentication

no area *area-id* **authentication**

Example :

!Configure authentication of OSPF area 0 to be MD5

QTECH(config-router-ospf)#area 0 authentication message-digest

10.2.5 Configure interface type

OSPF protocol calculating route is based on neighbor network topology of current router. Each router describes the network topology of its neighbor network and transmits it to other routers. According to link layer protocol type, divide network into following 4 types :

Broadcast : when link layer protocol is Ethernet or FDDI, the network is defaulted to be Broadcast

Non Broadcast MultiAccess (NBMA) : when link layer protocol is ATM, the network is defaulted to be NBMA.

Point-to-Multipoint : none link layer protocol will be defaulted to be Point-to-Multipoint. Point-to-Multipoint must be changed from other types of network. Generally, change non-entire connectivity NBMA to Point-to-Multipoint.

Point-to-Point : when link layer protocol is PPP, LAPB or POS, the network is defaulted to be Point-to-Point.

NBMA network is non-broadcasting, point-to-multipoint network, specially as ATM. Configure poll-interval to send poll-interval Hello packet time range before specifying neighborhood established by router and neighbor routers.

In broadcasting network without multiple accessing, configure interface to be nonbroadcast.

If in NBMA network, not all routers can reach each other. Configure interface to be point-to-multipoint.

If there is one opposite end in NBMA network, configure interface to be point-to-point.

The difference between NBMA and point-to-multipoint :

In OSPF protocol, NBMA is connectivity, non-broadcasting, multipoint reaching network. Point-to-multipoint network need not entire connectivity.

In NBMA, it needs selecting DR and BDR , while in point-to-multipoint, there is no DR and BDR.

NBMA is a default network, such as : if link layer protocol is ATM , OSPF will defaulted to think the interface is NBMA (no matter this network is entire connectivity or not) . Point-to-multipoint is not default network type. None link protocol is thought to be point-to-multipoint. Point-to-multipoint must be forced to be changed from other network type. Generally, change non-entire connectivity to be point-to-multipoint.

NBMA uses unicast sending packet which needs configure neighbor manually. Point-to-multipoint uses multicast sending packet.

The link layer protocol of switch is Ethernet, OSPF thinks network type is broadcast. Generally, not change its network type,

Configure it in interface configuration mode.

Configure network type of interface

ip ospf network { **broadcast** | **non-broadcast** | **point-to-multipoint** | **point-to-point** }

Restore default network type of interface

no ip ospf network

Example :

!Configure VLAN interface 4 to be broadcast

QTECH(config-if-vlanInterface-4)#ip ospf network broadcast

10.2.6 Configure interface cost

User can configure cost of interface sending packet, or OSPF defaults cost to be 1. Configure it in interface configuration mode :

Configure cost of VLAN interface to send packet

ip ospf cost cost

Restore the default cost of VLAN interface to send packet

no ip ospf cost

Example :

!Configure cost of VLAN interface 3 to be 10

QTECH(config-if-vlanInterface-3)#ip ospf cost 10

10.2.7 Configure priority when selecting DR

The priority of router interface determines the competency in selecting “designated router”. The superior priority is firstly considered in conflict. Designated router (DR) is not determined by human, but selected by all routers in the network interface. The router in this network interface whose Priority > 0 can be the candidate. Choose the one with the superior priority to be the so called DR. If the priority is the same, choose the one with larger router ID. The vote is the Hello packet. Each router writes its own DR into Hello and sends it to each router in the network interface. When two of them declaring that they are the DR, choose the one with superior priority. If they have the same priority, choose the one with the larger router ID. The one with the priority being 0, he will not be selected to be DR or BDR.

If DR is failure because of some fault, routers must select DR again at the same time. It costs a long time. During this time, the calculation of router is not correct. In order to shorten it, BDR (Backup Designated Router) is brought up. BDR is a abackup for DR. Select BDR at the same time as DR. It establishes neighborhood and exchange routing information with the routers in the network interface. After the failure of DR, BDR is about to be DR because the neighborhood has been established. There will be reselected a new BDR which will not be effected the calculation of router though it needs a long time.

Caution :

DR is not always the router with the superlative priority and BDR is not always the one with the second superlative priority. After selecting DR and BDR, a new router adds, no matter how superlative its priority is, it will not be DR.

DR is the definition in a network interface which is for router interface. A router may be DR in an interface and may be BDR or DR other in another interface.

Selecting DR in broadcast or NBMA interface, it is unnecessary to select DR in poit-to-poit or poit-to-multipoit interface.

Configure it in interface configuration mode.

Configure the priority of interface to select “designated router”

ip ospf priority value

Restore the default value

no ip ospf priority

By default, the priority of VLAN interface to select “designated router” is in the range of 0 ~ 255

Example :

!Configure priority of VLAN interface 3 to be 100

QTECH(config-if-vlanInterface-3)#ip ospf priority 100

10.2.8 Configure Hello time interval

Hello packet is a geneally used packet which is periodically sent to neighbor router to search and maintain neighborhood and select DR and BDR. User can configure time interval of sending Hello packet. The smaller the hello-interval is, the faster the changes of network is found. The hello-interval of routers in the same network must be the same.

After enabling a router, it sends Hello packet to the neighbor node whose priority is larger than 0 (the routers can be selected as DR or BDR). After selecting of DR and BDR, they will send Hello packet to all neighbors to set up neighborhood. If neighborhood fails, router periodically send Hello packet according to the time interval of poll-interval command until neighbor router can be used again. The value of poll-interval is three times of the value of hello-interval. When the time interval of sending Hello packet is changed, configure the value of poll-interval.

Configure it in interface configuration mode :

Configure time interval of sending hello packet

ip ospf hello-interval *seconds*

Restore the default time interval of sending hello packet

no ip ospf hello-interval

By default, the time interval of point-to-point, broadcast interface sending Hello packet to be 10 seconds and point-to-multipoint, nonbroadcast interface sending Hello packet to be 30 seconds.

Example :

!Configure time interval of VLAN interface 3 sending hello packet to be 15

QTECH(config-if-vlanInterface-3)#ip ospf hello-interval 15

10.2.9 Configure interface invalid time of neighbour routers

The dead interval of OSPF neighbor is : in the time interval, if the Hello packet hasn't received, it is thought the neighbor is ineffective. dead-interval seconds must be 4 times of Hello-interval seconds, and the dead-interval must be the same in the same network interface.

Configure it in interface configuration mode :

Configure dead interval of neighbor routers.

ip ospf dead-interval *seconds*

Restore the default dead interval

no ip ospf dead-interval

The default dead interval of OSPF neighbor for Point-to-point and broadcast is 40 seconds; The default dead interval of OSPF neighbor for point-to-multipoint, non-broadcast is 120seconds.

Example :

!Configure the dead interval of interface 3 to be 60seconds

QTECH(config-if-vlanInterface-3)#ip ospf dead-interval 60



Caution : After modifying network type, hello-interval and dead-interval are restore to the default value.

10.2.10 Configure retransmission LSA time interval of neighbor router

When a router sending "Link Status Advertisement" (LSA), it needs to receive the confirm. If the confirm hasn't received in LSA retransmit interval, this LSA will be retransmit. User can configure retransmit-interval value.

Configure it in interface configuration mode :

Configure the retransmit interval of sending LSA between neighbour routers

ip ospf retransmit-interval *seconds*

Restore the default value of retransmit interval of sending LSA between neighbour routers

no ip ospf retransmit-interval

By default, the retransmit interval of sending LSA between neighbour routers is 5 seconds.

Example :

!Configure LSA retransmit interval of VLAN interface 3 to be 3 seconds.

QTECH(config-if-vlanInterface-3)#ip ospf retransmit-interval 3

10.2.11 Configure time needed when interface sending link state update packet

In LSU packet, the aging time of LSA will add a transmit-delay before sending. LSA will be aging (1 more minute per second) with time in Link Status DataBase (LSDB) of this router but it will not be aging in network transmission, so it is necessary to add the configured time before sending LSA. This configuration is very important in network with low speed.

Configure it in interface configuration mode :

Configure the time of sending LSU

ip ospf transmit-delay *seconds*

Restore the default LSU time

no ip ospf transmit-delay

By default, the time of sending LSU is 1 second.

Example :

!Configure LSA delay interval of VLAN interface 3 to be 3 seconds.

QTECH(config-if-vlanInterface-3)#ip ospf transmit-delay 3

10.2.12 Configure packet authentication key

OSPF supports simple or MD5 encryption authentication between neighbour routers.

Configure it in interface configuration mode :

Configure interface simple authentication key

ip ospf authentication-key *password*

Cancel interface simple authentication key

no ip ospf authentication-key

Configure interface MD5 authentication key

ip ospf message-digest-key *key-id md5 key*

Cancel interface MD5 authentication key

no ip ospf message-digest-key

By default, non-authentication is configured.

password is a character string of 1 ~ 8 bytes;

key-id is intergeral number between 0 ~ 255 ;

key is a character string of 1 ~ 16 bytes.

Example :

!Configure simple authentication key of VLAN interface 3 to be abc123

QTECH(config-if-vlanInterface-3)#ip ospf authentication-key abc123

10.2.13 Configure STUB area of OSPF

Stub area is special LSA area. ABR in stub area doesn't transmit the router outside of autonomy system. The scale of routing table and transmission number of routing packet in these areas will greatly reduced.

Stub area is optional configuration attribution, but not every area can suit the configuration condition. Generally, stub area locates in the edge of autonomy system which is the non-backbone area with one ABR; or there are many ABRs, but no virtual connection is configured between ABRs.

To guarantee the reachable of router out of autonomy system, ABR in this area will generate a default route (0.0.0.0) and distribute it to other non-ABR router in this area.

Pay attention to followings when configuring Stub area :

- Backbone area cannot configure to be Stub area and virtual connection cannot pass through Stub area.
- If configuring an area to be Stub area, all routers in this area must configure this attribution.
- There cannot be ASBR in Stub area, that is, external router cannot transmit in this area.

Configure it in OSPF protocol configuration mode :

Configure an area to be Stub area

area *area-id* stub [no-summary]

Cancel configured Stub area

no area *area-id* stub

Configure the cost to default router of Stub area

area *area-id* default-cost *cost*

Cancel the cost to default router of Stub area

no area *area-id* default-cost

By default, Stub area is not configured; the cost to default router of Stub area is 1.

There are two configuration command in STUB area : area stub and area default-cost. All routers connected to STUB area must use area stub command to configure to be STUB attribution. Command area default-cost only be effective in ABR configuration. This command can specify the cost for ABR to send default routing to STUB area.

For reducing the number of Link State Advertisement (LSA) sent to STUB, configure no-summary in ABR to forbid ABR to send summary LSAs (LSA type 3) to STUB area.

Example :

!Configure area 1.1.1.1 to be stub area and configure the cost to default router of Stub area to be 10

QTECH(config-router-ospf)#area 1.1.1.1 stub

QTECH(config-router-ospf)#area 1.1.1.1 default-cost 10

10.2.14 Configure route convergence in OSPF

Route convergence is : ABR can convergent the route information with the same prefix together and distribute one route to other area. One area can configure many convergent network interface so that OSPF can convergent many network interface. ABR sends route information to other area to generate Sum_net_Lsa (Type 3 LSA) with the unit of network interface. If there are some continuous network interface in area, use area range command to convergent these continous network interface to be one network interface. ABR sends one convergent LSA and LSA located in specified convergent network will not be sent separately which can reduce the scale of LSDB in other areas.

Convergent them to be one network interface : 202.38.0.0 255.255.0.0

Once adding convergent network interface of some network interface to area, the internal route whose IP address locates in this network interface will not be broadcasted to other area but broadcast the abstract information of the whole convergent network route. If the network interface is restricted by keyword not-advertise, the abstract information to this network route will not be broadcasted. This network is demonstrated by IP address/ mask. Receiving convergent network and restricting it will reduce the exchange volume of route information in areas.

Caution : Route convergence is effective in ABR configuration.

Configure it in OSPF protocol configuration mode :

Configure OSPF area route convergence

area *area-id* range *address mask* [advertise | notadvertise]

Cancel OSPF area route convergence

no area *area-id* range *address mask*

By default, route in areas will not be convergent.

Example :

!Convergent 202.38.160.0 255.255.255.0 and 202.38.180.0 255.255.255.0 to be one route 202.38.0.0 255.255.0.0

QTECH(config-router-ospf)#area 1.1.1.1 range 202.38.0.0 255.255.0.0

10.2.15 Configure OSPF virtual connection

After dividing SOPF areas, not all areas are equal. One area with the area-id being 0.0.0.0 is different which is called BackboneArea. The update of OSPF route in non- BackboneArea is through BackboneArea. OSPF protocol regulates : all non- BackboneArea must be connected with BackboneArea, that is, there must be at least one interface of ABR in area 0.0.0.0. If there is an area which is not physically connected with BackboneArea 0.0.0.0, there must establish a virtual connection.

If the physical connection cannot be proved because of the restriction of network topology, create virtual connection. Virtual connection means two ABRs set up a physical connection through interbal route area of a non-Back Bone Area. The ends must be ABR and it can be effective when configuring at both two sides. Virtual connection is marked by route ID of the opposite end. The internal area supported a non-Back Bone Area for the two ends of the virtual connection is called Transit Area and its area number must be demonstrated when configuring.

Virtual connection is activated after the calculation of transmitting area route which equals to form point-to-point link between two ends, so in this connection, it can also configure interface parameter as physical interface, such as sending HELLO packet interval.

“Logical channel” means several routers running OSPF between two ABRs which only transmit packet (the destination address of protocol packet are not these routers, so packet is transmitted as general IP packet) and two ABRs can straightly transmit router information. Here, router information means type 3 LSA generated by ABR, so the synchronization of routers in area do not changed.

Caution : If autonomy system is divided to be one or more areas, there must be one Back Bone area to guarantee the straightly or logically connection between other areas and Back Bone area and Back Bone area itself must be connected.

Configure it in OSPF protocol configuration mode :

Create and configure virtual connection

```
area area-id virtual-link router-id [ hello-interval seconds ] [ retransmit- interval seconds ]
[ transmit-delay seconds ] [ dead-interval seconds] { [ authentication-key key ] |
[ message-digest-key keyid md5 key ] }
```

Cancel created virtual connection

```
no area area-id virtual-link router-id [ hello-interval seconds ] [ retransmit- interval seconds ]
[ transmit-delay seconds ] [ dead-interval seconds] { [ authentication-key key ] |
[ message-digest-key keyid md5 key ] }
```

By default, area-id and router-id has no default value ; the value of hello-interval is 10 seconds ; the value of retransmit-interval is 5 seconds ; the value of transmit-delay is 1 second ; the value of dead-interval is 40 seconds.

Example :

!Configure a virtual connection with the transmission area being 1.1.1.1 , router-id of the opposite end being 10.11.5.2

```
QTECH(config-router-ospf)#area 1.1.1.1 virtual-link 10.11.5.2
```

10.2.16 Configure route introduced by OSPF other route protocol

Each dynamic routing protocol can share routing information. Because of OSPF, router found by other routing protocol always be handled as external routing information of autonomy.

OSPF uses following 4 kinds of different router which as priority order are :

- . Inter Area Routing
- . Area Border Routing
- . The first category external routing
- . The second category external routing

The description of routing in or between areas for network structure in Autonomy system. External routing describes how to choose destination routing out of Autonomy system.

The first category external routing is received IGP router (such as : RIP and STATIC). This kind of router is more credible, so the cost volume of external router and autonomy system is the same and can compare with the router of OSPF itself, that is, the cost to external router = the cost to its ASBR + the cost of ASBR to destination address.

The second category external routing is the received EGP router. This kind of router is less credible, so the cost volume of ASBR to the outside of autonomy system is far more expensive than that of autonomy system to ASBR, so the former is mainly considered, that is, the cost to the second external router = the cost of ASBR to destination address. If the cost is the same, consider the cost of this router to corresponded ASBR.

Configure it in OSPF protocol configuration mode :

Introduce route information of other protocol

redistribute *protocol* [**metric** *metric*] [**type** *1 | 2*] [**tag** *tag-value*]

Cancel route information of other protocol

no redistribute *protocol*

By default , OSPF doesn't introduce route information of other protocol.

protocol means introduced source routing protocol which can be connected, rip, static, RIP, IS-IS and BGP.

Example :

!Configure OSPF introduce RIP router

QTECH(config-router-ospf)#redistribute rip

10.2.17 Configure OSPF introduced default route

Use redistribute static command cannot introduce default routing. Use default-information originate command to introduce default router to OSPF routing domain.

Configure it in OSPF protocol configuration mode :

Introduce default route to OSPF

default-information originate [**always**] [**metric** *metric-value*] [**type** *type-value*]

Cancel introduced default route

no default-information originate

By default , OSPF will not introduce any default route.

Example :

!Configure OSPF introduce default route

QTECH(config-router-ospf)#default-information originate always

10.2.18 Configure external route parameter received by OSPF

When OSPF introducing route information found by other route protocol to autonomy system, configure some extra parameter, such as the default cost and mark of introduced router. Route mark can used to mark protocol related information,, such as the number to distinguish autonomy system when OSPF receiving BGP.

Configure it in OSPF protocol configuration mode :

Configure default cost when OSPF receiving external route

default redistribute metric *metric*

Restore metric of received external route

no default redistribute *metric*

Configure default type when OSPF receiving external route

default redistribute type { *1 | 2* }

Restore default type of received external route

no default redistribute type

By default, the metric of received external route is 1 and type is 2.

Example :

!Configure the metric of received external route to be 10

QTECH(config-router-ospf)#default redistribute metric 10

10.2.19 OSPF monitor and maintain

Followings are display command :

show ip ospf

Display OSPF information.

show router id

Display configured router ID

show ip ospf neighbor	Display OSPF neighbor.
show ip ospf database	Display OSPF LSDB.
show ip ospf virtual-link	Display OSPF virtual link
show ip ospf border-routers	Display OSPF edge router
show ip ospf interface	Display OSPF interface
show ip route ospf	Display OSPF routing table.
show ip ospf cumulative	Display OSPF statistic.
show ip ospf error	Display OSPF error
show ip ospf request-list	Display OSPF request list.
show ip ospf retrans-list	Display OSPF retransmit list

These commands can be used in any configuration mode.

Example :

```
!Display OSPF information
```

```
show ip ospf
```

```
QTECH(config-router-ospf)#show ip ospf
```

```
!Display OSPF neighbor information
```

```
show ip ospf neighbor
```

```
QTECH(config-router-ospf)#show ip ospf neighbor
```

```
!Display OSPF virtual link information
```

```
show ip ospf virtual-link
```

```
QTECH(config-router-ospf)#show ip ospf virtual-link
```

```
!Display LSDB information
```

```
show ip ospf database
```

```
QTECH(config-router-ospf)#show ip ospf database
```

Chapter 11 BGP Configuration

11.1 Brief Introduction of BGP

BGP(Border Gateway Protocol)is a dynamic route protocol in anonymous system whose basic function is automatically changing no-loop route information in AS and constructs topology from AS through changing network layer arrival information with AS routine attribution.

The early published BGP standard document are RFC1105 (BGP-1) ,RFC1163 (BGP-2) and RFC1267 (BGP-3). The RFC1771(BGP- 4)is commonly used and RFC4271(BGP- 4) is the latest edition which is suitable for distributed structure and supports CIDR (Classless InterDomain Routing) . BGP can also be used in strategy of carrying out user's configuration. BGP-4 quickly becomes the real standard of Internet outer route protocol. BGP is mostly used between ISP.

The specialty of BGP is as following:

BGP is an outer route protocol which is different from such inner route protocol as OSPF,RIP. It is used for controlling route transmission and choosing better route, not searching and calculating route.

Completely solve route cycling through AS path information in BGP route.

Use TCP as tis transmission layer protocol to improve reliability.

BGP-4 supports CIDR, which is an important improvement of BGP-3. CIDR treats IP address without classifying category A, B and C. The introduction of CIDR simplified Route Aggregation, which is the process of combining different routes and several routes become one to reduce the resource utilization and BGP cost.

When update route, BGP sends adding route to reduce the bandwidth of BGP broadcasting route to transmit plenty of route information in internet.

For management and security, BGP-4 provides abundant route strategy to realize agile filtration and choice.

BGP operates in s specific router as a high layer protocol. BGP router exchanges route information through sending whole BGP table and handle route changes through Update message. Sending and receiving keepalive to detect the connection.

The router sending BGP message is called BGP speaker which receives and generates new route information and send advertise to pther BGP speaker. When others receiving the new advertising, this route is better than the current one, or it hasn't received the route, this new advertising will be broadcasted to all other BGP speakers and the BGP speaker who changes the message to others is called peer.

BGP runs in router with following two ways:

IBGP (Internal BGP)

EBGP (External BGP)

When running in the internal of the same AS, it is called IBGP ; when running in different AS, is called EBGP.

BGP runs through message. The message can be divided into 4 kinds:

Open message

Update message

Notification message

Keepalive message

Open message is the first message after TCP connection, which is used to set up the relationship between BGP and peer. Notification message is error notification message. Keepalive message is used to detect the efficiency of connection. Update message is the mpsst important message for exchanging route information. Update message consists of unreachable route, path attributes and NLRI , Network Layer Reachability Information.

11.1.1 BGP Message Type

11.1.1.1 Format of a BGP packet header

BGP is message-driven. There are five types of BGP packets: Open, Update, Notification, Keepalive, and Route-refresh. They share the same packet header, the format of which is shown by Figure 1.

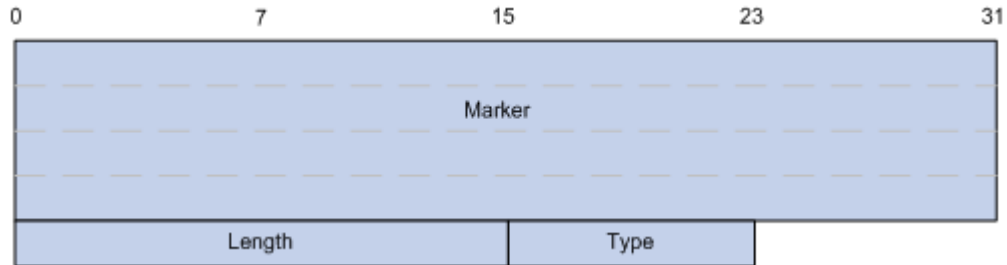


Figure 1 Packet header format of BGP messages

The fields in a BGP packet header are described as follows.

- Marker: 16 bytes in length. This field is used for BGP authentication. When no authentication is performed, all the bits of this field are 1.
- Length: 2 bytes in length. This field indicates the size (in bytes) of a BGP packet, with the packet header counted in.
- Type: 1 byte in length. This field indicates the type of a BGP packet. Its value ranges from 1 to 5, which represent Open, Update, Notification, Keepalive, and Route-refresh packets. Among these types of BGP packets, the first four are defined in RFC1771, and the rest one is defined in RFC2918.

11.1.1.2 Open

Open message is used to establish connections between BGP speakers. It is sent when a TCP connection is just established. Figure 2 shows the format of an Open message.

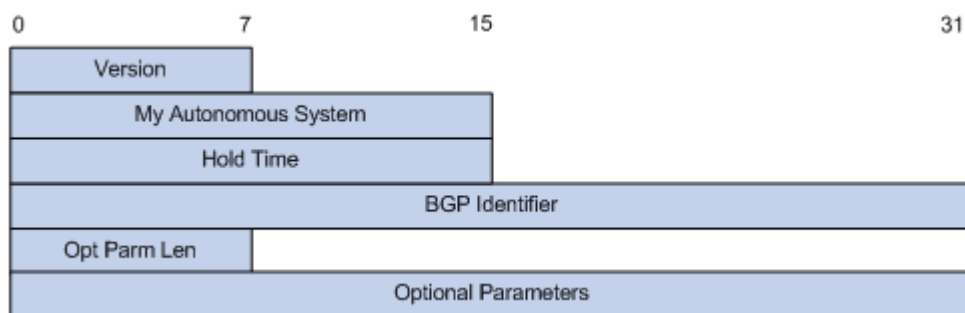


Figure 2 BGP Open message format

The fields are described as follows.

- Version: BGP version. As for BGP-4, the value is 4.

- My Autonomous System: Local AS number. By comparing this field of both sides, a router can determine whether the connection between itself and the BGP peer is of EBGP or IBGP.
- Hold time: Hold time is to be determined when two BGP speakers negotiate for the connection between them. The Hold times of two BGP peers are the same. A BGP speaker considers the connection between itself and its BGP peer to be terminated if it receives no Keepalive or Update message from its BGP peer during the hold time.
- BGP Identifier: The IP address of a BGP router.
- Opt Parm Len: The length of the optional parameters. A value of 0 indicates no optional parameter is used.
- Optional Parameters: Optional parameters used for BGP authentication or multi-protocol extensions.

11.1.1.3 Update

Update message is used to exchange routing information among BGP peers. It can propagate a reachable route or withdraw multiple pieces of unreachable routes. Figure 3 shows the format of an Update message.

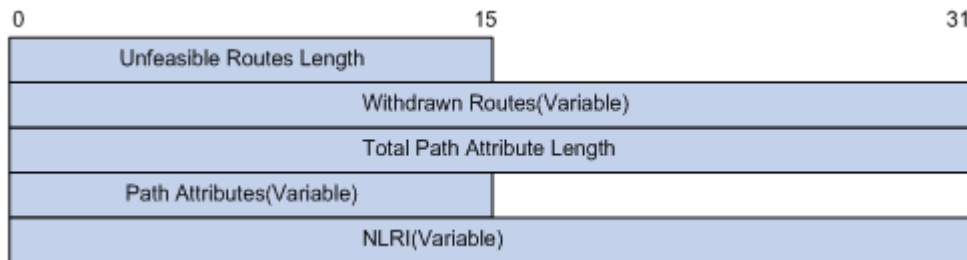


Figure 3 BGP Update message format

An Update message can advertise a group of reachable routes with the same path attribute. These routes are set in the NLRI (network layer reachability information) field. The Path Attributes field carries the attributes of these routes, according to which BGP chooses routes. An Update message can also carry multiple unreachable routes. The withdrawn routes are set in the Withdrawn Routes field.

The fields of an Update message are described as follows.

- Unfeasible Routes Length: Length (in bytes) of the unreachable routes field. A value of 0 indicates that there is no Withdrawn Routes field in the message.
- Withdrawn Routes: Unreachable route list.
- Total Path Attribute Length: Length (in bytes) of the Path Attributes field. A value of 0 indicates that there is no Path Attributes field in the message.
- Path Attributes: Attributes list of all the paths related to NLRI. Each path attribute is a TLV (Type-Length-Value) triplet. In BGP, loop avoidance, routing, and protocol extensions are implemented through these attribute values.
- NLRI (Network Layer Reachability Information): Contains the information such as reachable route suffix and the corresponding suffix length.

11.1.1.4 Notification

When BGP detects an error state, it sends the Notification message to peers and then tears down the BGP connection.

Figure 4 shows the format of an Notification message.

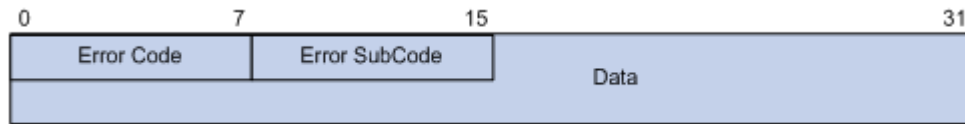


Figure 4 BGP Notification message format

The fields of a Notification message are described as follows.

- Error Code: Error code used to identify the error type.
- Error Subcode: Error subcode used to identify the detailed information about the error type.
- Data: Used to further determine the cause of errors. Its content is the error data which depends on the specific error code and error subcode. Its length is unfixed.

11.1.1.5 Keepalive

In BGP, Keepalive message keeps BGP connection alive and is exchanged periodically. A BGP Keepalive message only contains the packet header. No additional fields is carried.

11.1.1.6 Route-refresh

Route-refresh messages are used to notify the peers that the route refresh function is available and request the peers to resend the routing information of the specified address family. Figure 5 shows the format of a route-refresh message.

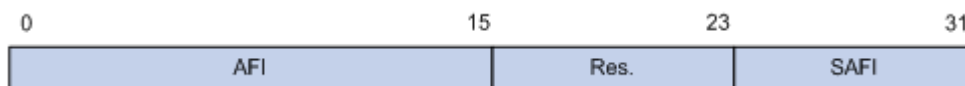


Figure 5 Route-refresh message format

The fields of a route-refresh message are described as follows.

- AFI: Address family identifier
- Res: Reserved. This field must not be set.
- SAFI: Subsequent address family identifier

11.1.2 BGP Route Attributes

11.1.2.1 Routes attributes classification

BGP route attributes describe route, so that BGP can filter and choose the routes.

In fact, all the BGP route attributes can be classified into the following four categories.

- Well-known mandatory attributes, which can be identified by any BGP routers. Route attributes of this type are carried in Update messages. Without these attributes, routing information goes wrong.
- Well-known discretionary attributes, which can be identified by any BGP routers. An Update message can travel with or without this type of attributes.
- Optional transitive attributes, which can be transmitted among ASs. Although attributes of this type may not be supported by any BGP routers, routes with them can still be received and be forwarded to BGP speakers.
- Optional non-transitive attributes, which is dropped on the BGP routers that do not support them. In this case, the attributes are not forwarded to other BGP routers.

Table 1 lists basic BGP route attributes and the categories they belong to.

BGP route attribute	Category
ORIGIN	Well-known mandatory
AS_PATH	Well-known mandatory
NEXT_HOP	Well-known mandatory
LOCAL_PREF	Well-known discretionary
ATOMIC_AGGREGATE	Well-known discretionary
AGGREGATOR	Optional transitive
COMMUNITY	Optional transitive
MULTI_EXIT_DISC(MED)	Optional non-transitive
ORIGINATOR_ID	Optional non-transitive
CLUSTER_LIST	Optional non-transitive

Table 1 BGP route attributes and the corresponding categories

11.1.2.2 Primary route attributes

1) ORIGIN

The ORIGIN attribute holds the source of routing information. It indicates how a route becomes a BGP route. The following describes the possible values of the ORIGIN attribute.

- IGP: BGP routes with their ORIGIN attributes being IGP have the highest priority. They are added to the BGP routing table through the **network** command.
- EGP: BGP routes with their ORIGIN attributes being EGP are obtained through EGP.
- Incomplete: BGP routes with their ORIGIN attributes being Incomplete have the least priority. This value does not indicate that the BGP route is unreachable; it means the source of the BGP route cannot be determined. The ORIGIN attribute of a BGP route imported through the **import-route** command is Incomplete.

2) AS_PATH

The AS_PATH attribute holds the numbers of all the ASs that a route passes from the source to the destination. AS numbers in this attribute are in the order the route passes the ASs. Before a BGP speaker advertises a route to the

BGP speakers of other ASs, it adds the local AS number to the head of the AS number queue in the AS_PATH attribute. According to the AS_PATH attribute of a received BGP route, a router can retrieve the information about the ASs the route passes. In AS_PATH attribute, AS numbers are listed by the distances between the ASs and the local AS. The number of the AS that is closest to the local AS is listed in the head, as shown in Figure 6.

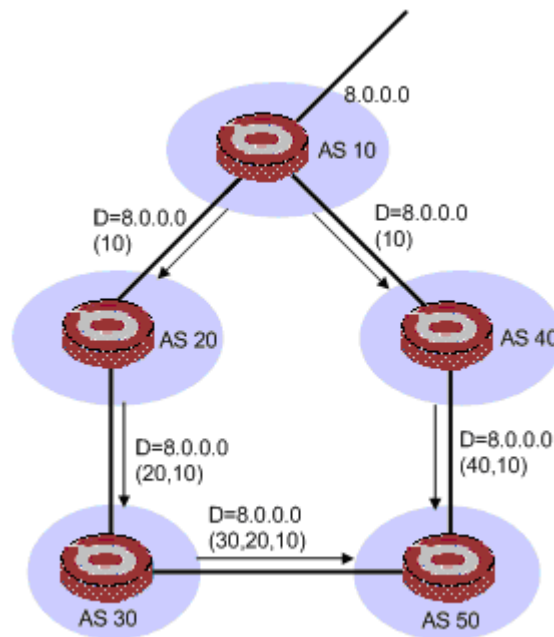


Figure 6 AS_PATH attribute

Normally, a router with BGP employed discards the routes that contain local AS number in the AS_PATH attribute. This eliminates routing loops.



Note

You can use the **peer allow-as-loop** command to allow AS number repetition to meet some special needs.

AS_PATH attribute can also be used to choose and filter routes. BGP chooses the routes containing less AS numbers with shorter path under the same circumstances. For example, in Figure 6, the BGP router in AS50 will choose the path passing through AS40 as the route to the router in AS 10.

In some applications, you can increase the number of AS numbers a BGP route contains through routing policy to control BGP routing in a flexible way.

By configuring AS path filtering list, you can have BGP routes filtered by the AS numbers contained in the AS-Path attribute.

3) NEXT_HOP

Different from that of the IGP, the NEXT_HOP attribute of a BGP route does not necessarily holds the IP address of the neighbor router.

The NEXT_HOP attribute is set in the following ways.

- When a BGP speaker advertises a route generated by itself to all its neighbors, it sets the NEXT_HOP attribute of the routing information to the address of its own interface connecting to the peer.
- When a BGP speaker sends a received route to one of its EBGP peer, it sets the NEXT_HOP attribute of the routing information to the address of its interface connecting to the EBGP peer.
- When a BGP speaker sends a route received from one of its EBGP peer to one of its IBGP neighbor, it does not change the NEXT_HOP attribute of the routing information. But with load balancing enabled, the NEXT_HOP attribute is changed when the BGP route is sent to a IBGP neighbor.

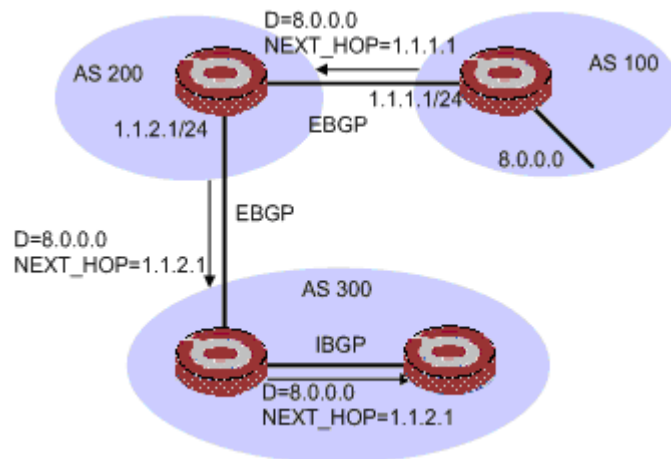


Figure 7 The NEXT_HOP attribute

4) MED (MULTI_EXIT_DISC)

The MED attribute is only valid between two neighboring ASs. The AS receiving this attribute will not advertise this attribute to a third AS.

The MED attribute is used to determine the optimal route for traffic flows to enter an AS. It acts the same as the metrics used in IGP. For multiple routes a BGP router receives from different EBGP peers, if they have the same destination address but different next hops, the route with the smallest MED value is chosen as the optimal route provided other conditions are the same. As shown in Figure 8, Router B is chosen as the ingress for traffic from AS 10 to AS 20.

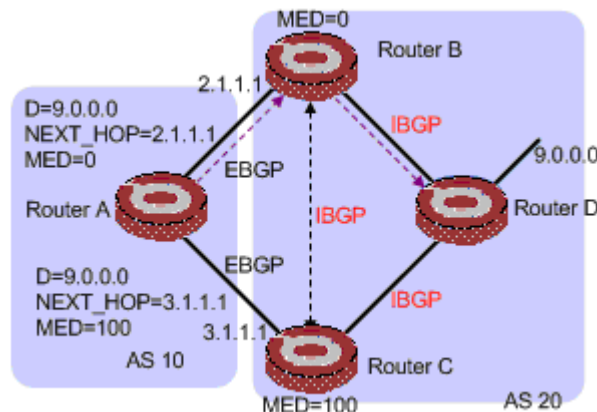


Figure 5-8 MED attribute

Normally, BGP only compares the MED attribute values of the routes received from the same AS.



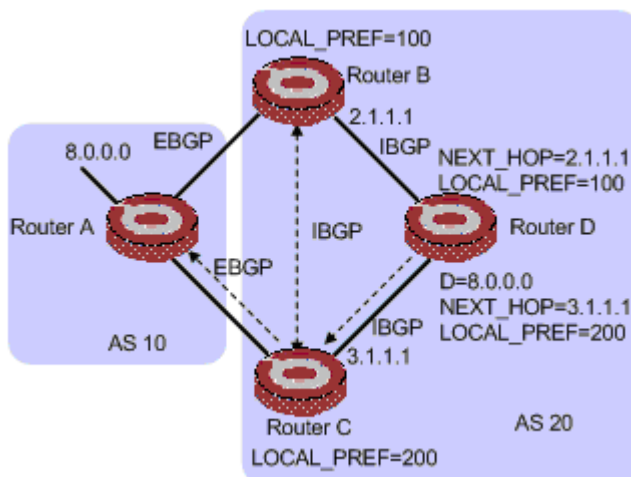
Note

You can force BGP to compare MED values of routes coming from different ASs.

5) LOCAL_PREF

The LOCAL_PREF attribute is only valid among IBGP peers. It is not advertised to other ASs. It indicates the priority of a BGP router.

LOCAL_PREF attribute is used to determine the optimal route for traffic leaving an AS. For multiple routes a BGP receives from different IBGP peers, if they have the same destination address but different next hops, the route with the smallest LOCAL_PREF value is chosen as the optimal route provided other conditions are the same. As shown in Figure 9, RouterC is chosen as the egress for traffic from AS 20 to AS 10.

**Figure 9** LOCAL_PREF attribute

7) COMMUNITY

The Community attribute is used to simplify routing policy application and ease the maintenance and management of routing policy. Community is a set of destination addresses with the same features. It is not restricted to physical boundary and is independent of AS. The Community attribute can be one of the following.

- INTERNET. By default, the value of the COMMUNITY attributes of all routes is INTERNET. That is, all routes belong to the Internet community by default. Routes with this attribute can be advertised to all BGP peers.
- No_EXPORT. Routes with this attribute cannot be sent to routers outside the local AS. With the presence of the confederation, routes of this kind

cannot be advertised outside the confederation, they can only be advertised in the sub-ASs in the confederation.

- No_ADVERTISE. Routes with this attribute cannot be advertised to any other BGP peers after being received by a BGP router.
- No_EXPORT_SUBCONFED. Routes with this attribute can neither be advertised outside the local AS nor be advertised to other sub-ASs inside the confederation after being received.

11.1.3 BGP Routing Policy

11.1.3.1 BGP routing policy

A BGP router filters routes in the following order.

- Drops the NEXT_HOP unreachable route.
- With Preferred-value specified, chooses the route with highest Preferred-value value.
- Prefers the route with highest LOCAL_PREF value.
- Prefers the routes starting from the local router.
- Prefers the route with the shortest AS path.
- Chooses routes in the order of the route ORIGIN type, that is, the order of IGP, EGP, and Incomplete.
- Prefers the route with the lowest MED value.
- Chooses the route learnt from EBGP, the route learnt from confederation and the route learnt from the IBGP in turn.
- Prefers the route with the smallest ORIGINATOR ID.
- Prefers the route with the smallest router ID.

11.1.3.2 BGP route advertising policy

A BGP router adopts the following policies to advertise routes.

- Sends the optimal route to its peers when multiple valid routes exist.
- Sends only the routes used by itself to its peers.
- Sends all the EBGP routes to all its BGP peers, including the EBGP peers and IBGP peers.
- Does not send IBGP routes to its IBGP peers.
- Sends IBGP routes to its EBGP peers.
- Sends all its BGP routes to the new peer once a new BGP connection is established.

11.1.4 Problems in Large-Scale BGP Networks

11.1.4.1 Route summarization

BGP routing tables in a large-scale network may be huge in size. Route summarization can largely diminish the size of a routing table.

Route summarization aggregates multiple routes to one route. It enables a BGP router to replace multiple specific routes with one summary route.

The switches support automatic route summarization and manual route summarization. In the manual route summarization mode, you can control the attribute of the summary routes and determine whether to send the specific routes or not.

11.1.4.2 BGP route dampening

BGP route dampening is used to solve the problem of route instability. Route instability mainly takes the form of route flaps, that is, a route appears and disappears repeatedly in the routing table.

When route flaps occur, a route sends route update to its neighbors. Routers receiving the update packets calculate the route over again and renew the routing table. Therefore, frequent route flaps consume much bandwidth and CPU time. They even affect the operation of network.

In most cases, BGP is applied in complicated networks where route changes are frequent. In order to avoid the unfavorable affection caused by route flaps, BGP uses route dampening to suppress the instable routes.

BGP route dampening uses penalty value to judge the stability of a route. A higher penalty value indicates a more instable route. Each time a route flaps, BGP adds a certain penalty value (fixed to 1000) to the route. When the penalty value exceeds the suppression threshold, the route will be suppressed and will neither be added to the routing table nor send update packets to other BGP peers.

The penalty value of a suppressed route is decreased by half in each specific period known as half-life. When the penalty value is decreased to a value less than the reuse threshold, the route gets valid and is added to the routing table again. At the same time, the BGP router sends corresponding update packets to its BGP peers.

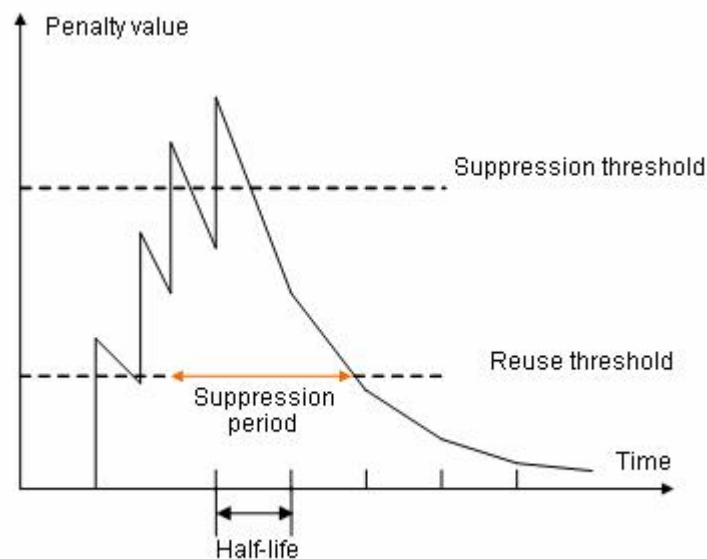


Figure 10 Diagram for BGP route dampening

11.1.4.3 Peer group

Peer group is a set of peers that are the same in certain attributes. When a peer joins into a peer group, the peer obtains the same configurations with those of the peer group. When the configuration of a peer group changes, those of the group members change accordingly.

A large-scale network can contain large amount of peers, lot of which adopt the same policies. Peer group simplifies your configuration when you configure peers adopting the same policy.

As the peers in a peer group adopt the same route updating policy, peer group gains more efficiency in route advertising.



Caution

If a BGP peer and the peer group containing the BGP peer are configured differently, the last configuration takes effect.

11.1.4.4 Community

Different from peer group, you can apply the same policy to BGP routers residing in different ASs through community. Community is a route attribute transmitted among BGP peers. It is independent of AS.

Before sending a route with the COMMUNITY attribute to its peers, a BGP router can change the original COMMUNITY attribute of the route.

Besides the well-known COMMUNITY attributes, you can also use the COMMUNITY attributes list to customize extended COMMUNITY attributes, so as to control the routing policy with more flexibility.

11.1.4.5 Router reflector

To ensure the connectivity among the IBGP peers in an AS, you need to make the IBGP peers fully connected. For an AS with the number of the routers in it being n , you need to establish at least $n*(n-1)/2$ IBGP connections to make them fully connected. This requires large amount of network resources and CPU time if large amount of IBGP peers exist in the AS.

You can decrease the use of network resources and CPU time through route reflection in this case. That is, use a router as a router reflector (RR) and establish IBGP connections between the RR and other routers known as clients. Routing information exchanged between the clients is passed/reflected by the RR. This eliminates the need to establish IBGP connections among the clients.

Note that a BGP router which is neither the RR nor a client is called a non-client. Non-clients and the RR must be fully connected, as shown in Figure 11.

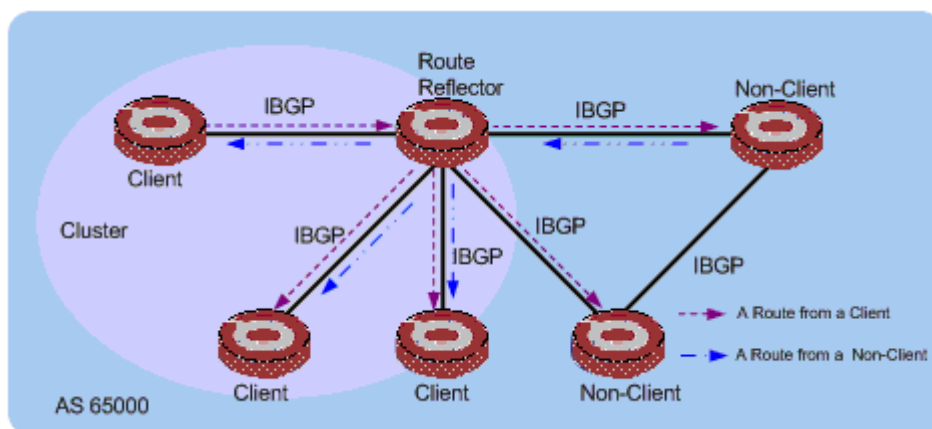


Figure 11 Diagram for the route reflector

An RR and all its clients form a cluster. To ensure network reliability and avoid single-point failure, you can configure more than one RR in a cluster. In this case, make sure all the RRs in the cluster are configured with the

same cluster ID to avoid routing loops. Figure shows a cluster containing two RRs.

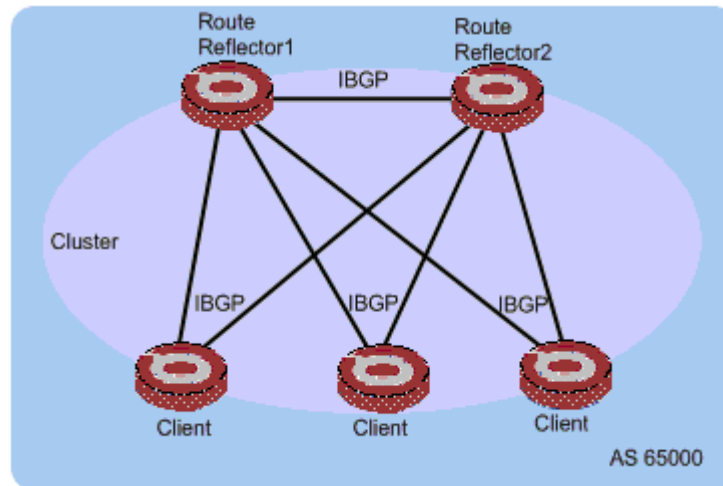


Figure 12 A cluster containing two RRs

RR is unnecessary for clients that are already fully connected. You can disable routing information reflection using corresponding commands provided by the switches.



Note

The configuration to disable routing information reflection only applies to clients. That is, routing information can still be reflected between a client and a non-client even if you disable routing information reflection.

11.1.4.6 Confederation

Confederation is another way to limit the number of IBGP connections in an AS. It divides an AS into multiple sub-ASs. The IBGP peers in each sub-AS are fully connected. The sub-ASs are connected through EBGP connections, Figure 13 shows a confederation implementation.

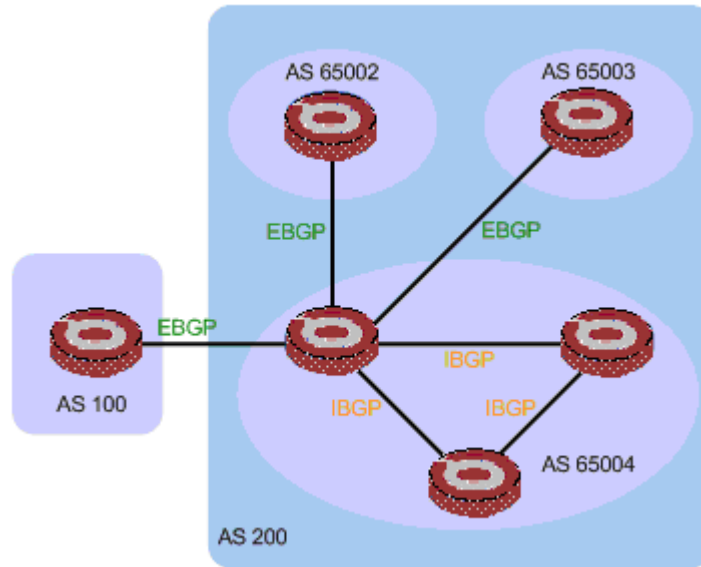


Figure 13 A confederation implementation

To a BGP speaker that does not belong to any confederation, the sub-ASs of a confederation are a whole, and the information about the sub-ASs is invisible to the BGP speaker. The confederation ID, which is usually the corresponding AS number, uniquely identifies a confederation. In Figure 13, AS 200 is a confederation ID.

The disadvantage of confederation is that when an AS changes from non-confederation to confederation, configurations are needed on the routers, and the topology changes.

In a large-scale BGP network, router reflector and confederation can be used simultaneously.

11.1.5 Protocol Standard

Protocol standards concerning BGP are:

- RFC1771: A border gateway protocol 4 (BGP-4)
- RFC2858: Multiprotocol extensions for BGP-4
- RFC3392: Capabilities advertisement with BGP-4
- RFC2918: Route refresh capability for BGP-4
- RFC2439: BGP route flap damping
- RFC1997: BGP communities attribute
- RFC2796: BGP route reflection
- RFC3065: Autonomous system confederations for BGP

Others are still in draft, such as the graceful restart feature and the extended COMMUNITY attribute.

11.2 BGP Configuration Task List

Complete the following tasks to configure BGP:

Task	Remarks
Basic BGP Configuration	Required

Task		Remarks
Configuring the Way to Advertise/Receive Routing Information	Importing Routes	Optional
	Configuring BGP Route	Optional
	Enabling Default Route Advertising	Optional
	Configuring BGP Route Distribution Filtering Policies	Optional
	Configuring BGP Route Reception Filtering Policies	Optional
	Disable BGP-IGP Route Synchronization	Optional
	Configuring BGP Route Dampening	Optional
Configuring BGP Route Attributes		Optional
Tuning and Optimizing a BGP Network		Optional
Configuring a Large-Scale BGP Network	Configuring BGP Peer Group	Required
	Configuring BGP Community	Required
	Configuring BGP RR	Optional
	Configuring BGP Confederation	Optional

11.3 BGP Configuration

BGP Configuration list

BGP Configuration list is as following:

- Enable/disable BGP
- Specify the network route BGP to be notified
- Configure BGP peer
- Configure BGP timer
- Configure local priority
- Configure AS MED
- Compare MED from different AS neighbors
- Configure BGP route aggregation
- Configure route information of IGP protocol introduced by BGP
- Configure BGP route filtration
- Define distribution list
- Define AS path list
- BGP monitor and maintenance

11.3.1 Enable/disable BGP

Specify local AS number when enabling BGP. After enable it, local router monitors and receives BGP request from neighbor routers. How to enable BGP request to neighbor router refers to neighbor Command. All established BGP connection will be cut off when disable it.

Configure it in global configuration mode:

Enable BGP to enter BGP mode

router bgp *as-number*

Disable BGP

no router bgp *as-number*

BGP is defaulted to be disabled.

Example:

!Enable BGP

QTECH(config)#router bgp 400

Specify the network route BGP to be notified

Use network command to specify the network route BGP to be notified and also mask.

Configure it in BGP configuration mode.

Configure the network route local BGP to be notified

network *ip-address* [**mask** *address-mask*]

Cancel the network route local BGP to be notified

no network *ip-address* [**mask** *address-mask*]

network command inserts the route whose destination is ip-address to BGP table and notify this route to peer. Only the route in IP address before configuration can be insert to local BGP table and notify to peer through Update Message.

The distributed route by network command is accurate, that is, the prefix and mask should be totally matched. If the mask is not specified, accurately match it according to natural network interface.

By default, local BGP will not notify any network route.

Example:

```
!Notify 192.168.5.0
```

```
QTECH(config)#network 192.168.5.0
```

11.3.2 Configure BGP peer

The two BGP speakers exchanging BGP packet are peers.

Configure it in BGP configuration mode.

11.3.2.1 Configure AS number

When configuring peers to be neighbors, know the AS number first.

If the AS number specified by neighbor remote-as and router bgp commands are the same, it is the internal neighbor (IBGP), or it is the external (EBGP).

Set up neighborhood and specified AS number

neighbor *neighbor-address* **remote-as** *as-number*

Delete neighborhood

no neighbor *neighbor-address* **remote-as** *as-number*

11.3.2.2 configure permit setting up connection to the indirect-connected peer of EBGP

generally, EBGP and peer should be physically connected, or configure it through following command.

Configure permit setting up connection to the indirect-connected peer of EBGP

neighbor *neighbor-address* **ebgp-multihop**

Configure to set up connection to the direct-connected peer of EBGP

no neighbor *neighbor-address* **ebgp-multihop**

by default, only permit set up connection to the direct-connected peer of EBGP.

11.3.2.3 Configure timer of specified peer

Use neighbor timers command to configure timer for specified BGP peer, including specified Keepalive packet sending interval and keep timer, whose priority is superior than that configured by timers bgp command.

Configure keepalive time interval and keep timer of peer

neighbor *neighbor-address* **timers** *keepalive-interval* *hold-time*

Restore default keepalive time interval and keep timer of peer

no neighbor *neighbor-address* **timers**

By default, keepalive sending interval is 30 seconds and keep timer is 180 seconds.

11.3.2.4 Configure the advertisement-interval

neighbor *neighbor-address* **advertisement-interval** *seconds*

Restore default advertisement-interval

no neighbor *neighbor-address* **advertisement-interval**

By default, advertisement-interval of IBGP is 15 seconds and advertisement-interval of EBGP is 30 seconds.

11.3.2.5 Configure to make its own address to be next-hop

When BGP router distributing route, use its own address to be next hop
Configure to make its own address to be next-hop

neighbor *neighbor-address* **next-hop-self**

Cancel to make its own address to be next-hop

no neighbor *neighbor-address* **next-hop-self**

By default, use default configuration to handle next hop.

11.3.2.6 Configure route filtration strategy based on IP ACL

Configure route filtration strategy based on IP ACL

neighbor *neighbor-address* **distribute-list** *access-list-number* { **in** | **out** }

Cancel route filtration strategy based on IP ACL

no neighbor *neighbor-address* **distribute-list** *access-list-number* { **in** | **out** }

By default, peer will not be route filtrated based on IP ACL

11.3.2.7 Configure route filtration strategy based on AS path-list

Configure route filtration strategy based on AS path-list

neighbor *neighbor-address* **filter-list** *aspath-list-number* { **in** | **out** }

Cancel route filtration strategy based on AS path-list

no neighbor *neighbor-address* **filter-list** *aspath-list-number* { **in** | **out** }

By default, peer will not be route filtrated based on AS path-list

11.3.3 Configure BGP timer

After BGP establishment between peers, they will send Keepalive message to each other to prevent BGP connection breaking up. If router does not receive Keepalive message or other types of packet from other side in the configured Holdtime, local router thought this BGP connection is broken up and it will exit this BGP connection and take corresponding handling to the routes received by this BGP connection. The interval of sending Keepalive message and BGP connection Holdtime are very important for BGP.

When setting up BGP connection between BGP router and its peer, there must be a negotiation. The negotiated hold-time should be the smaller one between this BGP router keepalive-interval and peer hold time. If the negotiated hold-time is 0, keepalive message will not be sent and Holdtime overtime will not be checked.

Configure it in BGP configuration mode.

Configure BGP timer

timers **bgp** *keepalive-interval* *hold-time*

Restore to default timer

no timers **bgp**

The default value of keepalive-interval and holdtime are 30 and 180 seconds.

Sending interval cannot less than 1 second. If configured Holdtime is not 0, it must be more than 3 seconds.

If keepalive interval is more than 1/3 of the negotiated hold-time, the keepalive interval will be auto-set to be the 1/3 of the negotiated hold-time.

11.3.4 Configure local preference

Configuring different local preference can influence the route choosing of BGP. The larger the local preference value is, the more chance for coorsponded route to be chosen.

Configure it in BGP configuration mode.

Configure local preference

bgp default local-preference *localpref*

Restore to default local preference

no bgp default local-preference

The default local preference is 100.

Local preference is sent when IBGP and peer exchanging their Update message.

11.3.5 Configure AS MED

Multi-Exit Discriminator (MED) is the external metric of route which is different from local preference. MED exchanges between AS, but the EMD entered AS will not leave it. MED is used for choosing the best route and the smaller one will be chosen. When a router running BGP which gains the route with the same destination address and different next hop through different External Peer, it will preferentially make a decision according to MED value of different route. The route will smaller MED will be chosen the external route of AS with the same condition.

Configure it in BGP configuration mode.

Configure system MED

default-metric *metric*

Restore to default MED

no default-metric

Default metric value is 0.

The above router only compares the MED from different EBGP peer route in the same AS. Use `bgp always-compare-med` command to compare the peer route in the different AS.

Example:

```
QTECH(config)#default-metric 20
```

11.3.6 Compare MED from different AS neighbors

MED is used for choosing the best path. The smaller MED value will be chosen. It is recommended to use this command when the IGP and route choosing method are the same in the corresponded AS.

Configure it in BGP configuration mode.

Compare MED from different AS neighbors

bgp always-compare-med

Stop comparing MED from different AS neighbors

no bgp always-compare-med

Only compare MED from the same AS route path by default.

Make sure different AS use the same IGP and route choosing method before using this command.

11.3.7 Configure BGP route aggregation

CIDR supports route aggregation. The command is for BGP local route.

Configure it in BGP configuration mode.

Configure local route aggregation

aggregate-address *address mask* [**summary-only**]

Stop local route aggregation

no aggregate-address *address mask* [**summary-only**]

Not to configure any convergent route by default.

11.3.8 Configure route information of IGP protocol introduced by BGP

BGP can send internal network information to other AS. A route protocol can introduce (or learn) route information collected by other route protocol.

Configure it in BGP configuration mode.

Configure BGP to introduce IGP protocol route

redistribute { **connected** | **static** | **rip** | **ospf** } [**metric** *metric*]

Cancel BGP to introduce IGP protocol route

no redistribute { **connected** | **static** | **rip** | **ospf** }

Not introduce route information of other protocol to BGP by default.
It can introduce the route learnt or generated by connected,static,rip,ospf.

11.3.9 Configure BGP distribution list

Use this command to configure BGP route filter list. Use the no command to delete it.

Configure it in global configuration mode:

Define distribute-list

ip distribute-list *list-number* { **permit** | **deny** } *net-addr wildcard-netmask*

Cancel defined distribute-list

no ip distribute-list *list-number* { **permit** | **deny** } *net-addr wildcard-netmask*

None distribute list is configured by default.

BGP route matching is completed by *net-addr* and *wildcard-netmask*. For those successfully matched, determine to accept route or not through deny or permit command. After defining BGP route distribute list, it can realize BGP strategy function by applying neighbor distribute-list command.

If ip distribute-list command is configured, there must be at least one command with ip distribute-list permit for the items with the same list number, or all route will be filtered when using neighbor distribute-list command.

Example:

```
QTECH(config)# ip distribute-list 3 deny 192.168.9.0 0.0.0.255
```

```
QTECH(config)# ip distribute-list 3 permit 0.0.0.0 255.255.255.255
```

11.3.10 Define AS path list

AS path matching is completed by AS path regular expression which matches AS-PATH in BGP route as ASCII string and determines to accept or deny route by deny or permit command for successfully matched AS path. After defining BGP route distribute list, it can realize BGP strategy function by applying neighbor filter-list command.

For the same list number, user can define multiple distribute list, that is, one distribute list number represents a group of distribute list. Each AS path list uses numbers to be their id.

If ip as-path access-list command is configured, there must be at least one command with ip as-path access-list permit for the items with the same list number, or all route will be filtered when using neighbor filter-list command.

Special symbol and its meanings in regular expression are in the following:

Symbol	Description
.	Any single character, including space.
*	Leader character never appears or continually appears many times in target object.
- (hyphen)	Any character in the range formed by the characters before or behind hyphen.
_ (underline)	Match comma, {, }, (,), the beginning of the string, the end of the string or a space
[]	Any character in square brackets.
[^]	Any character except listing in square brackets.(^ is in the front of the character)
	Alternation, matches either left side or right side
^	Match the beginning of the string
\$	Match the end of the string

Configure it in global configuration mode.

Configure an AS path access-list

ip as-path access-list *aspath-list-number* { **permit** | **deny** } *as-regular-expression*

Cancel defined AS path access-list

no ip as-path access-list *aspath-list-number* { **permit** | **deny** } *as-regular-expression*

None control list is configured by default.

In the process of matching, the relationship between *aspath-list-number* is "or", that is, route information matching one item of this list group means it matches the filtration of the distribute list of the as-path list id.

Example:

```
QTECH(config)# ip as-path access-list 10 deny ^700$
QTECH(config)# ip as-path access-list 10 permit .*
```

11.3.11 BGP monitor and maintenance

Use show command in any configuration mode.

Show information in BGP table

show ip bgp [*ip-address* | *A.B.C.D/M*]

```
QTECH# show ip bgp
```

```
Autonomous System number 400, local router ID 192.168.3.3 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 0.0.0.0          100       100         i
```

```
Show AS path list in BGP
```

```
show ip as-path access-list [ aspath-list-number ]
```

```
QTECH# show ip as-path 4
```

```
ip as path access list 4, 1 rule:
```

```
0 permit ^400$
```

```
show BGP peer
```

show ip bgp neighbors [*neighbor-address*]

```
QTECH# show ip bgp neighbors
```

```
BGP Neighbor 192.168.3.3 Status ENABLED remote AS 400, internal link
```

```
Local host 192.168.3.4 Mask 255.255.255.0 AS 400
```

```
Configured Timers: Hold 30 Keepalive 180 Connect Retry 30
```

```
Update 30 Update For IntraAS Route 15
```

```
Param: Local Preference 100 OutBound Metric 0
```

```
Route Reflector Client is DISABLED
```

```
BGP State = Established Socket State = ESTAB Remote Initialized
```

```
Remote Router ID = 192.168.3.3 Connection Up Times 0
```

```
Running Timers: Hold 180 Keepalive 30 Connect Retry DISABLED
```

```
Update 30 Update For IntraAS Route 15
```

```
show BGP peer summary
```

show ip bgp summary

```
QTECH# show ip bgp summary
```

Neighbor	V	AS	MsgRcvd	MsgSent	Up/Down	State/PfxRcd
192.168.3.3	4	400	1	2	04:41:13	Established
192.168.3.7	4	700	2	0	00:44:15	Established
192.168.3.8	4	400	4	1	06:27:29	Established

Chapter 12 Multicast Protocol Configuration

12.1 Overview

12.1.1 Background

Traditional IP communications fall into two modes: point-to-point communications between a source host and a destination host, known as unicast, and point-to-multipoint communications between a source host and all other hosts on the same subnet with the source host, known as broadcast. With broadcast, the information is delivered to all hosts, rather than some specific hosts that need the information, resulting in waste of network bandwidth. In addition, broadcasts are confined only to the local subnet. With unicast, as a separate copy of information is sent to each host, the duplicate IP packets not only use a tremendous amount of network bandwidth but also add to the burden of the source host. Therefore, the conventional unicast and broadcast technologies cannot effectively address the issue of point-to-multipoint data transmission.

Multicast provides a best-effort service to deliver data packets to a specific set of receiver hosts, known as a multicast group, on the network. With multicast, the source host, known as a multicast source, sends only one copy of data packets destined for a multicast group address, and each receiver host of the multicast group can receive the data packets. Only the hosts that have joined the multicast group can receive the traffic addressed to the multicast group, while hosts out of the multicast group cannot.

12.1.2 Benefits

Compared with unicast and broadcast, the multicast technique effectively addresses the issue of point-to-multipoint data transmission. By allowing high-efficiency point-to-multipoint data transmission over an IP network, multicast greatly saves network bandwidth and reduces network load. More importantly, multicast allows convenient deployments of new value-added services in Internet-based information service areas, such as live Webcasting, Web TV, distance learning, telemedicine, Web radio, and real-time videoconferencing.

12.2 Multicast Implementation

Multicast implementation needs to resolve the following issues:

- **Multicast Addressing Mechanism:** As a multicast source sends information to a certain group of receivers, a multicast addressing mechanism is needed to identify multicast groups.
- **Group Membership Management:** As a receiver host needs to join a multicast group before receiving the traffic destined for that group, a group membership management mechanism is needed to allow receiver hosts to join or leave a multicast group dynamically.
- **Multicast Packet Forwarding:** The process a multicast stream is forwarded and delivered to the receiver hosts over the network.
- **Multicast Routing Protocol:** Multicast routing protocols for constituting multicast forwarding trees.

12.2.1 Multicast Addressing Mechanism

12.2.1.1 IP Multicast Addresses

An IP multicast address identifies a specific IP multicast group. IANA has assigned the Class D address space (224.0.0.0 to 239.255.255.255) for multicast.

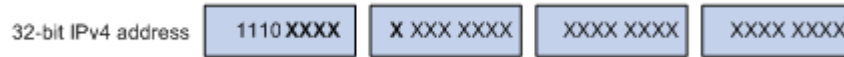


Figure 1 IP multicast address format

As shown in Figure 1, the high-order four bits of a multicast address are 1110. Figure 2 shows the specific address blocks and usages.

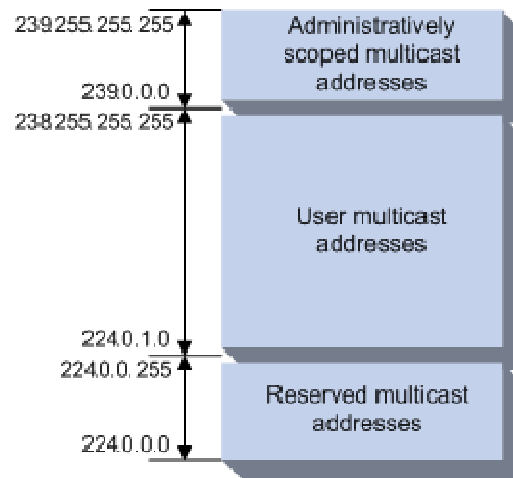


Figure 2 IP multicast address blocks

- 224.0.0.0 to 224.0.0.255 are reserved permanent group addresses. The IP address 224.0.0.0 is reserved and other IP address can be used by routing protocols and for topology searching, protocol maintenance, and so on. Addresses in this range identify local subnets, namely, a packet destined for an address in this block will not be forwarded beyond the local subnet regardless of the Time to Live (TTL) value.
- 224.0.1.0 to 238.255.255.255 are user-available, globally scoped group addresses, among which 232.0.0.0/8 are SSM group addresses while the others are ASM group addresses.
- 239.0.0.0 to 239.255.255.255 are administratively scoped multicast addresses, which are considered to be locally rather than globally unique ASM group addresses. In other words, these addresses can be reused in domains administered by different organizations without causing conflicts.

Note:

Some addresses in the 224.0.1.0/24 segment have also been reserved by IANA for particular multicast applications. For example, 224.0.1.1 has been reserved for the Network Time Protocol (NTP).

12.2.2 Multicast Address Mapping

Note:

As for link-layer multicast, this document focuses on the multicast implementations of the Ethernet protocol only, and the multicast implementations of other link-layer protocols are out of the scope of this document.

IANA assigned MAC addresses 01:00:5E:00:00:00 to 01:00:5E:7F:FF:FF for multicast. This requires the 28-bit IP multicast address space to be mapped to the 23-bit multicast MAC address space. To be specific, the low-order 23 bits of the IP multicast address space are mapped to the low-order 23 bits of the multicast MAC address space, as shown in Figure 3.

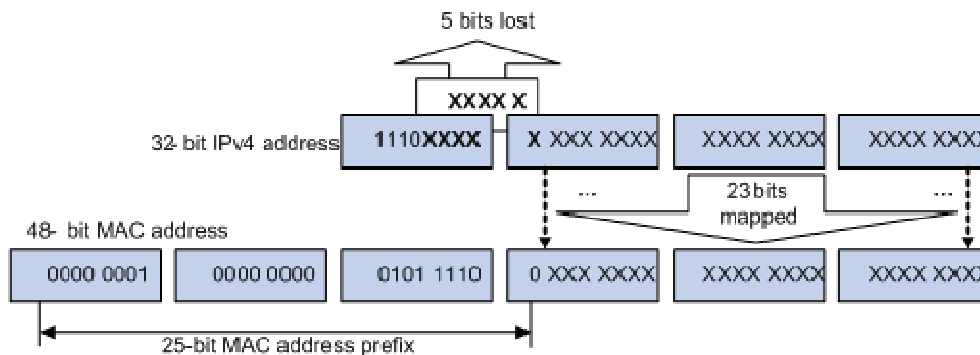


Figure 3 IP-to-MAC address mapping

As a result of this mapping, 32 multicast addresses are mapped to the same MAC address.

12.2.3 Group Membership Management

Group membership management means the establishment and maintenance of multicast group memberships on a multicast router or switch for the subnets directly connected to it, namely, the management of multicast group members attached to the interfaces or ports of the multicast device.

12.2.3.1 IGMP

The Internet Group Management Protocol (IGMP) runs between IP hosts and the immediately connected router. IGMP functions in both directions: A host informs the router of its interest in specific multicast groups through IGMP, and a multicast router uses IGMP to periodically query group memberships on the local subnet to collect and maintain the group memberships. Through IGMP, a router learns whether a specific multicast group has active members on the local subnet, rather than the interrelationships between a multicast group and a host.

So far, there are three IGMP versions:

- IGMPv1 (documented in RFC 1112) defines the group member query and report processes;
- IGMPv2 (documented in RFC 2236) adds a leave-group mechanism based on IGMPv1;

- A major enhancement in IGMPv3 (documented in RFC 3376) is that it allows hosts to specify a list of sources they expect or do not expect multicast data from. The Source-Specific Multicast (SSM) model needs the support of IGMPv3.

This section mainly describes how IGMPv2 works.

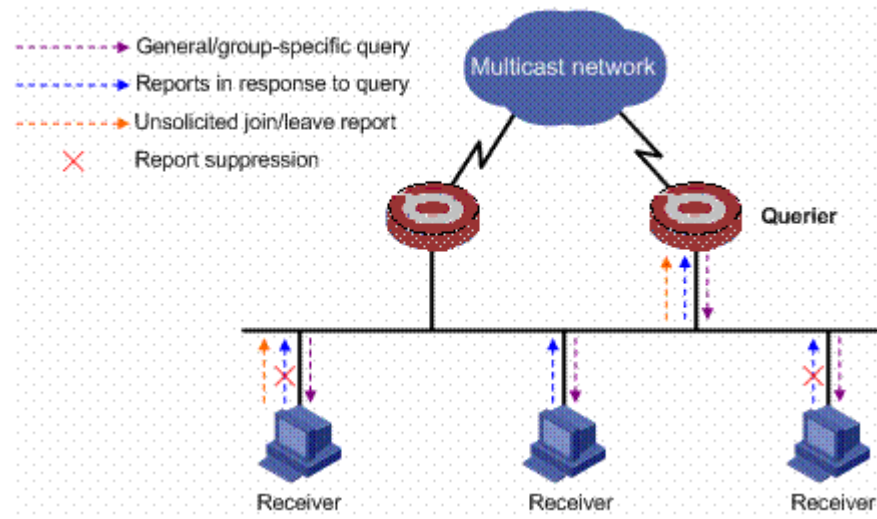


Figure 4 How IGMPv2 works

As shown in Figure 4, when multiple IGMP routers are connected to a subnet, a unique querier needs to be chosen from these routers through the querier election mechanism provided by IGMPv2. A querier periodically sends general query messages (often referred to as general queries) to learn the group memberships and a host responds with an IGMP report. A router also responds to the queries as a receiver host if it has joined a multicast group.

A host sends an unsolicited IGMP report when it needs to join a multicast group, without having to wait for an IGMP query. When a host leaves a multicast group, it sends a leave group message (often referred to as leave message). Upon receiving the leave message, the router sends group-specific queries to that group to determine whether the group still has any active members on the local subnet.

Through the mechanism discussed above, a router creates a table that records the multicast group members on the subnet attached to each of its interfaces. When receiving multicast traffic destined for multicast group G, the router forwards the traffic only to the interfaces with members of multicast group G. The forwarding of multicast traffic between multicast routers is not implemented by IGMP but by the multicast routing protocols.

12.2.3.2 IGMP Snooping

As a protocol designed for IP multicast at the network layer, IGMP maintains only the relationships between Layer 3 interfaces and IP multicast addresses. In most situations, however, multicast traffic inevitably goes through some Layer 2 switches. Without a mechanism to establish mappings between Layer 2 ports and multicast MAC addresses, multicast traffic will be flood to all ports of a switch, and this wastes a great deal of system resource.

IGMP Snooping is used to address this issue. When an IGMP Snooping switch hears an IGMP report message a host sent to the IGMP querier, the switch creates a mapping between the port connected to the host and the multicast MAC address corresponding to the reported multicast group. Then, upon receiving multicast data for that group, the switch forwards the multicast data just to that port based on the created mapping.

12.2.4 Multicast Packet Forwarding

12.2.4.1 Multicast Forwarding Tree

Multicast packets travel along tree-shaped forwarding paths known as multicast forwarding trees over the network to the receivers. Multicast forwarding trees fall into two types: source tree and shared tree.

(1) Source tree

Rooted at the multicast source, a source tree is a forwarding tree with the shortest path from the multicast source to the receivers; therefore, it is also called a shortest path tree (SPT). An SPT needs to be constructed per source per group.

As the shortest forwarding path between a multicast source and the receivers, the source tree minimizes the end-to-end transmission latency. However, this does come at a price. As the router must maintain the routing information for each multicast source, a great deal of system resource is used and the routing table is very large.

(2) Shared tree

Rooted at a router called a rendezvous point (RP), a shared tree is a forwarding tree with the shortest path from the RP to each receiver. It is also called a rendezvous point tree (RPT). There is only one RPT per multicast group on the network. All multicast sources and receivers use the RPT tree for multicast data transmission and reception. The multicast sources send data to the RP and the RP forwards the data down the RPT to all the receivers.

The main advantage of an RPT is that it allows a router to maintain a small number of routing entries. However, as the multicast traffic from a multicast source must pass through the RP before it reaches the receivers, this forwarding tree is not the shortest path from the source to the receivers, and the RP must be highly reliable and powerful.

12.2.4.2 Multicast Packet Forwarding Mechanism

Upon receiving a multicast packet, a router searches its multicast forwarding table according to the destination address and then forwards the packet accordingly. Forwarding a multicast packet is more complex than forwarding a unicast packet. In unicast, a router does not care about the source address; it cares only about the destination address of the packet, based on which the router determines the interface to forward the packet to. In multicast, multicast traffic is destined for to a group of receivers identified by a logical address known as multicast address. Upon receiving a multicast packet, a router checks whether the packet has arrived to the correct incoming interface, namely whether the incoming interface leads to the multicast source, based on the source address before forwarding the packet out the outgoing interface. This process is known as the reverse path forwarding (RPF) check.

The basis for the RPF check is the existing unicast routing table. The router forwards only those packets received on the interface connected to the upstream neighbor on the unicast route to the source. This incoming interface is called RPF interface. The RPF check not only ensures multicast data forwarding along the correct forwarding path but also helps avoid loops.

The RPF check process is as follows: The router searches the unicast routing table for the RPF interface toward the multicast source (when an SPT is used) or the RP (when an RPT is used). If the packet is received on the RPF interface, it passes the RPF check and then forwarded to downstream node; otherwise, the packet is discarded.

12.2.5 Multicast Routing Protocols

Similar to unicast protocols, multicast routing protocols fall into intra-domain and inter-domain protocols:

- Based on the group memberships maintained by IGMP, an intra-domain multicast routing protocol builds multicast distribution trees according to certain

- multicast routing algorithms and creates multicast routing state entries on multicast routers, which forward multicast traffic as per the routing state entries.
- Based on the inter-domain multicast routing policy configured in the network, an inter-domain multicast routing protocol propagates multicast source information and exchanges multicast routing information among autonomous systems (ASs), thus ensuing multicast forwarding among different domains.

12.2.5.1 Intra-Domain Multicast Routing Protocols

Among a variety of intra-domain multicast routing protocols, Protocol Independent Multicast (PIM) is a popular one. Based on the forwarding mechanism, PIM falls into two modes – dense mode (referred to as PIM-DM) and sparse mode (referred to as PIM-SM).

(1) PIM-DM

In a PIM-DM domain, routers use periodical PIM Hello messages for PIM neighbor discovery, determination of leaf routers and leaf networks, and designated router (DR) election on a multi-access network. Although PIM-DM does not require a DR, one must be elected among multiple routers on a multi-access network running IGMPv1 in a PIM-DM domain to act as the IGMPv1 querier on that multi-access network.

As a dense mode multicast routing protocol, PIM-DM uses the “push” mode for multicast forwarding, and is suitable for small-sized networks with densely distributed multicast members. PIM-DM works as follows:

- PIM-DM assumes that at least one multicast group member exists on each subnet of a network, and therefore multicast data is flooded to all nodes on the network. Then, branches without receivers downstream are pruned from the forwarding tree, leaving only those branches with receivers. This “flood and prune” process takes place periodically, that is, pruned branches resume multicast forwarding when the pruned state times out and then data is re-flooded down these branches, and then are pruned again.
- When a host attached to a pruned node joins the multicast group, the node sends a graft message toward the upstream node. Then the node resumes multicast traffic forwarding.

(2) PIM-SM

In a PIM-SM domain, routers periodically sends PIM Hello messages for PIM neighbor discovery and DR election on a multi-access network, where the DR sends join/prune messages toward the root of the multicast forwarding tree for the receiver host attached to it, or forwards multicast traffic from the directly connected multicast source onto the multicast distribution tree.

As a sparse mode multicast routing protocol, PIM-SM uses the “pull” mode for multicast forwarding, and is suitable for large- and medium-sized networks with sparsely and widely distributed multicast members. The basic implementation of PIM-SM is as follows:

- PIM-SM assumes that no hosts need to receive multicast data. In the PIM-SM mode, and delivers multicast data only to those hosts that have explicitly requested for the data. The core task for PIM-SM in multicast forwarding is to build and maintain rendezvous point trees (RPTs). An RPT is rooted at a router in the PIM domain as the common node, referred to as the rendezvous point (RP), through which the multicast data travels down the RPT to the receivers.
- When a receiver is interested in the multicast data addressed to a specific multicast group, the router connected to this receiver sends a join message to the RP for that multicast group. The path along which the message goes hop by hop to the RP forms a branch of the RPT.

- When a multicast source sends multicast data to a multicast group, the source-side DR first registers the multicast source with the RP by sending register messages to the RP by unicast. The arrival of a register message at the RP triggers the establishment of an SPT. Then, the multicast source sends subsequent multicast packets along the SPT to the RP. Upon reaching the RP, the multicast packet is duplicated and delivered to the receivers down the RPT.

12.2.5.2 Inter-Domain Multicast Routing Protocols

Inter-domain multicast routing protocols are used to propagate multicast information among different ASs. So far, mature solutions include:

- Multicast Border Gateway Protocol (MBGP) is used for exchanging multicast routing information between ASs.
- Multicast Source Discovery Protocol (MSDP) is used for advertising multicast source information between Internet service providers (ISPs).

(1) MBGP

The exchange of routing information (reachability information) between different ASs is the first-line issue to address for inter-domain communications. Because different ASs may be operated by different service providers, unlike intra-domain routing information, inter-domain routing information needs to contain not only the distance information but also the service provider policies.

The multicast topology may be different from the unicast topology due to both physical and policy-related reasons. Some routers on the network may support only unicast, while some others, though multicast capable, may be configured not to forward multicast packets. In order to construct inter-domain multicast forwarding trees, in addition to unicast routing information, the multicast topology information is also needed. In short, an inter-domain multicast routing protocol needs to meet the following requirements:

- Able to differentiate the unicast topology and the multicast topology.
- Having a set of stable methods for peering and policy control.

As the most popular inter-domain unicast routing protocol so far, the Border Gateway Protocol version 4 (BGP-4) already satisfies the latter requirement and is proven to be effective and stable. Therefore, a reasonable solution for inter-domain propagation of multicast routing information is to enhance and extend BGP-4 rather than to devise a set of entirely new protocols. Multiprotocol extensions for BGP are defined in RFC 2858. The extended BGP (known as MP-BGP or BGP-4+) can carry not only IPv4 unicast routing information but also the routing information for other network layer protocols (such as multicast and IPv6). The capability of carrying multicast routing information is only one of the functions of these extensions. The multicast extension of BGP is referred to as multicast BGP (MBGP).

With MBGP, both unicast routing information and multicast routing information can be exchanged in the same process, but are stored in different routing tables. MBGP is an enhanced version of BGP-4; therefore, all the common policies and configuration methods supported by BGP-4 can be applied to multicast.

(2) MSDP

In the basic PIM-SM mode, a multicast source registers only with the RP in the local PIM-SM domain, and the multicast source information of a domain is confined within the domain. As a result, the RP is aware of the source information only within the local domain and multicast distribution trees are built only within the local domain to deliver multicast data from a local multicast source to local receivers. A service provider does not want to rely on other ISPs' RP routers to forward multicast traffic to its own customers, but it does want to be able to get information from multicast sources, wherever the RPs for these sources are, and send the information to its customers.

The Multicast Source Discovery Protocol (MSDP) is an inter-domain multicast solution developed to discover multicast sources in other PIM-SM domains thus to allow interconnection among these domains. With

MSDP, an RP in one domain can establish peering relationships with RPs in other domains. Based on these peering relationships, RPs in different domains are interconnected and multicast source information is exchanged between the RPs.

In addition to inter-domain propagation of multicast source information, MSDP has a special application for PIM-SM: anycast RP. Anycast RP refers to such an application that implements load balancing and redundancy backup between two or more RPs within a PIM-SM domain by configuring the same IP address for, and establishing MSDP peering relationships between, these RPs.

12.2.6 Multicast Models

Based on how the receivers treat the multicast sources, there are two multicast models:

- **ASM model:** Any-source multicast model. In the ASM model, any sender can be a multicast source sending multicast information to a multicast group, and receivers can join a multicast group identified by a group address and obtain multicast information addressed to that multicast group. In this model, receivers are not aware of the location of the multicast source in advance. However, they can join or leave the multicast group at any time.
- **SSM model:** Source-specific multicast model. In actual situations, users may be interested in the multicast data only from certain specific multicast sources. The SSM model provides a transmission service that allows users to specify the multicast sources they are interested in at the client side.

The multicast routing protocols mentioned in the section above are mainly for the ASM model. In ASM, receivers cannot specify the multicast sources they are interested in; instead, they passively receive multicast streams from all multicast sources. Unlike the ASM model, The SSM model allows hosts to specify the multicast sources.

In the SSM model, the multicast address range is different from that in the ASM model and dedicated multicast forwarding paths between receivers and the specified multicast sources are established through PIM-SM. In SSM, receivers know exactly where a multicast source is located by means of advertisements, consultancy, and so on. Therefore, no RP is needed, no RPT is required, there is no source registration process, and there is no need of using MSDP for discovering sources in other PIM-SM domains. Moreover, routers with receivers on the subnet can learn the multicast source information specified by the receivers when joining a multicast group in the following two ways:

- With IGMPv3 running on the receivers, multicast source addresses are contained in IGMPv3 report messages.
- With IGMPv1 or IGMPv2 running on the receivers, no source addresses are specified in IGMPv1 or IGMPv2 report messages. In such cases, static SSM mappings must be configured on the router to map the (*, G) information carried in these reports to the (G, INCLUDE, (S1, S2...)) information.

12.3 GMRP Overview

GMRP (GARP Multicast Registration Protocol), based on GARP, is used for maintaining multicast registration information of the switch. All GMRP-capable switches can receive multicast registration information from other switches, dynamically update local multicast registration information, and send their own local multicast registration information to other switches. This information switching mechanism keeps consistency of the multicast information maintained by every GMRP-supporting device in the same switching network.

A host sends a GMRP Join message, if it is interested in joining a multicast group. After receiving the message,

the switch adds the port on which the message was received to the multicast group, and broadcasts the message throughout the VLAN where the receiving port resides. In this way, the multicast source in the VLAN gets aware of the existence of the multicast group member. When the multicast source sends multicast packets to a group, the switch only forwards the packets to ports connected to the members of that group, thereby implementing Layer 2 multicast in the VLAN.

12.3.1 GMRP Configuration

12.3.1.1 GMRP Configuration list

In all configurations, enable global GMRP first before enable GMRP on a port. GMRP Configuration list is as following :

- Enable/disable global GMRP
- Enable/disable GMRP on a port
- Display GMRP
- Add/delete multicast that can be dynamic learnt by GMRP
- Display multicast that can be learnt by GMRP

12.3.1.2 Enable/disable global GMRP

Please configure it in global configuration mode :

Enable global GMRP

gmrp

Disable global GMRP

no gmrp

By default, GMRP globally disables

For example :

! Enable GMRP globally

QTECH(config)#gmrp

12.3.1.3 Enable/disable GMRP on a port

Enable global GMRP before enable GMRP on a port. Please configure it in interface configuration mode :

Enable GMRP on a port

gmrp

Disable GMRP on a port

no gmrp

For example :

! Enable GMRP on Ethernet port 3

QTECH(config-if-ethernet-0/0/3)#gmrp



Caution : Enable global GMRP before enable GMRP on a port. By default, global GMRP deisables and GMRP on a port can be enabled in trunk mode interface.

12.3.1.4 Display GMRP

Use following command in any configuration mode to display global GMRP :

show gmrp

Use following command in any configuration mode to display GMRP on a port :

show gmrp interface [*interface-list*]

Interface-list keyword is optional. If this keyword unspecified, the command displays GMRP information for all the Ethernet ports. If specified, the command displays GMRP information on specified Ethernet port.

For example :

! Display GMRP information of Ethernet 0/0/2 to ethernet 0/0/4 ethernet 0/1/2

```
QTECH(config)#show gmrp interface ethernet 0/0/2 to ethernet 0/0/4 ethernet 0/1/2
port GMRP status
e0/0/2 enable
e0/0/3 enable
e0/0/4 enable
e0/1/2 enable
Total entries : 4.
```

12.3.1.5 Add/delete multicast that can be dynamic learnt by GMRP

Add configured static multicast group to GMRP for other switch learning it.

garp permit multicast [**mac-address** *mac* **vlan** *vlan-id*]

Example :

Add multicast group 01 : 00 : 5e : 00 : 01 : 01 vlan 1 to GMRP

```
QTECH(config)#garp permit multicast mac-address 01 : 00 : 5e : 00 : 01 : 01 vlan 1
```

12.3.1.6 Display multicast that can be learnt by GMRP

Display multicast group can be statically learnt by GMRP.

show garp permit multicast

For example : Display multicast group that can be statically learnt by GMRP

```
QTECH(config)#show garp permit multicast
```

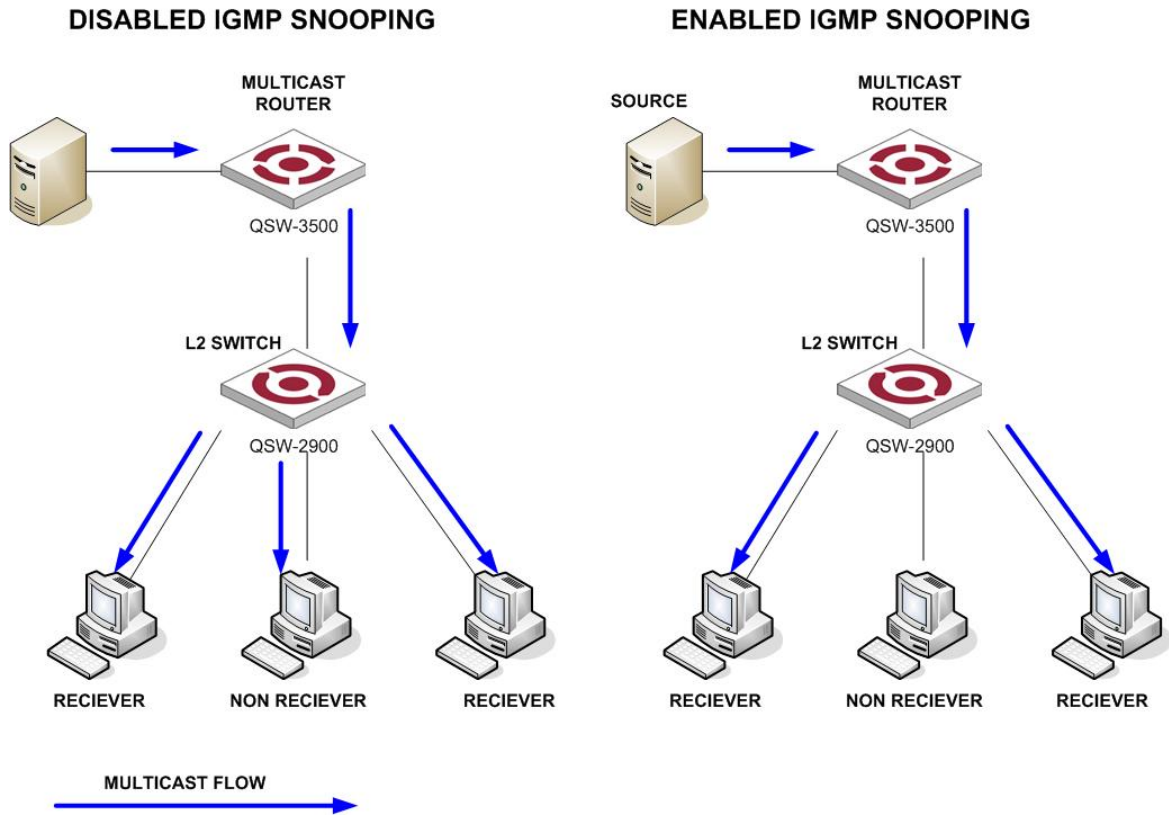
12.4 IGMP Snooping Overview

12.4.1 IGMP Snooping

Internet Group Management Protocol Snooping (IGMP Snooping) is a multicast constraining mechanism that runs on Layer 2 devices to manage and control multicast groups.

By listening to and analyzing IGMP messages, a Layer 2 device running IGMP Snooping establishes mappings between ports and multicast MAC addresses and forwards multicast data based on these mappings.

As shown in Figure below, when IGMP Snooping is not running on the switch, multicast packets are broadcast to all devices at Layer 2. When IGMP Snooping is running on the switch, multicast packets for known multicast groups are multicast to the receivers, rather than broadcast to all hosts, at Layer 2.



Before and after IGMP Snooping is enabled on the Layer 2 device

12.4.2 Basic Concepts in IGMP Snooping

12.4.2.1 IGMP Snooping related ports

As shown in Figure 2, Router A connects to the multicast source, IGMP Snooping runs on Switch A and Switch B, Host A and Host C are receiver hosts (namely, multicast group members).

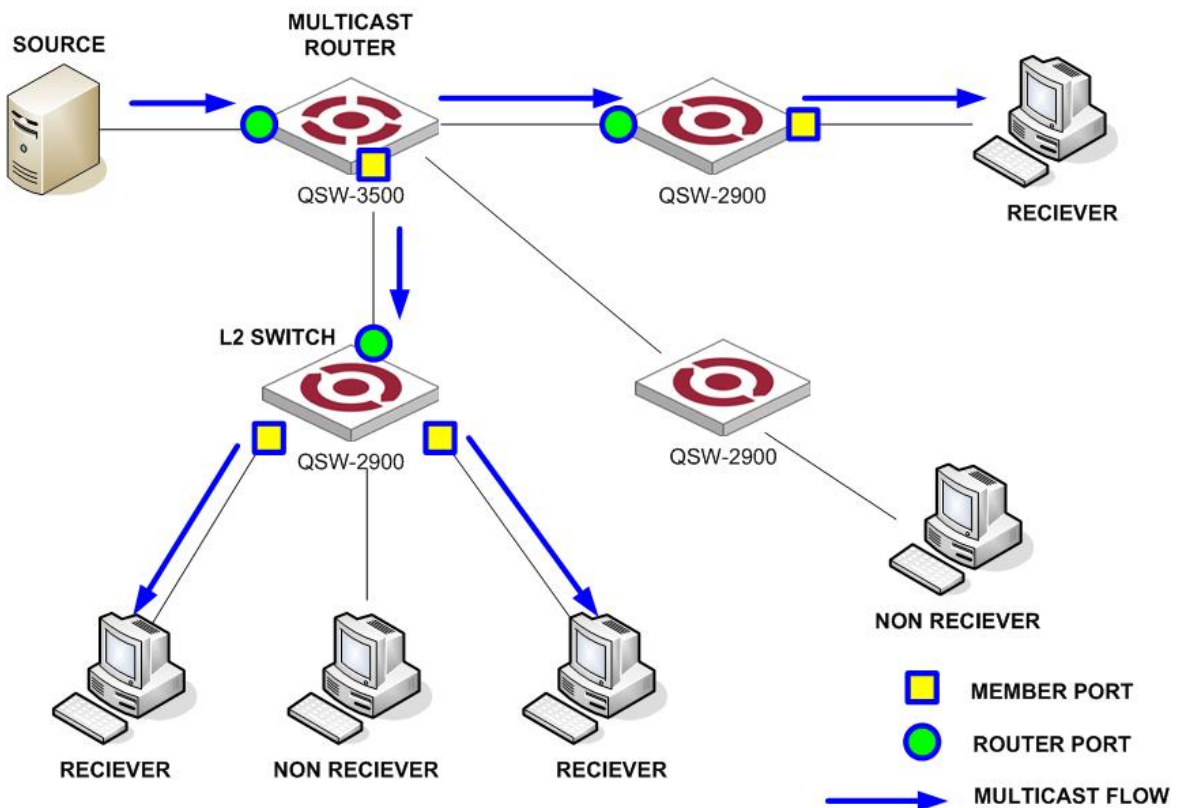


Figure 2 IGMP Snooping related ports

Ports involved in IGMP Snooping, as shown in Figure 2, are described as follows :

- Router port : A router port is a port on the Ethernet switch that leads switch towards the Layer 3 multicast device (DR or IGMP querier). In the figure, Ethernet 1/0 of Switch A and Ethernet 1/0 of Switch B are router ports. The switch registers all its local router ports (including static and dynamic router ports) in its router port list.
- Member port : A member port is a port on the Ethernet switch that leads switch towards multicast group members. In the figure, Ethernet 0/0/1/1 and Ethernet 1/2 of Switch A and Ethernet 0/0/1/1 of Switch B are member ports. The switch registers all the member ports (including static and dynamic member ports) on the local device in its IGMP Snooping forwarding table.

 Note :

- Whenever mentioned in this document, a router port is a port on the switch that leads the switch to a Layer 3 multicast device, rather than a port on a router.
- An IGMP-snooping-enabled switch deems that all its ports on which IGMP general queries with the source address other than 0.0.0.0 or PIM hello messages are received to be router ports.

12.4.2.2 Port aging timers in IGMP Snooping and related messages and actions

Table 1 Port aging timers in IGMP Snooping and related messages and actions

Timer	Description	Message before expiry	Action after expiry
Router port aging timer	For each router port, the switch sets a timer initialized to the aging time of the route port.	IGMP general query of which the source address is not 0.0.0.0 or PIM hello	The switch removes this port from its router port list.
Member port aging timer	When a port joins a multicast group, the switch sets a timer for the port, which is initialized to the member port aging time.	IGMP membership report	The switch removes this port from the multicast group forwarding table.

 Note :

The port aging mechanism of IGMP Snooping works only for dynamic ports; a static port will never age out.

12.4.3 How IGMP Snooping Works

A switch running IGMP Snooping performs different actions when it receives different IGMP messages, as follows :

12.4.3.1 When receiving a general query

The IGMP querier periodically sends IGMP general queries to all hosts and routers (224.0.0.1) on the local subnet to find out whether active multicast group members exist on the subnet.

Upon receiving an IGMP general query, the switch forwards it through all ports in the VLAN except the receiving port and performs the following to the receiving port :

- If the receiving port is a router port existing in its router port list, the switch resets the aging timer of this router port.
- If the receiving port is not a router port existing in its router port list, the switch adds it into

its router port list and sets an aging timer for this router port.

12.4.3.2 When receiving a membership report

A host sends an IGMP report to the multicast router in the following circumstances :

- Upon receiving an IGMP query, a multicast group member host responds with an IGMP report.
- When intended to join a multicast group, a host sends an IGMP report to the multicast router to announce that it is interested in the multicast information addressed to that group.

Upon receiving an IGMP report, the switch forwards it through all the router ports in the VLAN, resolves the address of the reported multicast group, and performs the following :

- If no forwarding table entry exists for the reported group, the switch creates an entry, adds the port as member port to the outgoing port list, and starts a member port aging timer for that port.
- If a forwarding table entry exists for the reported group, but the port is not included in the outgoing port list for that group, the switch adds the port as a member port to the outgoing port list, and starts a member port aging timer for that port.
- If a forwarding table entry exists for the reported group and the port is included in the outgoing port list, which means that this port is already a member port, the switch resets the member port aging timer for that port.

Note :

A switch does not forward an IGMP report through a non-router port. The reason is as follows : Due to the IGMP report suppression mechanism, if the switch forwards a report message through a member port, all the attached hosts listening to the reported multicast address will suppress their own reports upon hearing this report, and this will prevent the switch from knowing whether any hosts attached to that port are still active members of the reported multicast group.

12.4.3.3 When receiving a leave group message

When an IGMPv1 host leaves a multicast group, the host does not send an IGMP leave group message, so the switch cannot know immediately that the host has left the multicast group. However, as the host stops sending IGMP reports as soon as it leaves a multicast group, the switch deletes the forwarding entry for the member port corresponding to the host from the forwarding table when its aging timer expires.

When an IGMPv2 or IGMPv3 host leaves a multicast group, the host sends an IGMP leave group message to the multicast router.

When the switch hears a group-specific IGMP leave group message on a member port, it first checks whether a forwarding table entry for that group exists, and, if one exists, whether its outgoing port list contains that port.

- If the forwarding table entry does not exist or if its outgoing port list does not contain the port, the switch discards the IGMP leave group message instead of forwarding it to any port.
- If the forwarding table entry exists and its outgoing port list contains the port, the switch forwards the leave group message to all router ports in the VLAN. Because the switch does not know whether any other hosts attached to the port are still listening to that group address, the switch does not immediately remove the port from the outgoing port list of the forwarding table entry for that group; instead, it resets the member port aging timer for the port.

Upon receiving the IGMP leave group message from a host, the IGMP querier resolves from the message the address of the multicast group that the host just left and sends an IGMP group-specific query to that multicast group through the port that received the leave group message. Upon hearing the IGMP group-specific query, the switch forwards it through all its router ports in the VLAN and all member ports for that multicast group, and performs the following :

- If any IGMP report in response to the group-specific query is heard on a member port

before its aging timer expires, this means that some host attached to the port is receiving or expecting to receive multicast data for that multicast group. The switch resets the aging timer of the member port.

- If no IGMP report in response to the group-specific query is heard on a member port before its aging timer expires, this means that no hosts attached to the port are still listening to that group address : the switch removes the port from the outgoing port list of the forwarding table entry for that multicast group when the aging timer expires.

12.4.4 Processing of Multicast Protocol Messages

With Layer 3 multicast routing enabled, an IGMP Snooping switch processes multicast protocol messages differently under different conditions, specifically as follows :

1) If only IGMP is enabled, or both IGMP and PIM are enabled on the switch, the switch handles multicast protocol messages in the normal way.

2) In only PIM is enabled on the switch :

- The switch broadcasts IGMP messages as unknown messages in the VLAN.
- Upon receiving a PIM hello message, the switch will maintain the corresponding router port.

3) When IGMP is disabled on the switch, or when IGMP forwarding entries are cleared (by using the reset igmp group command) :

- If PIM is disabled, the switch clears all its Layer 2 multicast entries and router ports.
- If PIM is enabled, the switch clears only its Layer 2 multicast entries without deleting its router ports.

4) When PIM is disabled on the switch :

- If IGMP is disabled, the switch clears all its router ports.
- If IGMP is enabled, the switch maintains all its Layer 2 multicast entries and router ports.

Table 2-3 IGMP Snooping messages

Message	Sender	Receiver	Purpose	Switch action		
IGMP general query message	Multicast router and multicast switch	Multicast member switch and host	Query if the multicast groups contain any member	Check if the message comes from the original router port	<p>If yes, reset the aging timer of the router port</p> <p>If not, notify the multicast router that a member is in a multicast group and start the aging timer for the router port</p>	
IGMP group-specific query message	Multicast router and multicast switch	Multicast member switch and host	Query if a specific IGMP multicast group contains any member	Send an IGMP group-specific query message to the IP multicast group being queried.		
IGMP host report message	Host	Multicast router and multicast switch	Apply for joining a multicast group, or respond to an IGMP query message	Check if the IP multicast group has a corresponding MAC multicast group	<p>If yes, check if the port exists in the MAC multicast group</p> <p>If not, add the port to the MAC multicast group, reset the aging timer of the port and check if the corresponding IP multicast group exists.</p>	<p>If yes, add the IP multicast group address to the MAC multicast group table.</p> <p>If yes, add the port to the IP multicast group.</p> <p>If not, create an IP multicast group and add the port to it.</p>
				<p>If not :</p> <p>Create a MAC multicast group and notify the multicast router that a member is ready to join the multicast group.</p> <p>Add the port to the MAC multicast group and start the aging timer of the port.</p> <p>Add all ports in the VLAN owning this port to the forward port list of the MAC multicast group.</p> <p>Add the port to the IP multicast group.</p>		
IGMP leave message	Host	Multicast router and multicast switch	Notify the multicast router and multicast switch that the host is leaving its multicast group.	Multicast router and multicast switch send IGMP group-specific query packet(s) to the multicast group whose member host sends leave packets to check if the multicast group has any members and enable the corresponding query timer.	<p>If no response is received from the port before the timer times out, the switch will check whether the port corresponds to a single MAC multicast group.</p> <p>If yes, remove the corresponding MAC multicast group and IP multicast group</p> <p>If no, remove only those entries that correspond to this port in the MAC multicast group, and remove the corresponding IP multicast group entries</p> <p>If no response is received from the multicast group before the timer times out, notify the router to remove this multicast group node from the multicast tree</p>	



Caution :

An IGMP-Snooping-enabled Ethernet switch judges whether the multicast group exists when it receives an IGMP leave packet sent by a host in a multicast group. If this multicast group does not exist, the switch will drop the IGMP leave packet instead of forwarding it.

12.4.5 Protocols and Standards

IGMP Snooping is documented in :

RFC 4541 : Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches

12.5 MLD Snooping Configuration

12.5.1 MLD Snooping protocol overview

MLD (Multicast Listener Discovery) is a part of IPv6 protocol which supports and manages IP multicast between host and multicast route. IP multicast permits transferring IP packets to a host clump which constructing a multicast group. Multicast group members are dynamic. Host can dynamically add or leave the group to reduce the network loading.

MLD Snooping is for detecting MLD packet between host and router. It can dynamically create, maintain and delete multicast address table according to the adding and leaving of the group members. Multicast packet is transferred according to their own multicast address.

12.5.2 MLD Snooping Configuration

Use following commands to enable/disable MLD Snooping create mac address multicast transferring table in L2.

Configure it in global configuration mode :

Enable MLD Snooping

mld-snooping

Disable MLD Snooping

no mld-snooping

By default, MLD Snooping is disabled.

Display MLD Snooping status

show mld-snooping

For example :

! Display MLD Snooping status

QTECH(config)#show mld-snooping

Enable/disable MLD Snooping of some VLAN

In VLAN mode, use following commands :

Enable MLD Snooping under VLAN

mld-snooping

Disable MLD Snooping under VLAN

no mld-snooping

By default, MLD Snooping under vlan is disabled.

12.5.3 MLD Snooping host aging time

Use following command to configure MLD-snooping host aging time in global configuration mode :

mld-snooping host-aging-time

Use following command to show mld-snooping host aging time :

```
show mld-snooping
```

For example :

```
! Configure mld snooping host aging time to be 10s
QTECH(config)#mld-snooping host-aging-time 10
```

12.5.4 MLD Snooping Max response time

Configure max response time when receiving leave packet in global configuration mode :

mld-snooping max-response-time *seconds*

For example :

```
! Configure MLD-Snooping max response time to be 13s
QTECH(config)#mld-snooping max-response-time 13
```

12.5.5 MLD Snooping fast leave

This function is for controlling aging port after receiving leave packet in interface mode :

mld-snooping fast-leave

For example :

```
! Enable mld-snooping fast-leave
QTECH(config-if-ethernet-0/1)#mld-snooping fast-leave
```

12.5.6 MLD Snooping max learnt multicast number

Configure max learnt multicast number in global configuration mode :

mld-snooping group-limit *limit*

For example :

```
! Configure max learnt multicast number to be 10
QTECH(config-if-ethernet-0/1)#mld-snooping group-limit 10
```

12.5.7 MLD Snooping permit/deny group

Configure mld-snooping permit/deny group and learning rules.

In interface mode :

mld-snooping permit/deny group *group-address*

In global configuration mode :

mld-snooping deny/permit group all

For example :

```
! Configure port deny group 33:33:00:00:01:01
QTECH(config-if-ethernet-0/1)#mld-snooping deny group 33:33:00:00:01:01
! Configure learning group all
QTECH(config)#mld-snooping permit group all
```

12.5.8 Configure MLD Snooping route-port forward

The port receiving MLD query packet is called multicast route port.

Configure add route-port to MLD Snooping learnt dynamic multicast in global configuration mode. By default, it is not added.

mld-snooping route-port forward
no mld-snooping route-port forward

For example :

```
! Add route port to MLD Snooping learnt dynamic multicast
QTECH(config)#mld-snooping route-port forward
```

12.5.9 MLD Snooping multicast VLAN

Specify a vlan for a port. All MLD packets detected by MLD Snooping will be considered from this vlan. The vlan ID in MLD packet will be ignored.

Multicast VLAN will be effective after creation. Configure it in interface configuration mode :

mld-snooping multicast vlan *vlan-id*
no mld-snooping multicast vlan

For example :

```
! Configure multicast vlan of e0/1 to be vlan 2
QTECH(config-if-ethernet-0/1)#mld-snooping multicast vlan 2
```

12.5.10 Display MLD Snooping group

Show MLD Snooping group in any mode

show mld-snooping group

For example:

Show MLD Snooping group

QTECH(config)#show mld-snooping group

12.6 IGMP Snooping configuration

Use following command to control IGMP Snooping to establish the MAC address multicast transmission table in layer 2.

Use following command in global configuration mode :

Enable IGMP Snooping

igmp-snooping

Disable IGMP Snooping

no igmp-snooping

By default, IGMP Snooping disables.

Display IGMP Snooping

Use following command in any mode to see IGMP Snooping :

For example :

! Display IGMP snooping information

QTECH(config)#show igmp-snooping

12.6.1 IGMP Snooping multicast interface aging time configuration

Use following command in global configuration mode to configure host-aging-time dynamic multicast group learnt by igmp-snooping :

igmp-snooping host-aging-time

Use following command to display host-aging-time dynamic multicast group learnt by igmp-snooping :

show igmp-snooping

For example :

! Configure host-aging-time of the dynamic multicast group learnt by igmp-snooping to be 10 seconds

QTECH(config)#igmp-snooping host-aging-time 10

12.6.2 IGMP Snooping max-response-time configuration

Configure the max response time to delete group interface when receiving a leave packet :

igmp-snooping max-response-time *seconds*

Use this command in global configuration mode.

For example :

! Configure the max-response-time of igmp-snooping is 13 seconds

```
QTECH(config)#igmp-snooping max-response-time 13
```

12.6.3 IGMP Snooping interface fast-leave configuration

Configure interface fast-leave when fast-leave enables, if the fast-leave packet is received, the interface leaves the aging group, or the time to leave is determined by the max-response-time :

igmp-snooping fast-leave

Use this command in interface configuration mode.

For example :

! Enable igmp-snooping fast-leave

```
QTECH(config-if-ethernet-0/0/1)#igmp-snooping fast-leave
```

12.6.4 Configure the number of the multicast group allowed learning

Use igmp-snooping group-limit command to configure the number of the multicast group allowed learning.

igmp-snooping group-limit *limit*

Use this command in global configuration mode.

For example :

! Configure the igmp-snooping group-limit to be 10

```
QTECH(config-if-ethernet-0/0/1)#igmp-snooping group-limit 10
```

12.6.5 IGMP Snooping permit/deny group configuration

Configure igmp-snooping permit/deny group and default group learning regulation.

Configure igmp-snooping permit/deny group in interface configuration mode :

igmp-snooping permit/deny group *group-address*

Configure igmp-snooping default group learning regulation in global configuration mode :

igmp-snooping deny/permit group *all*

For example :

! Configure Ethernet 0/0/1 not to learn multicast 01 : 00 : 5e : 00 : 01 : 01

```
QTECH(config-if-ethernet-0/0/1)#igmp-snooping deny group 01 : 00 : 5e : 00 : 01 : 01
```

! Configure the learning regulation of default group to allow all multicast group

```
QTECH(config)#igmp-snooping permit group all
```

12.6.6 IGMP Snooping route-port forward configuration

Multicast routers interface is the interface received IGMP inquiring packet (It is also called mix router interface.).

Use `igmp-snooping route-port forward` command to configure whether to add router interface to IGMP snooping learning group. By default, router interface to IGMP snooping learning group is not added.

Use following command in global configuration mode :

igmp-snooping route-port forward

no igmp-snooping route-port forward

For example :

! Enable igmp-snooping route-port forward

```
QTECH(config)#igmp-snooping route-port forward
```

12.6.7 Enable/disable IGMP Snooping querier

To set up multicast route table, send IGMP query packet. The unit to send the packet is called querier.

Enable or disable querier sending IGMP query packet. It is defaulted not to send.

Configure it in global configuration mode :

igmp-snooping querier

no igmp-snooping querier

Example :

! Enable igmp-snooping querier

```
QTECH(config)# igmp-snooping querier
```

12.6.8 Configure IGMP Snooping query-interval

Configure interval of sending IGMP query. It is defaulted to be 60s.

Configure it in global configuration mode :

igmp-snooping query-interval *seconds*

no igmp-snooping query-interval

Example :

! Configure interval of sending IGMP query to be 90s

```
QTECH(config)# igmp-snooping querier 90
```

12.6.9 Configure IGMP Snooping querier vlan

Sending IGMP query must specify vlan. Packet will be transferred to all ports of this vlan.

Configure vlan which IGMP query sent by querier to be sent to. It is defaulted to be vlan 1

Configure it in global configuration mode :

igmp-snooping querier-vlan *vlanID*

no igmp-snooping querier-vlan

Example :

! Configure querier sending query to vlan 10

```
QTECH(config)# igmp-snooping querier-vlan 10
```

12.6.10 Configure IGMP Snooping query max response

Configure the max response after receiving query, that is the response value in IGMP query. It is defaulted to be 10s.

Configure it in global configuration mode :

igmp-snooping query-max-respon *second*

no igmp-snooping query-max-respon

Example :

! Configure the max response after receiving query to be 15s

```
QTECH(config)# igmp-snooping query-max-respon 15
```

12.6.11 Configure IGMP Snooping query source IP

Configure IGMP query source IP to demonstrate the destination IP to response to. It is defaulted to be 0.0.0.0

Configure it in global configuration mode :

igmp-snooping general-query source-ip *ipaddress*

no igmp-snooping general-query source-ip

Example :

! Configure IGMP query source IP to be 1.1.1.111

```
QTECH(config)# igmp-snooping general-query source-ip 1.1.1.111
```

12.6.12 Configure IGMP Snooping route port aging

The port receiving IGMP query is called multicast route port.

Configure the aging of route port. It is defaulted to be aging.

Configure it in global configuration mode :

no igmp-snooping router-port-age

igmp-snooping router-port-age

Example :

Configure the route port aging

no igmp-snooping router-port-age

12.6.13 Add IGMP Snooping route port

Added route port demonstrates the transferred port of leave or report packet of the host in the same multicast.

Configure uplink route port of host responding packet.

Configure it in global configuration mode :

igmp-snooping route-port vlan *vlanID* **interface** *port-number*

no igmp-snooping route-port vlan *vlanID* interface *port-number*

Example :

Configure e0/0/1 of vlan 2 to be route port of current group(determined by source IP of querie)

```
QTECH(config)# igmp-snooping route-port vlan 2 interface ethernet 0/0/1
```

12.7 Static Multicast Configuration

12.7.1 Brief introduction of Static Multicast

Static multicast configuration command is used to create multicast group and add interfaces to it. If the switch supports multicast, when receiving multicast packet, detect whether there is multicast group. If it doesn't exist, transfer the multicast packet as broadcast packet. If it exists, transfer the multicast packet to all interface members of this multicast group.

12.7.2 Static Multicast Configuration

Static Multicast Configuration list

Configure static multicast in following turns :

- Create multicast group
- Add interfaces to multicast group
- Display multicast group information
- Delete interface members from multicast group
- Delete multicast group

12.7.3 Create multicast group

Use following command in global configuration mode to create a multicast group :

multicast mac-address *mac* vlan *vlan-id*

mac : The mac address of multicast group displayed in the form of multicast address, such as : 01 : 00 : 5e : ** : ** : **. *vlan-id* ranges from 1 to 4094. If the VLAN doesn't exist, the multicast group adding fails.

Example :

! Create a multicast group to VLAN 1 with the mac address being 01 : 00 : 5e : 01 : 02 : 03

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

12.7.4 Add interfaces to multicast group

Use multicast mac-address vlan interface command in global configuration mode to add interface to existed multicast group :

multicast mac-address *mac* vlan *vlan-id* interface { all | *interface-list* }

mac : Means mac address of existed multicast which is in the form of multicast mac-address, such as : 01 : 00 : 5e : ** : ** : **. *vlan-id* ranges from 1 to 4094. Multicast group is assembled by *vlan-id* and *mac-address*. *Interface-list* is optional. If all is chosen, all interfaces in system in multicast mac-address vlan interface command. If the VLAN doesn't exist, the multicast group adding fails.

For example :

! Add interface Ethernet 0/0/2 to ethernet 0/0/4 ethernet 0/0/8 to existed multicast group

```
QTECH(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
```

12.7.5 Display multicast group information

Use show multicast command to display the information of the specified or all existed multicast group which includes multicast group interface information, IGMP interface list information :

show multicast [mac-address *mac*]

Mac is the mac address existed in multicast group. If mac-address is not specified, input show multicast command, information of the entire multicast group is displayed.

For example :

! Display the information of multicast group with the MAC address to be 01 : 00 : 5e : 01 : 02 : 03

```
QTECH(config)#show multicast mac-address 01 : 00 : 5e : 01 : 02 : 03
show multicast table information
```

```
MAC Address      : 01 : 00 : 5e : 01 : 02 : 03
VLAN ID         : 1
Static port list : e0/0/2, e0/0/3, e0/0/4, e0/0/8.
IGMP port list
Dynamic port list
Total entries : 1.
```

12.7.6 Delete interface members from multicast group

Use following command in global configuration mode to delete multicast interface member :

no multicast mac-address *mac* vlan *vlan-id* interface { all | *interface-list* }

The meaning of mac, vlan-id and interface-list is the same as that in adding interfaces. Interface in interface-list means the interface member existed in multicast group. All means all the members in multicast group.

For example :

! Delete interface ethernet 5, 6 from existed multicast group.

```
QTECH(config)#no multicast mac-address 01 : 00 : 5e : 01 : 02 : 03 vlan 1 interface
ethernet 0/0/5 ethernet 0/0/6
```

12.7.7 Delete multicast group

Use following command in global configuration mode to delete specified mac address and the multicast group of specified VLAN ID or all multicast groups :

no multicast [mac-address *mac* vlan *vlan-id*]

The meaning of mac, vlan-id and interface-list is the same as that above. They are corresponded to be existed multicast group.

For example :

! Delete multicast group with the mac address being 01 : 00 : 5e : 01 : 02 : 03 and VLAN ID being 1

```
QTECH(config)#no multicast mac-address 01 : 00 : 5e : 01 : 02 : 03 vlan 1
```

12.8 IGMP configuration

12.8.1 IGMP configuration list

Enable multicast route before configuring IGMP.
IGMP configuration list is as following:
Enable multicast protocol
Specify interface running IGMP protocol
Specify interface running IGMP version
Configure the time interval switch sending query packet
Configure switch sending the last member query interval
Configure switch robustness-variable
Configure the number of the multicast group restricted switch interface to add
Configure IGMP the max response time of query packet
Configure interface accessing control list
Configure switch interface to add to multicast group
Configure ingress vlanid of static multicast group members
Enable IGMP proxy

IGMP monitor and maintenance
Enable SSM-Mapping
Enter IGMP mode
Configure SSM-Mapping static group address mapping rules

12.8.2 Enable multicast protocol

Only after enabling multicast protocol, other configuration related to multicast can be effective.
Configure it in global configuration mode:
Enable multicast protocol

ip multicast-routing

Disable multicast protocol

no ip multicast-routing

By default, multicast protocol disables

12.8.3 Specify interface running IGMP protocol

Enable IGMP protocol in interface before switch sending multicast packet.
Configure it in interface mode (include VLAN and SuperVlan interface) :
Run IGMP in specified interface

ip igmp

Disable IGMP in interface

no ip igmp

By default, IGMP is run in any interface.



Caution: Enable IGMP protocol in interface before switch sending multicast packet.

12.8.4 Specify interface running IGMP version

All system run in the same subnetwork must support the same IGMP version. switch can find the switch with other version automatically and inform sys-log, but it cannot shift it automatically.

Configure it in interface mode (include VLAN and SuperVlan interface) :
Configure the version of run IGMP in switch interface

ip igmp version { 1 | 2 | 3 }

Restore the default version of run IGMP in switch interface

no ip igmp version

By default, switch interface runs IGMP Version 2.



Caution: Before configuring interface IGMP, interface must run IGMP protocol. Following commands which configure interface attribution should be attention.

12.8.5 Configure the time interval switch sending query packet

Switch need periodically send Membership Query Message to the network it connected. The time interval id determined by Query Interval timer. User can modify the time interval of IGMP host sending query packet by configuring Query Interval timer.

Configure it in interface mode (include VLAN and SuperVlan interface) :
Configure the time interval of IGMP host sending query packet.

ip igmp query-interval *seconds*

Restore default time interval of IGMP host sending query packet.

no ip igmp query-interval *seconds*

By default, time interval of IGMP host sending query packet is 125 seconds.

12.8.6 Configure switch sending the last member query interval

When switch receives leave packet, it will send special group query packet to know whether there is group member. User can modify the time interval of switch sending special group query packet.

Configure it in interface mode (include VLAN and SuperVlan interface) :
Configure the time interval of switch sending last member query packet

ip igmp last-member-query-interval *seconds*

Restore default time interval of switch sending last member query packet

no ip igmp last-member-query-interval

By default, time interval of switch sending last member query packet is 1 second.



Caution: Only when switch interface running IGMP V2/V3, this configuration is effective. (though running IGMP Version 1 this command can be configured.)

12.8.7 Configure switch robustness-variable

The robustness-variable is a very important parameter to express the operation of IGMP which is used to control the number of sending packets to prevent the loss of the packet in network to strengthen the operation of network protocol.

Configure it in interface mode (include VLAN and SuperVlan interface) :

Configure switch robustness-variable

ip igmp robustness-variable num

Restore default robustness-variable

no ip igmp robustness-variable

By default, robustness-variable is 2.

12.8.8 Configure the number of the multicast group

restricted switch interface to add

Use this command to restrict the number of IGMP groups added in interface, the router will not handle IGMP adding packets if it is beyond the restriction. By default, the maximum number of IGMP groups added in interface is the maximum number of multicast group number (that is maximum hardware table items, considering it can use up all hardware table items through one interface). In configuration, if the added number of IGMP groups is beyond the configuration, the added IGMP group will not be deleted. Repeat this command, the new configuration will cover the original.

Configure it in interface mode (include VLAN and SuperVlan interface) :

Configure the number of the multicast group restricted switch interface to add.

ip igmp limit-group num

Restore the default number of the multicast group restricted switch interface to add.

no ip igmp limit-group

By default, the number of the multicast group restricted to add is 1024.

12.8.9 Configure IGMP the max response time of query packet

After host receiving the query packet from switch, it will enable a Delay Timers for each multicast group it added to, and use a random number between (0, Max Response Time] to be the start value, and Max Response Time is the maximum response time specified by query packet (the maximum response time of IGMP Version 1 is 10 seconds). Host

should inform switch multicast group members before the time is up. If switch hasn't received any multicast member report after the max response time, it thought there is no members in local group and it will never transmit multicast packet it received to network.

Configure it in interface mode (include VLAN and SuperVlan interface)

Configure the max response time of query packet of host members

ip igmp query-max-response-time *seconds*

Restore the default max response time.

no ip igmp query-max-response-time

By default, the max response time in query packet of host member is 10 seconds.



Caution: Only when switch interface running IGMP V2/V3, this configuration is effective. (though running IGMP Version 1 this command can be configured.)

12.8.10 Configure interface accessing control list

Multicast switch makes sure which multicast group contains local group members which connected to switch by sending IGMP query packet. Configure a filtration in interface to make host add to multicast group regulated by IP standard ACL.

Configure it in interface mode (include VLAN and SuperVlan interface)

Control switch receiving the addition of multicast group

ip igmp access-group access-list-number [port-list]

Cancel addition of configured multicast group

no ip igmp access-group access-list-number [port-list]

By default, host can add to any multicast group.

12.8.11 Configure switch interface to add to multicast group

Configure Ethernet switch interface to add to multicast group to make switch transmit multicast packet to it and specify source address list.

Configure it in interface mode (including VLAN interface and superVlan interface):

Configure switch interface to add to multicast group

ip igmp static-group groups-address port-list sourcelist sourcelist

Cancel interface to add to multicast group

no ip igmp static-group groups-address port-list sourcelist sourcelist

12.8.12 **vlanid** Configure ingress **vlanid** of static multicast group members

This command is used with `ip igmp static-group` command. This command specifies ingress vlan id and creates a complete static multicast member table to realize the packet transmission of static multicast members.

Configure it in interface mode (include VLAN and SuperVlan interface)

In VLAN interface mode:

Configure ingress **vlanid** of static multicast group members of Ethernet switch.

ip igmp create-group *groups-address*

Cancel ingress **vlanid** of static multicast group members of Ethernet switch.

no ip igmp create-group *groups-address*

In SuperVlan interface mode:

ip igmp create-group groups-address vlan *vlanid*

Cancel ingress **vlanid** of static multicast group members of Ethernet switch.

no ip igmp create-group groups-address vlan *vlanid*

12.8.13 **Enable IGMP proxy**

After enabling IGMP proxy, switch is the same as a host which reports collected multicast information to uplink multicast router through IGMP proxy, thus, uplink multicast router will transfer corresponded multicast traffic to switch before user getting it. When switch is on the edge of network and only one port connecting to multicast router, enable IGMP proxy to transfer multicast instead of multicast routing protocol.

Configure it in VLAN interface configuration mode:

Enable IGMP proxy

igmp-proxy

Disable IGMP proxy on a port

no igmp-proxy

IGMP proxy is defaulted to be disabled.



Caution: Enable multicast proxy before using interface IGMP proxy.

12.8.14 **IGMP monitor and maintenance**

Display IGMP configuration and running in command line configuration:

Display IGMP interface information:

show ip igmp interface [{ **vlan-interface** *vid* } | { **supervlan-interface** *number* }]

Display static configuration and multicast group information learnt by IGMP:

show ip igmp groups [*multicast-ip*]

Display IGMP proxy information

show ip igmp proxy

Display SSM-Mapping mapping rules

show ip igmp ssm-mapping [*multicast-ip*]

12.8.15 Enable SSM-Mapping

In SSM network, for all kinds of restrictions, some receiving host can only run IGMPv1 or IGMPv2. Configure IGMP SSM Mapping in router to provide SSM service for them.

Configure it in interface mode (including VLAN interface and superVlan interface), specifying subvlan in superVlan mode:

Enable SSM-Mapping

ip igmp ssm-mapping

Disbale SSM-Mapping in interface

no ip igmp ssm-mapping

By default, SSM-Mapping is disabled.



Caution: Enable multicast proxy and run IGMP protocol in specified interface before enable SSM-Mapping.

12.8.16 Enter IGMP mode

Configuring global parameter with IGMP needs entering IGMP view.

Configure it in any configuration mode:

Enter IGMP mode

mroute igmp

Exit igmp mode

exit

12.8.17 Configure SSM-Mapping static group address mapping rule

Realize the same multicast group map to multiple multicast source by multiple configuration in specific SSM multicast group.

Configure it in IGMP mode:

Configure SSM-Mapping static group address mapping rule

ssm-mapping static { *access-control-list source-address* }

Delete SSM-Mapping static group address mapping rule

no ssm-mapping static { *access-control-list source-address* | **all** }

By default, there is no SSM-Mapping static group address mapping rule.



Caution: SSM multicast group address range is specified by configuring ssm multicast group range in PIM mode.

12.9 Brief introduction of PIM

PIM-DM (Protocol Independent Multicast-Dense Mode) is intensive multicast route protocol. PIM-DM suits small scaled network and multicast members are intensive.

12.9.1 Working theory of PIM-DM

The working process of PIM-DM are: Neighbor Discovering, DM forwarding to pruning and grafting.

12.9.1.1 Neighbor Discovering

When enabling PIM-DM router, it needs Hello packet to be neighbor discovering. Each network node running PIM-DM uses Hello packet to keep connection. PIM-DM sends Hello packet periodically.

12.9.1.2 Forwarding & Pruning

PIM-DM supposes all hosts in network are ready to receive multicast data. When some multicast source S sends data to multicast group G, router will be RPF examining after receiving multicast packet. If the examination is passed, router will create a (S,G) item and forward data to all downlink PIM-DM nodes. If the examination fails, that is, multicast packet inputted from error interface, the packet is dropped. After this process, a (S,G) item will be created in PIM-DM multicast area.

If there is no multicast member in downlink node, Prune packet will be sent to uplink node which informs the uplink not to transmit data to downlink node. After receiving Prune packet, uplink node will delete corresponded interface from the outputting interface list of its multicast transmission item (S,G) to set up a SPT (Short Path Tree) with the source S being the root. Prune is originated by leave router.

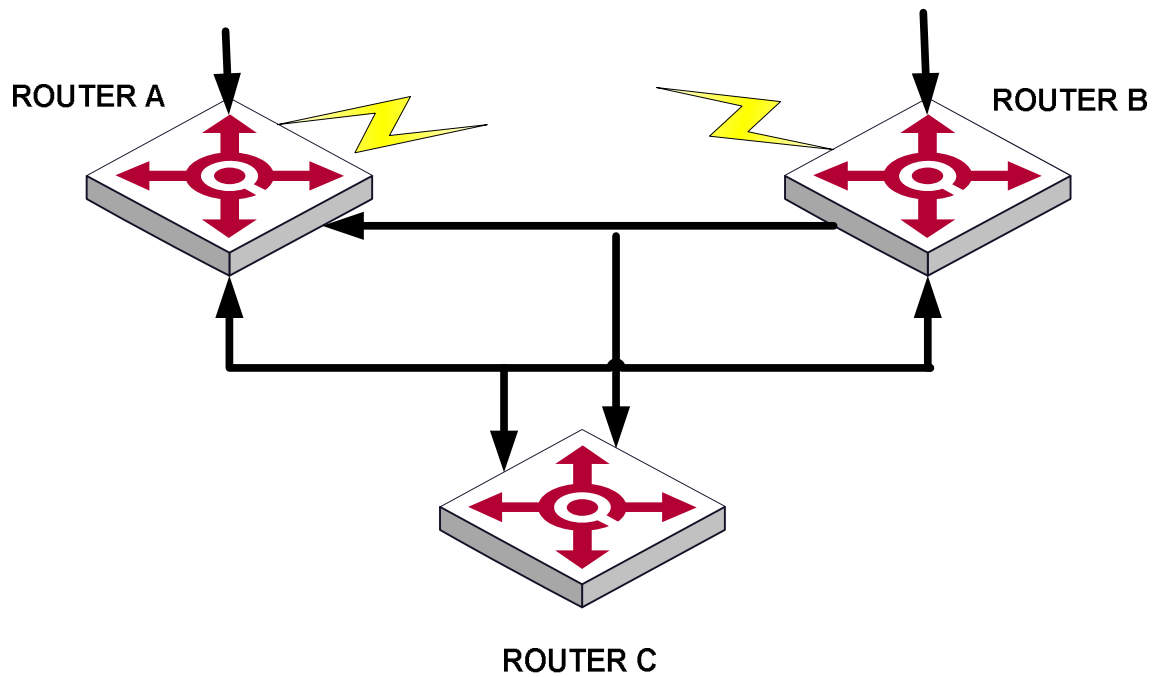
This process is called Forwarding and Pruning. Each pruned node provides overtime mechanism. When pruning is overtime, each router restarts Forwarding and Pruning. The process of PIM-DM Forwarding and Pruning is periodically running.

In this process, PIM-DM adopts RPF examination to establish a multicast transmission tree originated with data source by using current unicast route table. When a multicast packet arrives, router will judge the correctness of arrival path. If arrival interface is the one to multicast source demonstrated by unicast route, the multicast packet is from the correct path; or, this multicast packet will be dropped as redundant packet without transmission. The unicast route information as path judging can be from any unicast route protocol, such as RIP or route information found by OSPF, but not dependent on specified unicast route protocol.

12.9.1.3 Assert mechanism

As following picture, if there are two multicast routers A and B in a LAN network interface and they have their own receiving paths to multicast source S, they will transmit this multicast packet to LAN after receiving multicast packet from multicast source S. multicast router C in downlink node will receive two same multicast packet.

Multicast packet router from uplink node detect this situation, it needs Assertmechanism to select a unique transferrer. By sending Assert packet, select a best path. If the priority and metric of the two or more paths are the same, the one with larger IP address will be the uplink neighbor of (S, G) item, which is responsible for the transmission of (S, G) multicast packet.



Picture for Assert Mechanism

12.9.1.4 Graft

When the pruned downlink node needs to restore to transmission state, this node uses graft packet to inform uplink node. Enable multicast route before configuring IGMP protocol.

12.9.1.5 SRM

To avoid repeat forwarding –pruning, new protocol standard adds this mechanism. Router connected to multicast source timely sends SRM, and PIM will refresh pruning state after receiving it.

12.9.2 Working theory of PIM-SM

The working process of PIM-SM are: Neighbor Discovering, RP sharing tree generating, multicast source register and SPT shift. Neighbor Discovering is the same as that of PIM-DM.

12.9.2.1 RP sharing tree (RPT) generating

When host adding to a multicast group G, leave router connected with this host know there is receiver of multicast group G through IGMP packet, it will calculate the corresponded convergent point RP and send join packet to the upper node. Passing each router from leave router to RP, (*, G) item will be generated in transmission table. No matter where it sending from, it was sent to multicast group G. When RP received the packet to multicast group G, packet will arrive leave router along the established path to host. Above forms RPT with the root of RP.

12.9.2.2 Multicast source register

When multicast source S sends a multicast packet to multicast group G, PIM-SM multicast router will encapsulate received multicast packet to register packet and send it to corresponded RP in the form of unicast. If there are many PIM-SM multicast router in a network interface, DR (Designated Router) will send this multicast packet.

12.9.3 Working theory of PIM-SSM

According to protocol standard, SSM's realization relies on PIM-SM. It can exist with PIM-SM in the same router. Using SSM or PIM-SM is determined by multicast address in data and protocol packet. IANA distributes address segment from 232.0.0.0 to 232.255.255.255, the multicast group of which will not add to share tree but handle by SSM. SSM needs to realize neighbor detect and DR election.

It is realized by IGMPv3 for router host. IGMPv3 adds source filtration and permits host specifying to receive data from specific group, and specifying to receive data of specific source in this group. When SSM receives IS_IN packet of IGMPv3 to know there is host in the network which interface who received IGMP packet connected to want to receive data packet sending to multicast group G from source S. Send PIM (S,G) source group adding packet to the first-hop router connected multicast source according to unicast route towards the direction of source hop-by-hop to establish the shortest path tree between multicast source and the last-hop router connected to receiver. When multicast source sending multicast data, data will be received along the shortest path.

For the host supporting IGMPv1/IGMPv2 but not IGMPv3, configure ssm-mapping in connected router, add packet mapping of group sent by IGMPv1/IGMPv2 to source group to use SSM in network.

12.10 PIM configuration

12.10.1 PIM configuration list

Configuring PIM needs following operation: when router runs in PIM-DM protocol domain, it is suggested enabling PIM-DM in all interface of non-border router. PIM-SM need not enable PIM-SM in all interface.

Basic configuration of PIM:

1. Enable multicast protocol
2. Enable PIM-DM or PIM-SM protocol

Advanced configuration of PIM:

1. Configure Hello packet sending interval
2. Configure BSR border
3. Enter PIM mode
4. Configure multicast source (group) filtrate
5. Configure PIM neighbor filtrate
6. Configure max number of PIM neighbor
7. Configure static RP
8. Specify candidate BSR
9. Specify candidate RP
10. Configure SPT threshold
11. Configure SSM multicast range

12.10.2 Specified interface to run PIM-DM protocol

PIM-DM protocol needs enabling in each interface.

After configuring PIM-DM, it will send Hello packet timely and handle protocol sent by PIM neighbor.

Configure it in VLAN interface configuration mode:

Specified interface to run PIM-DM protocol

ip pim dense-mode

Disable PIM-DM protocol

no ip pim dense-mode

By default, not any interface run PIM-DM protocol. Generally, it is suggested each interface configure PIM-DM. This configuration must be effective in global configuration mode. After enabling PIM-DM, it cannot enable PIM-SM, vice versa.



Caution: Enable multicast protocol before enabling PIM-DM.

12.10.3 Specified interface to run PIM-SM protocol

PIM-SM protocol needs enabling in each interface.

After configuring PIM-SM, it will send Hello packet timely and handle protocol sent by PIM neighbor.

Configure it in VLAN interface configuration mode:

Specified interface to run PIM-SM protocol

ip pim sparse-mode

Disable PIM-SM protocol

no ip pim sparse-mode

By default, not any interface run PIM-SM protocol. Generally, it is suggested each interface configure PIM-SM. This configuration must be effective in global configuration mode. After enabling PIM-SM, it cannot enable PIM-DM, vice versa.



Caution: Enable multicast protocol before enabling PIM-SM.

12.10.4 Configure Hello packet sending interval

After enabling PIM, Hello packet will be sent timely. The time interval sending Hello packet can be modified according to the bandwidth and type of network.

Configure it in VLAN interface configuration mode:

Configure sending interval of Hello packet

ip pim query-interval *seconds*

Restore the default time interval

no ip pim query-interval

Default sending time interval of Hello packet is 30 seconds. User can configure it according to the different network environment.

Generally, this parameter need not modify.



Caution: Enable interface running PIM before configuring PIM attribution. Following commands should pay attention to it.

12.10.5 Configure BSR border

Configure interface to be BSR border of PIM. After configuring this command in some interface, all Bootstrap Message cannot be through the border, but other PIM packets can. Through this, user can divide the network operating PIM-SM into many areas and use different Bootstrap Router in each area. Caution: This command cannot establish multicast border but a PIM Bootstrap Router border.

Configure it in interface configuration mode (including VLAN interface and superVlan interface)
Configure interface to be BSR border

ip pim bsr-border

Delete BSR border

no ip pim bsr-border

By default, bsr-border disables.

12.10.6 Enter PIM mode

Configuring global parameter related to PIM need to enter PIM view.
Configure it in global configuration mode:

mroute pim

Be back from PIM mode

exit

12.10.7 Configure multicast source (group) filtrate

Filtrate according to the source address encapsulated in multicast packet to improve security of network.
Configure it in Pim configuration mode:
Source filtrate received multicast data packet

source-policy *acl-number*

Cancel configuration

no source-policy

If configured ACL, match source address for received all multicast data packet, the failed will be dropped.

12.10.8 Configure PIM neighbor filtrate

Configure basic ACL to restrict all passed routers to be the neighbor of current PIM.
Configure it in VLAN interface configuration:
Filtrate PIM neighbor

ip pim neighbor-policy *acl-number*

Cancel neighbor

no ip pim neighbor-policy

By default, non-filtration.

12.10.9 Configure max number of PIM neighbor

To prevent establishing large number of PIM neighbourship to occupy router's memory and lead to router's failure, it can restrict the number of PIM neighbour in router's interface. The restriction of PIM neighbor is defined by system, user cannot change it by command.

Configure it in VLAN interface configuration mode:

Configure restriction of PIM neighbor number

ip pim neighbor-limit *limit*

Restore default configuration

no ip pim neighbor-limit

By default, the max number of PIM neighbor is 128.

When configuring, PIM neighbor number in interface is beyond the configured value, the original PIM neighbor will not be deleted.

12.10.10 Configure static RP

Static RP can be used as backup of dynamic RP too improve the strength of network.

Configure it in PIM configuration mode:

Configure static RP

static-rp *address*

Delete static RP

no static-rp

By default, PIM do not configure static RP.

If using static RP, all routers in PIM area must use the same configuration. If configured static RP address is the interface address of the state being UP, it is static RP. The static RP interface need not enable PIM protocol.

When selecting RP efficiently by BSR mechanism, static RP is ineffective. If gaining dynamic RP fails, static RP is effective.

12.10.11 Candidate BSR

In a PIM-SM area, there must be unique BSR (Bootstrap Router) to guarantee the normal working of PIM-SM network devices (such as router and switch) . BSR collects and publishes RP information. Many C-BSR select unique and acknowledged BSR through Bootstrap Message. Before it, C-BSR take themselves BSR and they broadcast Bootstrap Message in PIM-SM area. Bootstrap Message contains C-BSR address and priority. PIM selects BSR through C-BSR address and priority . The candidate BSR with superior priority will be selected to be BSR; the one with larger IP address with the same priority will be selected to be BSR.

After configured to be C-BSR, switch sends Bootstrap Message to all PIM neighbors and C-BSR address is specified yo be interface IP address. Each neighbor compares C-BSR address with the former received Bootstrap Message. Address. If C-BSR address is equal to or larger than the former one, PIM neighbor will store the address and transmit Bootstrap Message; or PIM neighbor will drop the received Bootstrap Message. Before receiving Bootstrap Message from a C-BSR with superior priority (or the same priority but larger IP address), C-BSR takes itself BSR. Caution: configure it in backbone network device (router or switch) well connected to other network devices in PIM area.

Generally, configure a C-BSR and C-RP in one network device (switch or router) , which is the core of the network. There is one C-BSR in the same network device, and it will cover the former one.

Configure it in PIM mode:

Configure C-BSR

bsr-candidate *interface-type interface-number hash-mask-length priority*

Delete C-BSR

no bsr-candidate

By default, no candidate BSR is specified.

12.10.12 Specify candidate RP

After selecting BSR, all C-RP will send C-RP Advertisements timely to BSR. BSR collects and publish RP information (There may be many RP and they have different group service range), so that all switches can get RP information.

When configuring C-RP, RP service range can be specified which can serve for all multicast group or some. Configure it in PIM mode:

Configure C-RP

rp-candidate *interface-type interface-number group-list acl-number priority*

Delete C-RP

rp-candidate *interface-type interface-number group-list acl-number*

By default, no candidate RP is specified.

12.10.13 Configure SPT threshold

In PIM-SM mode, receiving host adds to RPT before gaining needed multicast packet through RP. Generally, path in RPT is not the shortest one from receiving host to multicast source. In this case, DR where the receiving host locates can shift to add to SPT to avoid transmission delaying of multicast packet. It supports 2 kinds of fixed thresholds : immediately and infinity, and immediately is the default one.

Configure it in PIM mode:

Configure threshold

spt-threshold { *immediately* | *infinity* }

Restore to default threshold

no spt-threshold

immediately is the default threshold.

12.10.14 Monitor and maintenance of PIM

Display PIM configuration and running:

Display interface information of running PIM:

show ip pim interface [**vlan-interface** vid]

Display PIM neighbor information

show ip pim neighbor

Display multicast route table learnt by PIM

show ip mroute *group-address*

Display PIM current RP information, including dynamic learnt RP and configured static RP.

show ip pim rp-info *group-address*

Display BSR information, including: selected BSR and information about local configured candidate BSR.

show ip pim bsr

Display configured SSM group address range

show ip pim ssm range

12.10.15 Configure SSM multicast group range

Using PIM-SM or PIM-SSM in the process of transmitting information from multicast source to receivers is determined by the multicast group in receiver's subscription (S, G) is in SSM multicast group range or not. All interfaces enabled PIM-SM are considered the multicast group in the range will adopt PIM-SSM.

Configure it in all devices in PIM-SM domain:

Configure it in PIM mode:

Configure SSM multicast group range

ssm {default | range *access-list*}

Delete SSM multicast group range

no ssm {default | range *access-list*}

By default, there is no SSM multicast group range.



Caution: Make sure the configured SSM multicast group range of all devices in the domain are the same, or multicast information cannot transmit through SSM.

If a multicast group is in SSM multicast group range, but the member uses IGMPv1 and IGMPv2 to send adding packet, the device will not use (*, G) adding packet.

Chapter 13 DHCP Configuration

13.1 Brief introduction of DHCP

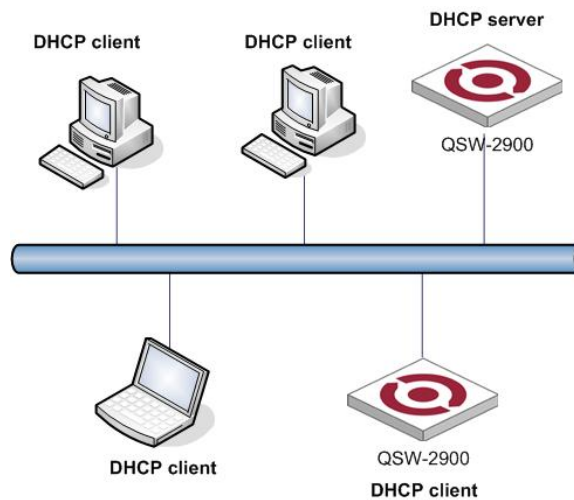
Dynamic Host Configuration Protocol (DHCP) is a network application protocol used by devices (*DHCP clients*) to obtain configuration information for operation in an Internet Protocol network. This protocol reduces system administration workload, allowing devices to be added to the network with little or no manual intervention. DHCP was initially defined in [RFC 1531](#) as a standard-track protocol in October 1993, succeeding the BOOTP. The next update, [RFC 2131](#) released in 1997 is the current DHCP definition for IPv4 networks. The extensions of DHCP for IPv6 (DHCPv6) were published as [RFC 3315](#).

Technical change of IP addresses to DHCP, and each client computer on the LAN has its [IP](#) software configured to request an IP address from the DHCP server during network initialization. The request-and-grant process uses a lease concept with a controllable time period, allowing the DHCP server to reclaim (and then reallocate) IP addresses that are not renewed (dynamic re-use of IP addresses).

With networks getting larger in size and more complicated in structure, lack of available IP addresses becomes the common situation the network administrators have to face, and network configuration becomes a tough task for the network administrators. With the emerging of wireless networks and the using of laptops, the position change of hosts and frequent change of IP addresses also require new technology. Dynamic host configuration protocol (DHCP) is developed in this background.

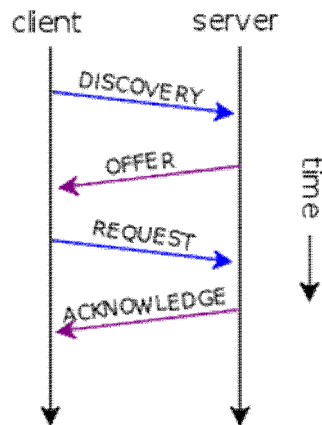
DHCP adopts a client/server model, where DHCP clients send requests to DHCP servers for configuration parameters; and the DHCP servers return the corresponding configuration information such as IP addresses to configure IP addresses dynamically.

A typical DHCP application includes one DHCP server and multiple clients (such as PCs and laptops), as shown below.



Typical DHCP application

13.2 Technical details



Schema of a typical DHCP session

DHCP uses the same two IANA assigned ports as BOOTP : 67/udp for the server side, and 68/udp for the client side.

DHCP operations fall into four basic phases. These phases are IP discovery, IP lease offer, IP request, and IP lease acknowledgement.

After the client obtained an IP address, the client may start an address resolution (ARP) query to prevent IP conflicts caused by address pool overlapping of DHCP servers.

13.2.1 DHCP discovery

The client broadcasts on the physical subnet to find available servers. Network administrators can configure a local router to forward DHCP packets to a DHCP server on a different subnet. This client-implementation creates a UDP packet with the broadcast destination of 255.255.255.255 or subnet broadcast address.

A client can also request its last-known IP address (in the example below, 192.168.1.100). If the client is still in a network where this IP is valid, the server might grant the request. Otherwise, it depends whether the server is set up as [authoritative](#) or not. An authoritative server will deny the request, making the client ask for a new IP immediately. A non-authoritative server simply ignores the request, leading to an implementation-dependent timeout for the client to give up on the request and ask for a new IP address.

13.2.2 DHCP offers

When a DHCP server receives an IP lease request from a client, it reserves an IP address for the client and extends an IP lease offer by sending a DHCPOFFER message to the client. This message contains the client's MAC address, the IP address that the server is offering, the subnet mask, the lease duration, and the IP address of the DHCP server making the offer.

The server determines the configuration, based on the client's hardware address as specified in the CHADDR (Client Hardware Address) field. Here the server, 192.168.1.1, specifies the IP address in the YIADDR (Your IP Address) field.

13.2.3 DHCP requests

A client can receive DHCP offers from multiple servers, but it will accept only one DHCP offer and broadcast a DHCP request message. Based on Transaction ID field in the request, servers are informed whose offer the client has accepted. When other DHCP servers receive this message, they withdraw any offers that they might have made to the client and return the offered address to the pool of available addresses.

13.2.4 DHCP acknowledgement

When the DHCP server receives the DHCPREQUEST message from the client, the configuration process enters its final phase. The acknowledgement phase involves sending a DHCPACK packet to the client. This packet includes the lease duration and any other configuration information that the client might have requested. At this point, the IP configuration process is complete.

The client is expected to configure its network interface with the negotiated parameters.

13.2.5 DHCP information

The client to the DHCP server : either to request more information than the server sent with the original DHCP OFFER; or to repeat data for a particular application - for example, browsers use *DHCP Inform* to obtain web proxy settings via WPAD. Such queries do not cause DHCP server to refresh the IP expiry time in its database.

13.2.6 DHCP releasing

The client sends a request to the DHCP server to release the DHCP information and the client deactivates its IP address. As clients usually do not know when users may unplug them from the network, the protocol does not mandate the sending of *DHCP Release*.

13.2.7 Client configuration parameters

A DHCP server can provide optional configuration parameters to the client. [RFC 2132](#) describes the available DHCP options defined by Internet Assigned Numbers Authority (IANA) - [DHCP and BOOTP PARAMETERS](#).

13.2.8 Options

To identify the vendor and functionality of a DHCP client. The information is a variable-length string of characters or octets which has a meaning specified by the vendor of the DHCP client. One method that a DHCP client can utilize to communicate to the server that it is using a certain type of hardware or firmware is to set a value in its DHCP requests called the Vendor Class Identifier (VCI) (Option 60). This method allows a DHCP server to differentiate between the two kinds of client machines and process the requests from the two types of modems appropriately. Some types of set-top boxes also set the VCI (Option 60) to inform the DHCP server about the hardware type and functionality of the device. The value that this option is set to give the DHCP server a hint about any required extra information that this client needs in a DHCP response.

Ordinary option looks like : |id|len|v1|v2|...|

between || is exactly one byte

len=size in bytes of option value

v1 v2 ... = value in bytes.

Special options are :

id=0x00 has no meaning. It is just byte alignment and has NO LENGTH followed by.

id=0xFF means end of DHCP options and has no length

13.2.9 DHCP IP Address Assignment

13.2.9.1 IP Address Assignment Policy

Currently, DHCP provides the following three IP address assignment policies to meet the requirements of

different clients :

- Manual assignment. The administrator statically binds IP addresses to few clients with special uses (such as WWW server). Then the DHCP server assigns these fixed IP addresses to the clients.
- Automatic assignment. The DHCP server assigns IP addresses to DHCP clients. The IP addresses will be occupied by the DHCP clients permanently.
- Dynamic assignment. The DHCP server assigns IP addresses to DHCP clients for predetermined period of time. In this case, a DHCP client must apply for an IP address at the expiration of the period. This policy applies to most clients.

13.2.9.2 Obtaining IP Addresses Dynamically


A DHCP client undergoes the following four phases to dynamically obtain an IP address from a DHCP server :

1) Discover : In this phase, the DHCP client tries to find a DHCP server by broadcasting a DHCP-DISCOVER packet.

2) Offer : In this phase, the DHCP server offers an IP address. Each DHCP server that receives the DHCP-DISCOVER packet chooses an unassigned IP address from the address pool based on the IP address assignment policy and then sends a DHCP-OFFER packet (which carries the IP address and other configuration information) to the DHCP client. The transmission mode depends on the flag field in the DHCP-DISCOVER packet. For details, see section DHCP Packet Format.

3) Select : In this phase, the DHCP client selects an IP address. If more than one DHCP server sends DHCP-OFFER packets to the DHCP client, the DHCP client only accepts the DHCP-OFFER packet that first arrives, and then broadcasts a DHCP-REQUEST packet containing the assigned IP address carried in the DHCP-OFFER packet.

4) Acknowledge : Upon receiving the DHCP-REQUEST packet, the DHCP server returns a DHCP-ACK packet to the DHCP client to confirm the assignment of the IP address to the client, or returns a DHCP-NAK packet to refuse the assignment of the IP address to the client. When the client receives the DHCP-ACK packet, it broadcasts an ARP packet with the assigned IP address as the destination address to detect the assigned IP address, and uses the IP address only if it does not receive any response within a specified period.

 Note : The IP addresses offered by other DHCP servers (if any) are not used by the DHCP client and are still available to other clients.

13.2.9.3 Updating IP Address Lease

After a DHCP server dynamically assigns an IP address to a DHCP client, the IP address keeps valid only within a specified lease time and will be reclaimed by the DHCP server when the lease expires. If the DHCP client wants to use the IP address for a longer time, it must update the IP lease.

By default, a DHCP client updates its IP address lease automatically by unicasting a DHCP-REQUEST packet to the DHCP server when half of the lease time elapses. The DHCP server responds with a DHCP-ACK packet to notify the DHCP client of a new IP lease if the server can assign the same IP address to the client. Otherwise, the DHCP server responds with a DHCP-NAK packet to notify the DHCP client that the IP address will be reclaimed when the lease time expires.

If the DHCP client fails to update its IP address lease when half of the lease time elapses, it will update its IP address lease by broadcasting a DHCP-REQUEST packet to the DHCP server again when seven-eighths of the lease time elapses. The DHCP server performs the same operations as those described in the previous section.

13.2.9.4 DHCP Packet Format

DHCP has eight types of packets. They have the same format, but the values of some fields in the packets are different. The DHCP packet format is based on that of the BOOTP packets. The following table describes the packet format (the number in the brackets indicates the field length, in bytes) :

DHCPDISCOVER	DHCPOFFER	DHCPREQUEST	DHCPACK
UDP Src=0.0.0.0 sPort=68 Dest=255.255.255.255 dPort=67	UDP Src=192.168.1.1 sPort=67 Dest=255.255.255.255 dPort=68	UDP Src=0.0.0.0 sPort=68 Dest=255.255.255.255 dPort=67	UDP Src=192.168.1.1 sPort=67 Dest=255.255.255.255 dPort=68
OP HTYPE HLEN HOPS	OP HTYPE HLEN HOPS	OP HTYPE HLEN HOPS	OP HTYPE HLEN HOPS
0x01 0x01 0x06 0x00	0x02 0x01 0x06 0x00	0x01 0x01 0x06 0x00	0x02 0x01 0x06 0x00
XID	XID	XID	XID
0x3903F326	0x3903F326	0x3903F326	0x3903F326
SECS FLAGS	SECS FLAGS	SECS FLAGS	SECS FLAGS
0x0000 0x0000	0x0000 0x0000	0x0000 0x0000	0x0000 0x0000
CIADDR	CIADDR	CIADDR	CIADDR (Client IP Address)
0x00000000	0x00000000	0x00000000	0x00000000
YIADDR	YIADDR	YIADDR	YIADDR (Your IP Address)
0x00000000	0xC0A80164	0x00000000	0xC0A80164
SIADDR	SIADDR	SIADDR	SIADDR (Server IP Address)
0x00000000	0x00000000	0x00000000	0x00000000
GIADDR	GIADDR	GIADDR	GIADDR (Gateway IP Address switched by relay)
0x00000000	0x00000000	0x00000000	0x00000000
CHADDR	CHADDR	CHADDR	CHADDR (Client Hardware Address)
0x00053C04	0x00053C04	0x00053C04	0x00053C04
0x8D590000	0x8D590000	0x8D590000	0x8D590000
0x00000000	0x00000000	0x00000000	0x00000000
0x00000000	0x00000000	0x00000000	0x00000000
192 octets of 0's. BOOTP legacy	192 octets of 0's. BOOTP legacy	192 octets of 0's. BOOTP legacy	192 octets of 0's. BOOTP legacy
Magic Cookie	Magic Cookie	Magic Cookie	Magic Cookie
0x63825363	0x63825363	0x63825363	0x63825363
DHCP Options	DHCP Options	DHCP Options	DHCP Options
DHCP option 53 : DHCP Discover	DHCP option 53 : DHCP Offer	DHCP option 53 : DHCP Request	DHCP option 53 : DHCP ACK
DHCP option 50 : 192.168.1.100 requested	DHCP option 1 : 255.255.255.0 subnet mask	DHCP option 50 : 192.168.1.100 requested	DHCP option 1 : 255.255.255.0 subnet mask
DHCP option 55 : Parameter Request List :	DHCP option 3 : 192.168.1.1 router	DHCP option 54 : 192.168.1.1 DHCP server.	DHCP option 3 : 192.168.1.1 router
Request Subnet Mask (1), Router (3), Domain Name (15), Domain Name Server (6)	DHCP option 51 : IP lease time in seconds, 1 day = 86400 s		DHCP option 51 : 1 day IP lease time
	DHCP option 54 : 192.168.1.1 DHCP server		DHCP option 54 : 192.168.1.1 DHCP server
	DHCP option 6 : DNS servers 9.7.10.15, 9.7.10.16, 9.7.10.18		DHCP option 6 : DNS servers 9.7.10.15, 9.7.10.16, 9.7.10.18

Format of DHCP packets

The field meanings are illustrated as follows :

- op : Operation types of DHCP packets : 1 for request packets and 2 for response packets.
- htype, hlen : Hardware address type and length of the DHCP client.
- hops : Number of DHCP relay agents which a DHCP packet passes. For each DHCP relay agent that the DHCP request packet passes, the field value increases by 1.
- xid : Random number that the client selects when it initiates a request. The number is used to identify an address-requesting process.
- secs : Elapsed time after the DHCP client initiates a DHCP request.
- flags : The first bit is the broadcast response flag bit. It is used to identify that the DHCP response packet is sent in the unicast or broadcast mode. Other bits are reserved.

- ciaddr : IP address of a DHCP client.
- yiaddr : IP address that the DHCP server assigns to a client.
- siaddr : IP address of the DHCP server.
- giaddr : IP address of the first DHCP relay agent that the DHCP client passes after it sent the request packet.
- chaddr : Hardware address of the DHCP client.
- sname : Name of the DHCP server.
- file : Name of the start configuration file that the DHCP server specifies for the DHCP client.
- option : Optional variable-length fields, including packet type, valid lease time, IP address of a DNS server, and IP address of the WINS server.

13.2.9.5 DHCP Packet Processing Modes

After the DHCP server is enabled on a device, the device processes the DHCP packet received from a DHCP client in one of the following three modes depending on your configuration :

- Global address pool : In response to the DHCP packets received from DHCP clients, the DHCP server picks IP addresses from its global address pools and assigns them to the DHCP clients.
- Interface address pool : In response to the DHCP packets received from DHCP clients, the DHCP server picks IP addresses from the interface address pools and assigns them to the DHCP clients. If there is no available IP address in the interface address pools, the DHCP server picks IP addresses from its global address pool that contains the interface address pool segment and assigns them to the DHCP clients.
- Trunk : DHCP packets received from DHCP clients are forwarded to an external DHCP server, which assigns IP addresses to the DHCP clients.

You can specify the mode to process DHCP packets. For the configuration of the first two modes, see DHCP Server Configuration. For the configuration of the trunk mode, see DHCP Relay Agent Configuration.

One interface only corresponds to one mode. In this case, the new configuration overwrites the previous one.

13.2.9.6 Protocols and Standards

Protocol specifications related to DHCP include :

- RFC2131 : Dynamic Host Configuration Protocol
- RFC2132 : DHCP Options and BOOTP Vendor Extensions
- RFC1542 : Clarifications and Extensions for the Bootstrap Protocol

13.3 DHCP server configuration list

DHCP packet is broadcasting packet so in layer 3 network structure and using DHCP to distribute IP address, each broadcasting domain needs a DHCP server. For layer 3 network structure by using QTECH QSW-3900 to establish a layer 3 network, each VLAN needs a DHCP server which greatly wastes of resources. A better way to solve this problem is to configure DHCP relay in QTECH QSW-3900 to relay DHCP packet to DHCP server which can need at least only one DHCP server.

Following DHCP functions are supported :

- Support DHCP relay function
- Support specifying DHCP server for each layer 3 interface
- support built-in DHCP server
- support at most 12 address pools and at most 8 network interfaces for each address pool
- support DHCP client to obtain system IP

DHCP configuration list is as following :

- Enable DHCP relay
- Configure DHCP server

- Specify DHCP server for layer 3 interface
- Display DHCP server configuration
- Hide DHCP server
- Support relay option82

13.3.1 Enable DHCP relay

By default, DHCP relay is disabled. Enable DHCP relay in global configuration mode :
Enable DHCP relay

dhcp-relay

Disable DHCP relay

no dhcp-relay

Display DHCP relay in any configuration mode

show dhcp-relay

13.3.2 Configure DHCP server

After enabling DHCP relay, configure DHCP server and specify it to corresponded interface. If IP address of DHCP server is configured to be IP address of any interface or 127.0.0.1, use built-in DHCP server. Configure IP address pool before using built-in DHCP server.

Enable DHCP server and specify corresponded interface IP

dhcp-server group-num ip ip-address

Disable DHCP server

no dhcp-server group-num

Example :

!Configure IP address of DHCP server 1 to be 192.168.0.100

QTECH(config)#dhcp-server 1 ip 192.168.0.100

!Disable DHCP server 1

QTECH(config)#no dhcp-server 1

13.3.3 Specify DHCP server for layer 3 interface

After creating DHCP server, specify DHCP server for each layer 3 interface, and system will relay DHCP packet to DHCP server of this interface after receiving DHCP packet. Use this command in interface configuration mode.

Specify DHCP server for layer 3 interface

dhcp-sever group-num

Delete DHCP server for current layer 3 interface

no dhcp-server

Example :

!Specify DHCP server 1 for VLAN interface 1

QTECH(config-if-vlanInterface-1)#dhcp-server 1

!Delete DHCP server for VLAN interface 1

QTECH(config-if-vlanInterface-1)#no dhcp-server

13.3.4 Display DHCP server configuration

After configuring DHCP server, there are two ways to display DHCP server configuration : one is displaying all DHCP server group, the other is displaying DHCP server of layer 3 interface.

Display DHCP server configuration of all or specified group

show dhcp-server [group-num]

Display DHCP server configuration of layer 3 interface

show dhcp-server interface [{ **supervlan-interface** | **vlan-interface** } *vlan-id*]

vlan-id is layer 3 interface number, if there is no keyword after interface, all DHCP server configuration of all layer 3 interface will be displayed.

Example :

!Display all DHCP server

QTECH(config)#show dhcp-server

!Display DHCP server 1

QTECH(config)#show dhcp-server 1

!Display DHCP server of VLAN interface 1

QTECH(config)#show dhcp-server interface vlan-interface 1

13.3.5 Hide DHCP server

After enabling this function, IP address of DHCP server in IP address information requested by DHCP client cannot be the real IP address of DHCP server, but primary IP address of current interface of QTECH QSW-3900 to hide DHCP server directory.

When DHCP relay of multi-levels exist and this function enables, all-around relay needs enable this function, or the first or the last relay enables, or the network will be abnormal.

Hide IP address of DHCP server

dhcp-relay hide server-ip

13.4 Local IP Address Pool Configuration

Local IP pool is the database which records the IP address DHCP server distributed to DHCP clients which can enquire IP address information DHCP server distributed. In local IP address pool configuration mode, configure parameter of DHCP clients distributed by DHCP server. The configuration options are : gateway and netmask of DHCP client, DNS server, WINS server, lease, IP address range distributed to DHCP client and IP address which is forbidden to distribute and specify. It needs configure local IP address pool before system built-in DHCP server distributing IP address to DHCP client. Enable ip-bind before applying specified IP address in dhcp-client configured client.

- Enter IP address pool configuration mode
- Configure gateway and netmask of local IP address pool
- Configure local IP address pool network interface
- Disable/enable specified IP address in IP address pool
- Configure lease
- Configure DNS
- Configure WINS
- Display IP address pool configuration
- Configure ip bind
- Display IP bind
- Add or delete dhcp client
- Display dhcp client

13.4.1 Enter IP address pool configuration mode

Please configure it in global configuration mode :

Enter ip address pool configuration mode

ip pool *ippoolname*

If IP address pool specified by ippoolname doesn't exist, create this pool.

Delete ip address pool

no ip pool *ippoolname*

Example :

!Enter IP address pool configuration mode

QTECH(config)#ip pool nic

!Delete IP address pool nic

QTECH(config)#no ip pool nic

13.4.2 Configure gateway and netmask of local IP address pool

Use this command in local IP address pool configuration mode :

Configure gateway and netmask of local IP address pool

gateway *ip-address mask*

Parameter *ip-address* is IP address and *mask* is its netmask.

All IP address in local IP address pool must be in the address domain determined by this gateway and netmask and IP address in address pool cannot contain gateway

Example :

!Configure gateway and netmask of local IP address pool

QTECH(config-ip-pool-nic)#gateway 192.168.0.100 255.255.255.0

13.4.3 Configure local IP address pool network interface

Please configure it in local IP address pool configuration mode :

Create local IP address pool network interface

section *section-id from-ip to-ip*

Delete local IP address pool network interface

no section *section-id*

section-id is the section id of this address pool which can configure at most 8 groups. *from-ip* is the start address of this address segment and *to-ip* is the end address. These two addresses must be in the address domain determined by this gateway and netmask and IP address in address pool cannot contain gateway.

Example :

!Create network interface of local IP address pool nic

QTECH(config-ip-pool-nic)#section 0 192.168.0.100 192.168.0.200

!Delete network interface 0 of local IP address pool nic

QTECH(config-ip-pool-nic)#no section 0

13.4.4 Disable/enable specified IP address in IP address pool

Configure it in local IP address pool configuration mode :

Disable/enable specified IP address in local IP address pool network interface

ip { *disable* | *enable* } *ip-address*

ip-address must contain some network interface of local IP address pool.

Example :

!Disable specified IP address 192.168.0.100 in local IP address pool network interface

QTECH(config-ip-pool-nic)#ip disable 192.168.0.100

!Enable specified IP address 192.168.0.100 in local IP address pool network interface

QTECH(config-ip-pool-nic)#ip enable 192.168.0.100

13.4.5 Configure lease time

Configure it in local IP address pool configuration mode :

Configure lease

lease *day : hour : min*

day : hour : min is the lease time which is accurated to minute. The shortest is 0 : 0 : 1 and the longest is 999 : 23 : 59. It is defaulted to be 1 day.

For example :

!Configure lease time to be 1 day 1 hour 1minute

QTECH(config-ip-pool-nic)#lease 1 : 1 : 1

13.4.6 Configure DNS

Configure it in local IP address pool configuration mode :

Configure primary and second DNS

dns { **primary-ip** | **second-ip** } *ip-address*

Delete primary and second DNS

no dns { **primary-ip** | **second-ip** }

Configure DNS suffix

dns suffix *suffix-name*

Delete DNS suffix

no dns suffix

Example :

!Configure primary DNS

QTECH(config-ip-pool-nic)#dns primary-ip 192.168.0.100

!Delete primary DNS

QTECH(config-ip-pool-nic)#no dns primary-ip

13.4.7 Configure WINS

Configure it in local IP address pool configuration mode :

Configure primary and second WINS

wins { **primary-ip** | **second-ip** } *ip-address*

Delete primary and second WINS

no wins { **primary-ip** | **second-ip** }

Example :

!Configure primary WINS

QTECH(config-ip-pool-nic)#wins primary-ip 192.168.0.100

!Delete primary WINS

QTECH(config-ip-pool-nic)#no wins primary-ip

13.4.8 Display IP address pool configuration

Use this command in any configuration mode :

show ip pool [*ippool-name* [*section-num*]]

Display configuration information of specified or all IP address pool

Example :

!Display all IP address pool configuration

QTECH(config)#show ip pool

13.4.9 Configure ip-bind

Configure it in global configuration mode :

Enable ip bind

ip-bind

Disable ip-bind

no ip-bind

Example :

!Enable ip bind

QTECH(config)# ip-bind

!Disable ip-bind

QTECH(config)#no ip-bind

13.4.10 Display ip-bind

Configure it in any mode :

show ip-bind

Display ip-bind configuration.

Example :

!Display ip-bind configuration

QTECH(config)#show ip-bind

13.4.11 Add dhcp client

Configure it in global configuration mode :

Add dhcp client

dhcp-client mac ip vlanid

Delete dhcp client

no dhcp-client mac vlanid

Example

!Add client with mac address being 01 : 00 : 5e : 22 : 22 : 22 , vlan being 2 , ip addrss being 5.5.1.2

QTECH(config)#dhcp-client 01 : 00 : 5e : 22 : 22 : 22 5.5.1.2 2

!Delete client with mac address being 01 : 00 : 5e : 22 : 22 : 22 , vlan being 2

QTECH(config)#no dhcp-client 01 : 00 : 5e : 22 : 22 : 22 2

13.4.12 Show dhcp client

Configure it in any configuration mode :

show dhcp-client

Use this command to display specified IP address , MAC or all client configuration.

Example :

!Display all dhcp client configuration

QTECH(config)#show dhcp-client

13.5 Introduction to DHCP Relay Agent

13.5.1 Usage of DHCP Relay Agent

In small networks DHCP typically uses broadcasts. However, in some circumstances, unicast addresses will be used : when networks have a single DHCP server that provides IP addresses for multiple subnets. When a router for such a subnet receives a DHCP broadcast, it converts it to unicast (with a destination MAC/IP address of the configured DHCP server, source MAC/IP of the router itself). The GIADDR field of this modified request is populated with the IP address of the router interface on which it received the original DHCP request. The DHCP server uses the GIADDR field to identify the subnet of the originating device in order to select an IP address from the correct pool. The DHCP server then sends the DHCP OFFER back to the router via unicast. The router then converts the DHCP OFFER back to a broadcast, sent out on the interface of the original device.



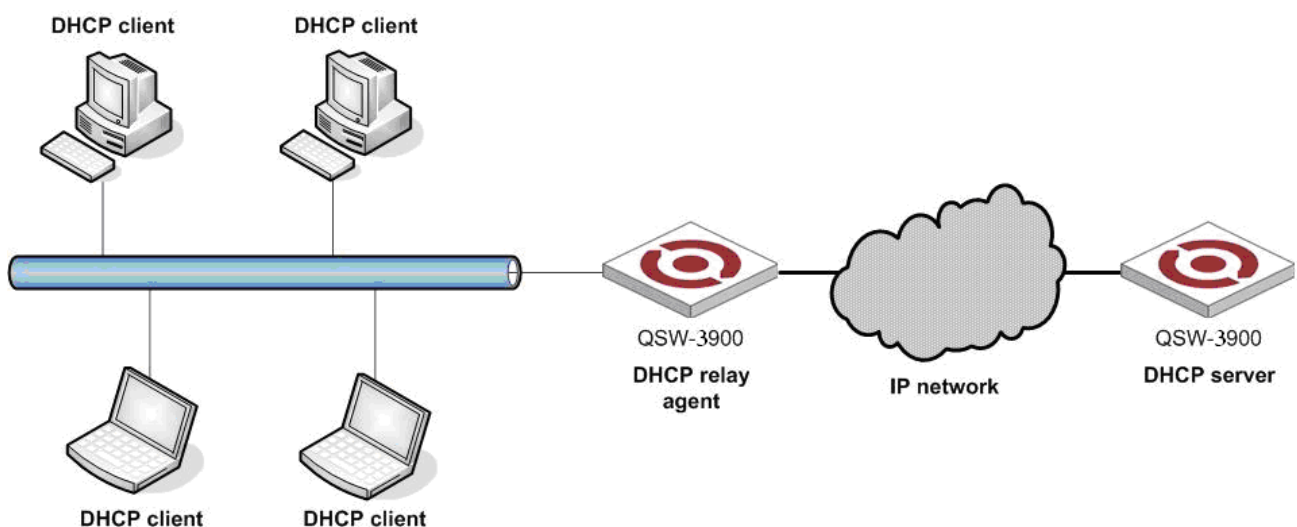
Caution : DHCP messages are usually broadcast packets. So to use DHCP to allocate IP for hosts in a three-level architected network, there need be a DHCP server in every broadcast domain. In a three-level architected network constructed with QTECH QSW-3500 or QSW-3900, a DHCP server is put in each VLAN. This is a greate waste of resources.A solution to this is to use the DHCP relay feature of QTECH QSW-3900, which relays DHCP messages to DHCP servers.Thus only one DHCP server is needed at least.

Since the packets are broadcasted in the process of obtaining IP addresses, DHCP is only applicable to the situation that DHCP clients and DHCP servers are in the same network segment, that is, you need to deploy at least one DHCP server for each network segment, which is far from economical.

The DHCP relay agent is designed to address this problem. It enables DHCP clients in a subnet to communicate with the DHCP server in another subnet so that the DHCP clients can obtain IP addresses. In this case, the DHCP clients in multiple networks can use the same DHCP server, which can decrease your cost and provide a centralized administration.

13.5.2 DHCP Relay Agent Fundamentals

Figure below illustrates a typical DHCP relay agent application.



Typical DHCP relay agent application

DHCP relay agents can transparently transmit broadcast packets on DHCP clients or servers to the DHCP servers or clients in other network segments.

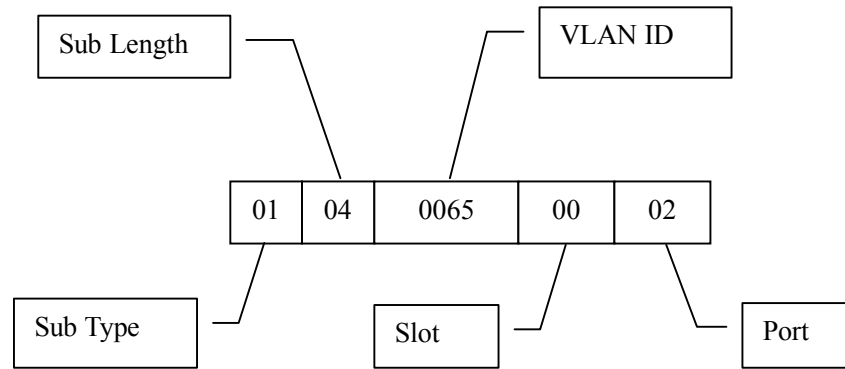
In the process of dynamic IP address assignment through the DHCP relay agent, the DHCP client and DHCP server interoperate with each other in a similar way as they do without the DHCP relay agent. The following sections only describe the forwarding process of the DHCP relay agent. For the interaction process of the packets, see [Obtaining IP Addresses Dynamically](#).

- 1) The DHCP client broadcasts the DHCP-DISCOVER packet.
- 2) After receiving the packets, the network device providing the DHCP relay agent function unicasts the packet to the designated DHCP server based on the configuration.
- 3) The DHCP server assigns IP addresses and transmits the configuration information to the clients through the DHCP relay agent so that the clients can be configured dynamically. The transmission mode depends on the flag field in the DHCP-DISCOVER packet. For details, see section [DHCP Packet Format](#).

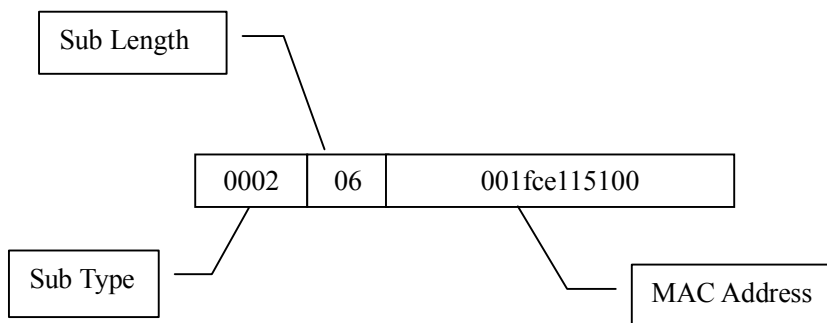
13.5.3 Option 82 Supporting

Option82 is the Relay Agent Information option in DHCP packet defined by rfc 3046. When DHCP client sending requiry packet to DHCP relay, option82 will be added to packet. Option82 in this chapter supports sub-option1, sub-option2 and sub-option5. sub-option1 is one of sub-option of option82 which is Circuit ID with the content being interface VID and MAC address of receiving packet. sub-option2 is also the sub-option of option82 which is Remote ID and is MAC address of relay devices. sub-option5 is also a sub-option of option82 which is Link Selection and is IP address of interface.

The form of sub-option1 is as following :



The form of sub-option2 is as following :



13.5.3.1 Introduction to option 82 supporting

Option 82 is a relay agent information option in DHCP packets. When a request packet from a DHCP client travels through a DHCP relay agent on its way to the DHCP server, the DHCP relay agent adds option 82 into the request packet. Option 82 includes many sub-options, but the DHCP server supports only sub-option 1 and sub-option 2 at present. Sub-option 1 defines agent circuit ID (that is, Circuit ID) and sub-option 2 defines remote agent ID (that is, Remote ID).

Option 82 enables a DHCP server to track the address information of DHCP relay agents, through which and other proper software, you can achieve the DHCP assignment limitation and accounting functions.

13.5.3.2 Primary terminologies

- 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 agent. Generally, sub-option 1 and sub-option 2 must be used together to identify information about a 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 agent, and is usually configured on the DHCP relay agent. Generally, sub-option 1 and sub-option 2 must be used together to identify information about a DHCP source.

13.5.3.3 Mechanism of option 82 supporting on DHCP relay agent

The procedure for a DHCP client to obtain an IP address from a DHCP server through a DHCP relay agent is similar to that for the client to obtain an IP address from a DHCP server directly. The following are the mechanism of option 82 supporting on DHCP relay agent.

- 1) A DHCP client broadcasts a request packet when it initiates.
- 2) The DHCP relay agent on the local network receives the request packet, and then checks whether the packet contains option 82 and processes the packet accordingly.
- 3) If the packet contains option 82, the DHCP relay agent processes the packet depending on the configured policy (that is, discards the packet, replaces the original option 82 in the packet with its own, or leaves the original option 82 unchanged in the packet), and forwards the packet (if not discarded) to the DHCP server.
- 4) If the packet does not contain option 82, the DHCP relay agent adds option 82 to the packet and forwards the packet to the DHCP server. The forwarded packet contains the port number of the switch to which the DHCP client is connected, the VLAN to which the DHCP client belongs, and the MAC address of the DHCP relay agent.
- 5) Upon receiving the DHCP request packet forwarded by the DHCP relay agent, the DHCP server stores the information contained in the option field and sends a packet that contains DHCP configuration information and option 82 to the DHCP relay agent.
- 6) Upon receiving the packet returned from the DHCP server, the DHCP relay agent strips option 82 from the packet and forwards the packet with the DHCP configuration information to the DHCP client.



Note : Request packets sent by a DHCP client fall into two categories : DHCP-DISCOVER packets and DHCP-REQUEST packets. As DHCP servers coming from different manufacturers process DHCP request packets in different ways (that is, some DHCP servers process option 82 in DHCP-DISCOVER packets, whereas the rest process option 82 in DHCP-REQUEST packets), a DHCP relay agent adds option 82 to both types of packets to accommodate to DHCP servers of different manufacturers.

13.6 DHCP relay configuration list

DHCP Configuration list is as following :

- Enable DHCP Relay
- Configure vlan interface
- Show DHCP relay status

13.6.1 Enable DHCP relay

By default, DHCP relay is disabled. To enable DHCP relay, use the following command :

```
Enable DHCP relay
```

dhcp-relay

```
Disable DHCP relay
```

no dhcp-relay

To show DHCP relay status, try the command in any configuration mode :

```
Show DHCP relay status
```

show dhcp-relay

Example :

```
! Enable DHCP relay
```

```
QTECH(config)#dhcp-relay
```

```
! Disable DHCP relay
```

```
QTECH(config)#no dhcp-relay
! Show DHCP relay status
QTECH(config)#show dhcp-relay
```

13.6.2 Configure vlan interface

Configure specified VLAN for relaying DHCP packets. It MUST be the same VLAN, like the PVID of client's port.

Use for example this configuration for set the IP address of DHCP server and specify the interface VLAN aliase :

```
QTECH(config)#vlan vlannumber
QTECH(config-if-vlan)#interface ipaddress mask gateway
QTECH(config-if-vlan)#dhcpserver ip ipaddress
```

13.6.3 Support relay option82

When relay devices receive the DHCP_DISCOVER and DHCP_REQUEST packet sent by client, add option82 and send to server. After receiving the request packet of server, strip option82 before transmitting to client.

Enable option82 support

dhcp option82

Disable option82 support

no dhcp option82

Configure handling strategy of requiry packet contained option82

dhcp option82 strategy {drop|keep|replace}

Display configuration of option82

show dhcp option82

13.7 Introduction DHCP snooping

When DHCP servers are allocating IP addresses to the clients on the LAN, **DHCP snooping** can be configured on LAN switches to harden the security on the LAN to only allow clients with specific IP/MAC addresses to have access to the network.

DHCP snooping is a series of layer 2 techniques. It works with information from a DHCP server to :

- Track the physical location of hosts.
- Ensure that hosts only use the IP addresses assigned to them.
- Ensure that only authorized DHCP servers are accessible.
- In short, DHCP snooping ensures IP integrity on a Layer 2 switched domain.

With DHCP snooping, only a whitelist of IP addresses may access the network. The whitelist is configured at the switch port level, and the DHCP server manages the access control. Only specific IP addresses with specific MAC addresses on specific ports may access the IP network.

DHCP snooping also stops attackers from adding their own DHCP servers to the network. An attacker-controlled DHCP server could wreak havoc in the network or even control it.

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.

- Layer 3 switches can track DHCP client IP addresses through a DHCP relay agent.
- Layer 2 switches can track DHCP client IP addresses through the DHCP snooping function, which listens to DHCP broadcast packets.

When an unauthorized DHCP server exists in the network, a DHCP client may obtain 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 trusted port or an untrusted port through 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.
- Trusted ports forward any received DHCP packet to ensure that DHCP clients can obtain IP addresses from valid DHCP servers. Untrusted ports drop all the received packets.

Figure 1 illustrates a typical network diagram for DHCP snooping application, where Switch B is an QSW-3900 series switch.

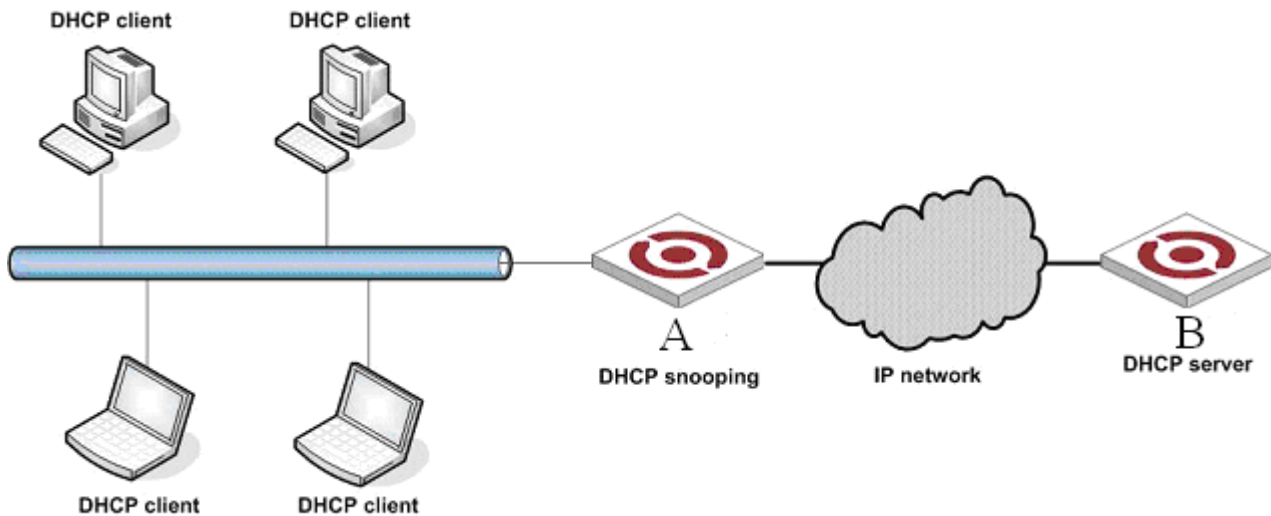


Figure 1 Typical network diagram for DHCP snooping application

Figure 2 illustrates the interaction between a DHCP client and a DHCP server.

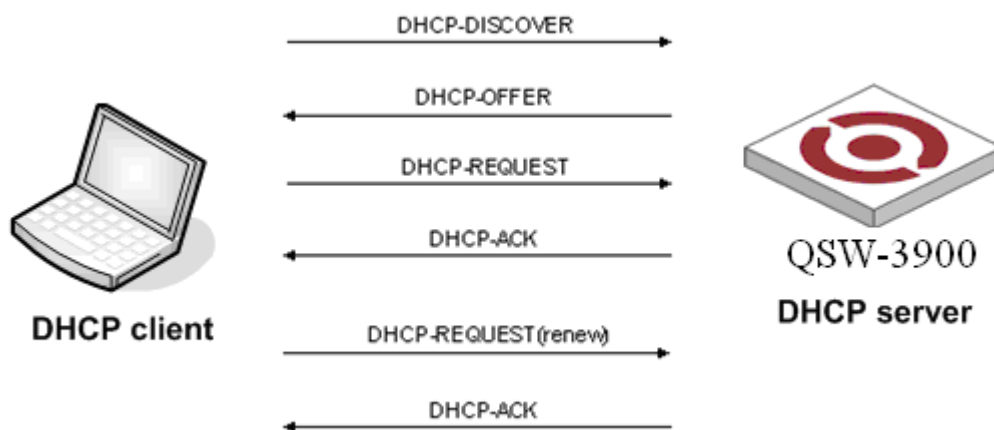


Figure 2 Interaction between a DHCP client and a DHCP server

DHCP snooping listens to 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

For security, DHCP snooping can limit the max number of hosts for a port or for a VLAN in order to avoid animus attacktion.

13.8 DHCP snooping configuration list

13.8.1 Enable DHCP snooping

Enable DHCP snooping

dhcp-snooping

13.8.2 Configure trust ports

Specify some port as trust port. In general, valid servers are connected to the trust ports.
Specify port as trust port

dhcp-snooping trust

13.8.3 Configure max host number

With max host number specified for ports or VLAN, we can avoid animus hosts'ip obtian attacking by DOS and protect servers.

- 1) Configre port/VLAN max host number

dhcp-snooping max-clients *num*

13.8.4 Configure IP source guard

Prevent IP address stolen through IP source guard.

Configure interface IP source guard

ip-source-guard

13.8.5 Show DHCP snooping of ports

DHCP snooping of ports configuraton can be displayed by this command.
Show DHCP snooping configuration of ports

show dhcp-snooping interface [*interface-num*]

13.8.6 Show DHCP snooping configuration of VLANs

DHCP SOOPING configuraton of VLANs can be displayed by this command.
Show DHCP snooping configuration of VLANs

show dhcp-snooping vlan

13.8.7 Show information of clients

Show clients' information of ip address, mac address and port number.
Show information of clients

show dhcp-snooping clients

Chapter 14 ARP Configuration

14.1 Brief Introduction of ARP

In computer networking, the **Address Resolution Protocol (ARP)** is the method for finding a host's link layer (hardware) address when only its Internet Layer (IP) or some other Network Layer address is known. ARP is defined in [RFC 826](#) It is Internet Standard STD 37.

ARP has been implemented in many types of networks; it is not an IP-only or Ethernet-only protocol. It can be used to resolve many different network layer protocol addresses to interface hardware addresses, although, due to the overwhelming prevalence of IPv4 and Ethernet, ARP is primarily used to translate IP addresses to Ethernet MAC addresses. It is also used for IP over other LAN technologies, such as Token Ring, FDDI, or IEEE 802.11, and for IP over ATM.

In the next generation Internet Protocol, IPv6, ARP's functionality is provided by the Neighbor Discovery Protocol (NDP).

ARP is a Link Layer protocol because it only operates on the local area network or point-to-point link that a host is connected to.

ARP is also very often discussed in terms of the Open Systems Interconnect (OSI) networking model, because that model addresses hardware-to-software interfaces more explicitly and is preferred by some equipment manufacturers. However, ARP was not developed based on the design principles and strict encapsulation hierarchy of this model and, therefore, such discussions create a number of conflicts as to the exact operating layer within this model. Most often ARP is placed into the Data Link Layer (Layer 2), but it also requires the definitions of network

The following is the packet structure used for ARP requests and replies. On Ethernet networks, these packets use an EtherType of 0x0806, and are sent to the broadcast MAC address of FF : FF : FF : FF : FF : FF. Note that the EtherType (0x0806) is used in the Ethernet header, and should not be used as the PTYPE of the ARP packet. The ARP type (0x0806) should never be used in the PTYPE field of an ARP packet, since a hardware protocol address should never be linked to the ARP protocol. Note that the packet structure shown in the table has SHA and THA as 48-bit fields and SPA and TPA as 32-bit fields but this is just for convenience — their actual lengths are determined by the hardware & protocol length fields.

bit offset	0 - 7	8 - 15	16 - 31
0	Hardware type (HTYPE)		Protocol type (PTYPE)
32	Hardware length (HLEN)	Protocol length (PLEN)	Operation (OPER)
64	Sender hardware address (SHA) (first 32 bits)		
96	Sender hardware address (SHA) (last 16 bits)		Sender protocol address (SPA) (first 16 bits)
128	Sender protocol address (SPA) (last 16 bits)		Target hardware address (THA) (first 16 bits)
160	Target hardware address (THA) (last 32 bits)		
192	Target protocol address (TPA)		

- **Hardware type (HTYPE)**
Each data link layer protocol is assigned a number used in this field. For example, Ethernet is 1.
- **Protocol type (PTYPE)**
Each protocol is assigned a number used in this field. For example, IP is 0x0800.
- **Hardware length (HLEN)**
Length in bytes of a hardware address. Ethernet addresses are 6 bytes long.
- **Protocol length (PLEN)**
Length in bytes of a logical address. IPv4 address are 4 bytes long.
- **Operation**
Specifies the operation the sender is performing : 1 for request, 2 for reply, 3 for RARP request, and 4 for

- RARP reply.
- **Sender hardware address (SHA)**
Hardware address of the sender.
 - **Sender protocol address (SPA)**
Protocol address of the sender.
 - **Target hardware address (THA)**
Hardware address of the intended receiver. This field is ignored in requests.
 - **Target protocol address (TPA)**
Protocol address of the intended receiver.

14.1.1 ARP announcements

An ARP announcement (also known as *Gratuitous ARP*) is a packet containing valid sender hardware and protocol addresses (SHA and SPA) for the host that sent it, with identical destination and source addresses (TPA = SPA). Such a request is not intended to solicit a reply, but merely updates the ARP caches of other hosts that receive the packet. Gratuitous ARP is usually an ARP request, but it may also be an ARP reply.

Many operating systems perform this during startup. It helps to resolve problems which would otherwise occur if, for example, a network card was recently changed (changing the IP-address-to-MAC-address mapping) and other hosts still have the old mapping in their ARP caches.

Gratuitous ARP is also used by some drivers to ensure load balancing on incoming traffic. In a team of network cards, it is used to announce a different MAC address in the team to receive incoming packets.

ARP announcements can be used to defend link-local IP addresses in the (Zeroconf) protocol ([RFC 3927](#)), and for IP address takeover within high-availability clusters.

14.1.2 ARP probe

An **ARP probe** is an ARP request constructed with an all-zero *sender IP address*. The term is used in the *IPv4 Address Conflict Detection* specification ([RFC 5227](#)). Before beginning to use an IPv4 address (whether received from manual configuration, DHCP, or some other means), a host implementing this specification must test to see if the address is already in use, by broadcasting ARP probe packets.

14.1.3 ARP mediation

ARP mediation refers to the process of resolving Layer 2 addresses when different resolution protocols are used on multiple connected circuits, e.g., ATM on one end and Ethernet on the others.

14.1.4 Variants of the protocol

ARP has also been adapted to resolve many types of Layer 2 addresses; for example, ATMARP is used to resolve ATM NSAP addresses in the Classical IP over ATM protocol.

14.1.5 Inverse ARP and Reverse ARP

The *Inverse Address Resolution Protocol*, also known as *Inverse ARP* or *InARP*, is a protocol used for obtaining Layer 3 addresses (e.g., IP addresses) of other nodes from Layer 2 addresses (e.g. the DLCI in Frame Relay networks). It is primarily used in Frame Relay and ATM networks, where Layer 2 addresses of virtual circuits are sometimes obtained from Layer 2 signaling, and the corresponding Layer 3 addresses must be available before these virtual circuits can be used.

ARP translates Layer 3 addresses to Layer 2 addresses, therefore InARP can be viewed as its inverse. In addition, InARP is actually implemented as an extension to ARP. The packet formats are the same; only the operation code and the certain field values differ.

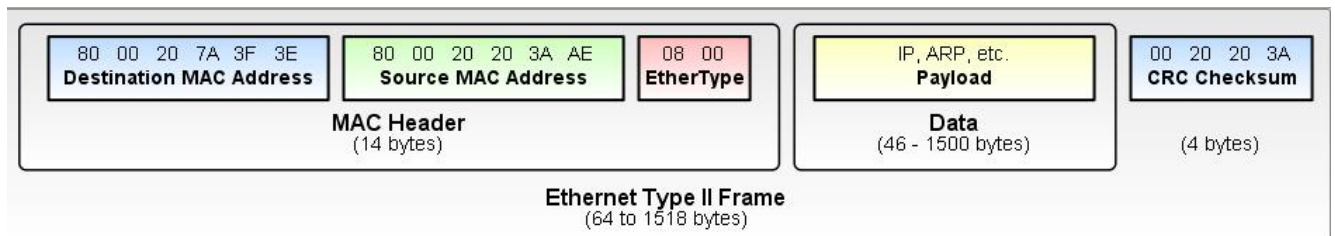
Reverse ARP (RARP), like InARP, also translates Layer 2 addresses to Layer 3 addresses. However, RARP is used to obtain the Layer 3 address of the requesting station itself, while in InARP the requesting station is querying the Layer 3 address of another node. RARP was obsolete by BOOTP which itself has been superseded by the Dynamic Host Configuration Protocol (DHCP).

14.2 ARP spoofing

Address Resolution Protocol (ARP) spoofing, also known as **ARP poisoning** or **ARP Poison Routing (APR)**, is a technique used to attack an Ethernet wired or wireless network. ARP Spoofing may allow an attacker to sniff data frames on a local area network (LAN), modify the traffic, or stop the traffic altogether (known as a denial of service attack). The attack can only be used on networks that actually make use of ARP and not another method of address resolution.

The principle of ARP spoofing is to send fake, or "spoofed", ARP messages to an Ethernet LAN. Generally, the aim is to associate the attacker's MAC address with the IP address of another node (such as the default gateway). Any traffic meant for that IP address would be mistakenly sent to the attacker instead. The attacker could then choose to forward the traffic to the actual default gateway (passive sniffing) or modify the data before forwarding it (man-in-the-middle attack). The attacker could also launch a denial-of-service attack against a victim by associating a nonexistent MAC address to the IP address of the victim's default gateway.

ARP spoofing attacks can be run from a compromised host, or from an attacker's machine that is connected directly to the target Ethernet segment.



A typical Ethernet frame. A spoofed frame could have false source MAC addresses to trick devices on the network.

14.2.1 How ARP spoofing works?

The attacker send fake arp message to the victim causing it to update its ARP table with false entries.

The ARP attack works as follow :

- 1) The attacket send ARP messages to the victim with false updates
- 2) The victim update its ARP table with the attacker MAC address and the false IP address provided by the attacker
- 3) When the victim is ready to send data (Ping in our case) it will send it using mac address listed in its ARP Table (the attacker's)

14.2.2 ARP Spoofing/poising Animation

The attacker is constently sending false ARP messages to the victim causing it to update its ARP table. When you ready to send Ping, watch closley where the ping goes.

14.3 ARP-Proxy

Proxy ARP (Address Resolution Protocol) is a technique by which a device on a given network answers the ARP queries for a network address that is not on that network. The ARP Proxy is aware of the location of the traffic's destination, and offers it's own MAC address in reply, effectively saying, "send it to me, and I'll get it to where it needs to go." Serving as an ARP Proxy for another host effectively directs LAN traffic to the Proxy. The "captured" traffic is then typically routed by the Proxy to the intended destination via another interface or via a

tunnel.

The process which results in the node responding with its own MAC address to an ARP request for a different IP address for proxying purposes is sometimes referred to as 'publishing'.

For more details of configuration of Proxy-ARP please refer to [ARP proxy configuration](#)

14.4 Anti-flood ARP

ARP anti-flood attack means to prevent the same MAC sending plenty of arp packets to influence handling for normal ARP packet. After enabling this function, if the received ARP packet number of fixed source MAC address is beyond configured threshold, it is thought the user of this MAC address is ARP attacking and system will filter this MAC address for delivering anti-attack table item. After delivering the anti-attack table item, this user is banned. By default, ARP anti-attack function is disabled.

14.5 ARP configuration list

Configuration list is as following :

- [Add and delete ARP table item](#)
- [Display ARP table item](#)
- [Configure ARP aging time](#)
- [Display ARP aging time](#)
- [Display ARP table item](#)
- [Enable/disable ARP anti-flood attack](#)
- [Configure deny action and threshold of ARP anti-flood](#)
- [Configure ARP anti-flood recover-time](#)
- [ARP anti-flood MAC recover](#)
- [Display ARP anti-flood attack information](#)
- [Enable/disable ARP anti-spoofing](#)
- [Configure unknown ARP packet handling strategy](#)
- [Enable/disable ARP anti-spoofing valid-check](#)
- [Enable/disable ARP anti-spoofing deny-disguiser](#)
- [Display ARP anti-spoofing](#)

14.5.1 Add and delete ARP table item

Use this command can add or delete a static or dynamic ARP table item. ARP table item not only include corresponding relations of IP and MAC, but also the local VLAN and port number the frame with keyword MAC being destination address has passed.

Add a static ARP table item with the IP address being 192.168.0.100 , MAC address being 00 : 01 : 02 : 03 :

04 : 05 , the corresponded VLAN interface being 1 , and port number being 3 :

QTECH(config)#arp 192.168.0.100 00 : 01 : 02 : 03 : 04 : 05 1 0/3

Delete the corresponded ARP table item of IP address 192.168.0.100 :

QTECH(config)#no arp 192.168.0.100

Delete all static ARP table item :

QTECH(config)#no arp static

Delete all dynamic ARP table item :

QTECH(config)#no arp dynamic

Delete all ARP table item :

QTECH(config)#no arp all

14.5.2 Display ARP table item

Use this command to display static, dynamic, specified IP address or all ARP table item.

Display all ARP table item :

show arp all

Display dynamic ARP table item :

show arp dynamic

Display static ARP table item :

show arp static

Display all ARP table item with the IP address being 192.168.0.100 :

QTECH(config)#show arp 192.168.0.100

14.5.3 Configure ARP aging time

Use this command to modify ARP aging time :

arp aging *seconds*

14.5.4 Display ARP aging time

Use this command to display ARP aging time :

show arp aging

14.5.5 Display ARP table item

Use this command to display static, dynamic, specified IP address or all ARP table item.

Display all ARP table item :

show arp all

Display dynamic ARP table item :

show arp dynamic

Display static ARP table item :

QTECH(config)#show arp static

Display all ARP table item with the IP address being 192.168.0.100 :

QTECH(config)#show arp 192.168.0.100

14.5.6 Enable/disable ARP anti-flood attack

Use following command in global configuration mode to enable it :

Enable ARP anti-flood attack

arp anti-flood

Disable ARP anti-flood attack

no arp anti-flood

14.5.7 Configure deny action and threshold of ARP anti-flood

ARP anti-flood attack has two kind of source mac deny for arp overspeed (the speed of sending arp packet is beyond threshold) : one is deny arp packet from this mac, the other is deny all packets from this mac. Configure following command in global configuration mode :

arp anti-flood action { deny-arp | deny-all } **threshold** *rate-limit*

Threshold range is from 1-100 pps. By default, the deny action is deny-arp and threshold is 16 pps.

Example :

! Configure deny action to be all packets deny and threshold to be 10 pps

```
QTECH(config)#arp anti-flood action deny-all threshold 10
```

14.5.8 Configure ARP anti-flood recover-time

The banned MAC in ARP anti-flood attack will be auto-recover after a certain time. Use this command in global configuration mode :

arp anti-flood recover-time *time*

The recover time can be configured in the range of 0-1440 minutes. If time is 0 , it means never auto-recover.

Example :

! Configure recover time to be 20 minutes

```
QTECH(config)#arp anti-flood recover-time 20
```

Default recover time is 10 minutes.

14.5.9 ARP anti-flood MAC recover

The banned MAC can auto-recover after recover time and specified and all banned MAC can cover manually. Use this command in global configuration mode :

arp anti-flood recover { mac | all }

Example :

! Recover banned mac : 00 : 1f : ce : 00 : 02 : 02

```
QTECH(config)#arp anti-flood recover 00 : 1f : ce : 00 : 02 : 02
```

! Recover all banned mac

```
QTECH(config)#arp anti-flood recover all
```

14.5.10 Display ARP anti-flood attack information

Use this command to show arp anti-flood :

show arp anti-flood

14.5.11 Bind blackhole mac generated by arp anti-flood to be general

Use this command to bind blackhole mac (non- decompiling) generated by arp anti-flood to be general (decompiling) :

arp anti-flood bind blackhole { mac | all }

For example :

! Bind mac : 00 : 1f : ce : 00 : 02 : 02

```
QTECH(config)#arp anti-flood bind blackhole 00 : 1f : ce : 00 : 02 : 02
```

! Bind all blackhole mac generated by all arp anti-flood

```
QTECH(config)#arp anti-flood bind blackhole all
```

14.5.12 Enable/disable ARP anti-spoofing

ARP anti-spoofing is used to check the match of ARP packet and configured static ARP. After enabling this function, all ARP through switch will be redirected to CPU. If source IP, source MAC, interface number, vlan id and static ARP are totally matched, it is thought to be valid and permitted normal handling and transmit. If not, drop it. If there is not corresponded static ARP table item, handle it as strategy of configuring unknown arp packet : drop it or flood (send to each interface) and ARP anti-flood is defaulted to be disabled. Use this command in global configuration mode to enable it :

Enable arp anti-spoofing

arp anti-spoofing

Disable arp anti-spoofing

```
QTECH(config)#no arp anti-spoofing
```

14.5.13 Configure unknown ARP packet handling strategy

Use following command to configure unknown ARP packet handling strategy.

```
arp anti-spoofing unknown { discard | flood }
```

Example :

! Configure unknown ARP packet handling strategy to be flood

```
QTECH(config)#arp anti-spoofing unknow flood
```

Strategy discard means to drop unknown arp packet without corresponded static arp. Strategy flood means to flood to each interface, transmit to each interface. The default strategy is discard.

14.5.14 Enable/disable ARP anti-spoofing valid-check

Source MAC of Ethernet data frame head of some ARP attack packet is different from that of ARP protocol packet. After enabling this function, it will check whether the source mac of arp packet sending to cpu is the as that in arp protocol packet. Drop it if they are different. This function is defaulted to be disabled. Use this command in global configuration mode to enable it :

Enable ARP anti-spoofing valid-check :

```
arp anti-spoofing valid-check
```

Disable ARP anti-spoofing valid-check :

```
QTECH(config)#no arp anti-spoofing valid-check
```

14.5.15 Enable/disable ARP anti-spoofing deny-disguiser

ARP gateway disguiser means attacker disguising gateway address to send free ARP packet whose gateway address is source IP address in LAN. After host in LAN receiving this packet, the original gateway address will be modified to be address of attacker to cause all hosts in LAN cannot visit network. Enable arp anti-spoofing deny-disguiser to solve this problem. After enabling this function, when switch cpu receives the ARP packet which is conflict with gateway address, push source mac of arp protocol packet to mac blackhole and send its own free arp. It will check arp broadcast packet. Those arp unicast packet not only for arp will not be checked for no uplink cpu. This function is defaulted to be disabled. Use following command to enable it :

Enable ARP anti-spoofing deny-disguiser :

arp anti-spoofing deny-disguiser

Disable ARP anti-spoofing deny-disguiser :

```
QTECH(config)#no arp anti-spoofing deny-disguiser
```

14.5.16 Display ARP anti-spoofing

Use this command to show ARP anti-spoofing :

```
show arp anti-spoofing
```

14.5.17 Configure trust port of ARP anti-attack

Use this command to set the port to be trust and ARP packet from this port will not be check attacking and spoofing.

!Configure e0/0/1 to be trust

```
QTECH(config-if-ethernet-0/0/1)#arp anti trust
```


Chapter 15 ACL Configuration

15.1 ACL Overview

An access control list (ACL) is used primarily to identify traffic flows. In order to filter data packets, a series of match rules must be configured on the network device to identify the packets to be filtered. After the specific packets are identified, and based on the predefined policy, the network device can permit/prohibit the corresponding packets to pass.

ACLs classify packets based on a series of match conditions, which can be the source addresses, destination addresses and port numbers carried in the packets.

The packet match rules defined by ACLs can be referenced by other functions that need to differentiate traffic flows, such as the definition of traffic [classification rules in QoS](#), [policy-based vlan](#), [selective QinQ](#) and others.

According to the application purpose, ACLs fall into the following four types :

- Standard ACL : rules are made based on the Layer 3 source IP addresses only.
- Extended ACL : rules are made based on the Layer 3 and Layer 4 information such as the source and destination IP addresses of the data packets, the type of protocol over IP, protocol-specific features, and so on.
- Link-based ACL : rules are made based on the Layer 2 information such as the source and destination MAC address, VLAN priority, Layer 2 protocol, and so on.
- User-based ACL : such rules specify a byte in the packet, by its offset from the packet header, as the starting point to perform logical AND operations, and compare the extracted string with the user-defined string to find the matching packets for processing.

15.1.1 ACL Match Order

An ACL may contain a number of rules, which specify different packet ranges. This brings about the issue of match order when these rules are used to filter packets.

An ACL supports the following two types of match orders :

- Configured order : ACL rules are matched according to the configured order.
- Automatic ordering : ACL rules are matched according to the “depth-first” order.

15.1.1.1 IP ACL depth-first order

With the depth-first rule adopted, the rules of an IP ACL (standard and extended) are matched in the following order :

- 1) Protocol range of ACL rules. The range of IP protocol is 1 to 255 and those of other protocols over IP are the same as the corresponding protocol numbers. The smaller the protocol range, the higher the priority.
- 2) Range of source IP address. The smaller the source IP address range (that is, the longer the mask), the higher the priority.
- 3) Range of destination IP address. The smaller the destination IP address range (that is, the longer the mask), the higher the priority.
- 4) Range of Layer 4 port number, that is, of TCP/UDP port number. The smaller the range, the higher the priority.

If rule A and rule B are the same in all the four ACEs (access control elements) above, and also in their numbers of other ACEs to be considered in deciding their priority order, weighting principles will be used in deciding their priority order.

The weighting principles work as follows :

- Each ACE is given a fixed weighting value. This weighting value and the value of the ACE itself will jointly decide the final matching order. The weighting values of ACEs rank in the following descending order : ToS, ICMP, established, precedence, fragment.
- The weighting value of each ACE of the rule is deducted from a fixed weighting value. The smaller the weighting value left, the higher the priority.
- If the number and type of ACEs are the same for multiple rules, then the sum of ACE values of a rule determines its priority. The smaller the sum, the higher the priority.

15.1.1.2 Layer 2 ACL depth-first order

With the depth-first order adopted, the rules of a Layer 2 ACL are matched in the order of the mask length of the source MAC address and destination MAC address, the longer the mask, the higher the match priority. If two mask lengths are the same, the priority of the match rule configured earlier is higher. For example, the priority of the rule with source MAC address mask FFFF-FFFF-0000 is higher than that of the rule with source MAC address mask FFFF-0000-0000.

15.1.2 Ways to Apply ACL on a Switch

15.1.2.1 ACLs activated directly on the hardware

In a switch, an ACL can be directly activated on the switch hardware for packet filtering and traffic classification in the data forwarding process. You can use the `acl order` command to specify the match order for the rules in the ACL. For detailed configuration, refer to [Matching Order of ACL Rules](#).

ACLs are directly activated on the switch hardware in the following situations : the switch references ACLs to implement the QoS functions, and forwards data through ACLs.

15.1.2.2 ACL referenced by the upper-level modules

The switch also uses ACLs to filter packets processed by software and implements traffic classification. In this case, there are two types of match orders for the rules in an ACL : `config` (user-defined match order) and `auto` (the system performs automatic ordering, namely according to the “depth-first” order). In this scenario, you can specify the match order for multiple rules in an ACL. You cannot modify the match order for an ACL once you have specified it. You can specify a new the match order only after all the rules are deleted from the ACL.

ACLs can also be referenced by route policies or be used to control login users.

15.1.3 ACLs Based on Time Ranges

A time range-based ACL enables you to implement ACL control over packets by differentiating the time ranges.

A time range can be specified in each rule in an ACL. If the time range specified in a rule is not configured, the system will give a prompt message and allow such a rule to be successfully created. However, the rule does not take effect immediately. It takes effect only when the specified time range is configured and the system time is within the time range. If you remove the time range of an ACL rule, the ACL rule becomes invalid the next time the ACL rule timer refreshes.

15.2 Configuring ACL

15.2.1 Matching order configuration

An ACL rule consists of many “permit | deny” syntax, and the range of data packet specified by each syntax is different. When matching a data packet and ACL rule, there should be order. Use following command to configure ACL matching order :

access-list *access-list-number* **match-order** { config | auto }

Parameter :

access-list-number : the number of ACL rule which is in the range of 1 to 399.

config : Specify user configured order when matching this rule.

auto : Specify auto-sequencing when matching this rule. (according to the deep precedency) It is defaulted to specify user configured order, that is “config”. Once user configures the matching order of an ACL rule, it cannot be changed unless delete the content of the rule and re-configure its order.

The deep precedency used by auto means locating the syntax with the smallest data range at the end, which can be realized by comparing address wildcard. The smaller the wildcard value is, the smaller range the host has. For example, 192.168.3.1 0 specifies a host : 192.168.3.1, while 192.168.3.1 0.0.255.255 specifies a network interface : 192.168.3.1 = 192.168.255.255. The former is before the latter in ACL. The concrete rule is : For standard ACL syntax, compare source address wildcard, if their wildcard is the same, use config order; for layer 2 ACL, the rule with “any” is in the front, others use config order; for extended ACL, compare source address wildcard, if they are the same, compare destination address wildcard, if they are the same, compare interface number range, the smaller is in the back, if the interface number range is the same, use config order; for user-defined ACL, compare the length of mask, the longer is in the back, if they are the same, use config order.

15.2.2 ACL support

ACL is the command control list applied to switch. These command is used to tell switch which data packet to receive and which to refuse. It consists of a series of judging syntax. After activating an ACL, switch will examine each data packet entering switch according to the judging condition given by ACL. The one which satisfies the ACL will be permit or dropped according to ACL. QOS introduces the permit rule configuration.

In system, the ACL can be classified as following :

- Standard ACL based on number ID
- Standard ACL based on name ID
- Extended ACL based on number ID
- Extended ACL based on name ID
- Layer 2 ACL based on number ID
- Layer 2 ACL based on name ID
- User-defined ACL based on number ID
- User-defined ACL based on name ID

The restriction to every ACL and number of QOS action is as following table :

Table 1 ACL number restriction

Standard ACL based on number ID	1-99	99
Extended ACL based on number ID	100-199	100
Layer 2 ACL based on number ID	200-299	100
Standard ACL based on name ID	--	1000
Extended ACL based on name ID	--	1000
Layer 2 ACL based on name ID	--	1000

Sub-rule number which can be configured by an ACL	0-127	128
The max sub-rule number which can be configured	--	3000
Time range	--	128
The absolute time range which can be configured by a time range	--	12
The periodic time range which can be configured by a time range	--	32
Sub-item of activating ACL	--	1416

15.3 ACL configuration

15.3.1 Configuration list

ACL configuration includes :

- Configure time range
- Define ACL
- Activate ACL

Above three steps should be in order. Configure time range at first, then define ACL which will introduce defined time range and activate ACL.

15.3.2 Configure time range

15.3.2.1 Enter time-range configuration mode

Use time-range command to enter time-range configuration mode. In this mode, you can configure time range.

Configure it in global configuration mode.

Command :

time-range *time-range-name*

There are two kinds of configuration : configure absolute time range and periodic time range. Configuring absolute is in the form of year, month, date, hour and minute. Configuring periodic time range is in the form of day of week, hour and minute.

15.3.2.2 Create absolute time range

Use following command to configure it.

Configure it in time-range configuration mode.

Configure absolute time range :

absolute [**start** *time date*] [**end** *time date*]

Delete absolute time range :

no absolute [**start** *time date*] [**end** *time date*]

If the start time is not configured, there is no restriction to the start time.; if endtime is not configured, the end time can be the max time of system. The end time must be larger than start time.

Absolute time range determines a large effective time and restricts the effective time range of periodic time. It can configure 12 absolute time range.

15.3.2.3 Create periodic time range

Use following command to configure periodic time range.

Configure it in time-range configuration mode.

Command :

periodic *days-of-the-week hh : mm : ss to [day-of-the-week] hh : mm : ss*

no periodic *days-of-the-week hh : mm : ss to [day-of-the-week] hh : mm : ss*

The effective time range of periodic time is a week. It can configure at most 32 periodic time range.

15.3.3 Standard ACL

Switch can define at most 99 standard ACL with the number ID (the number is in the range of 1 to 99), at most 1000 standard ACL with the name ID and totally 3000 sub-rules. It can define 128 sub-rules for an ACL (this rule can suit both ACL with name ID and number ID). Standard ACL only classifies data packet according to the source IP information of IP head of data packet and analyse the matching data packet. The construction of IP head refers to RFC791.

15.3.3.1 Define standard ACL based on number ID

Standard ACL based on number ID is using number to be ID of standard ACL. Use following command to define standard ACL based on number ID.

Configure it in global configuration mode.

Command :

access-list *access-list-number* { **deny** | **permit** } { *source-addr source-wildcard* | any } [**fragments**] [**time-range** *time-range-name*]

Define the matching order of ACL :

access-list *access-list-number* **match-order** { config | auto }

Delete all the subitems or one subitem in one ACL with number ID or name ID or all ACLs.

no access-list { all | { *access-list-number* | **name** *access-list-name* } [*subitem*] }

Use access-list command repeatedly to define more rules for the same ACL.

If parameter time-range is not used, this ACL will be effective at any time after activation.

Concrete parameter meaning refers to corresponded command line.

15.3.3.2 Define standard ACL with name ID.

Defining standard ACL with name ID should enter specified configuration mode : use **access-list standard** in global configuration mode which can specify matching order of ACL. Use exit command to be back from this mode.

Enter standard ACL with name ID configuration mode(global configuration mode)

access-list standard *name* [match-order { config | auto }]

Defining standard ACL rule (standard ACL with name ID configuration mode)

{ **permit** | **deny** } { *source-addr source-wildcard* | any } [**fragments**] [**time-range** *time-range-name*]

Delete all the subitems or one subitem in one ACL with number ID or name ID or all ACLs.(global configuration mode)

no access-list { all | { *access-list-number* | **name** *access-list-name* } [*subitem*] }

Use { permit | deny } command repeatedly to define more rules for the same ACL. Specifying matching order cannot be modified.

By default, the matching order is user configured order (config).

Concrete parameter meaning refers to corresponded command line.

15.3.4 Define extended ACL

Switch can define at most 100 extended ACL with the number ID (the number is in the range of 100 to 199),

at most 1000 extended ACL with the name ID and totally 3000 sub-rules. It can define 128 sub-rules for an ACL (this rule can suit both ACL with name ID and number ID). Extended ACL classifies data packet according to the source IP, destination IP, used TCP or UDP interface number, packet priority information of IP head of data packet and analyse the matching data packet. Extended ACL supports three types of packet priority handling : TOS(Type Of Service) priority, IP priority and DSCP. The construction of IP head refers to RFC791.

15.3.4.1 Define extended ACL with number ID

Extended ACL based on number ID is using number to be ID of extended ACL. Use following command to define extended ACL based on number ID.

```
access-list access-list-number2 { permit | deny } [ protocol ] [ established ] { source-addr source-wildcard | any }
[ port [ portmask ] ] { dest-addr dest-wildcard | any } [ port [ portmask ] ] [ icmp-type [ icmp-code ] ] [ fragments ]
{ [ precedence precedence ] [ tos tos ] | [ dscp dscp ] } [ time-range time-range-name ]
```

Define the matching order of ACL

```
access-list access-list-number match-order { config | auto }
```

Delete all the subitems or one subitem in one ACL with number ID or name ID or all ACLs.

```
no access-list { all | { access-list-number | name access-list-name } [ subitem ] }
```

Use **access-list** command repeatedly to define more rules for the same ACL.

Number ID of extended ACL is in the range of 100 to 199.

Caution : parameter port means TCP or UDP interface number used by all kinds of superior levels. For some common interface number, use corresponded mnemonic symbol to replace the real number, such as using “bgp” to instead of the TCP interface number 179 of BGP protocol. Details refer to corresponded command line.

15.3.4.2 Define extended ACL with name ID

Defining standard ACL with name ID should enter specified configuration mode : use access-list extended in global configuration mode which can specify matching order of ACL. Use exit command to be back from this mode.

Configure it in corresponded mode. Enter extended ACL with name ID (global configuration mode).

```
access-list extended name [ match-order { config | auto } ]
```

Define extended ACL (extended ACL with name ID configuration mode)

```
{ permit | deny } [ protocol ] [ established ] { source-addr source-wildcard | any } [ port [ portmask ] ] { dest-addr
dest-wildcard | any } [ port [ portmask ] ] [ icmp-type [ icmp-code ] ] { [ precedence precedence ] [ tos tos ] | [ dscp
dscp ] } [ fragments ] [ time-range time-range-name ]
```

Delete all the subitems or one subitem in one ACL with number ID or name ID or all ACLs.(global configuration mode)

```
no access-list { all | { access-list-number | name access-list-name } [ subitem ] }
```

Use { permit | deny } command repeatedly to define more rules for the same ACL. Specifying matching order cannot be modified.

Caution : parameter port means TCP or UDP interface number used by all kinds of superior levels. For some common interface number, use corresponded mnemonic symbol to replace the real number, such as using “bgp” to instead of the TCP interface number 179 of BGP protocol. Details refer to corresponded command line.

15.3.5 Define layer 2 ACL

Switch can define at most 100 layer 2 ACL with the number ID (the number is in the range of 200 to 299), at most 1000 layer 2 ACL with the name ID and totally 3000 sub-rules. It can define 128 sub-rules for an ACL (this rule can suit both ACL with name ID and number ID). Layer 2 ACL only classifies data packet according to the source MAC address, source VLAN ID, layer protocol type, layer packet received and retransmission interface and destination MAC address of layer 2 frame head of data packet and analyze the matching data packet.

15.3.5.1 Define layer 2 ACL based on number ID

Layer 2 ACL based on number ID is using number to be ID of layer 2 ACL. Use following command to define layer 2 ACL based on number ID.

Configure it in global configuration mode.

```
access-list access-list-number3 { permit | deny } [ protocol ] [ cos vlan-pri ] ingress { { [ source-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** time-range-name]

Define the matching order of ACL :

access-list *access-list-number* **match-order** { config | auto }

Delete all the subitems or one subitem in one ACL with number ID or name ID or all ACLs.

no access-list { all | { *access-list-number* | **name** *access-list-name* } [*subitem*] }

Use access-list command repeatedly to define more rules for the same ACL.

The number ID of layer 2 ACL is in the range of 200 to 299.

Interface parameter in above command specifies layer 2 interface, such as Ethernet interface. Concrete parameter meaning refers to corresponded command line.

15.3.5.2 Define layer 2 ACL with name ID.

Defining layer 2 ACL with name ID should enter specified configuration mode : use access-list link in global configuration mode which can specify matching order of ACL. Use exit command to be back from this mode.

Enter layer 2 ACL with name ID configuration mode(global configuration mode)

access-list link *name* [**match-order** { config | auto }]

Defining layer 2 ACL rule(layer 2 ACL with name ID configuration mode)

{ **permit** | **deny** } [*protocol*] [**cos** *vlan-pri*] **ingress** { { [*source-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** *time-range-name*]

Delete all the subitems or one subitem in one ACL with number ID or name ID or all ACLs.(global configuration mode)

no access-list { all | { *access-list-number* | **name** *access-list-name* } [*subitem*] }

Use { permit | deny } command repeatedly to define more rules for the same ACL. Specifying matching order cannot be modified.

By default, the matching order is user configured order (config).

Concrete parameter meaning refers to corresponded command line.

15.3.6 Activate ACL

After activating ACL, it can be effective. Use access-group command to activate accessing control list.

Configure it in global configuration mode.

Activate ACL

access-group { **user-group** { *access-list-number* | *access-list-name* } [**subitem** *subitem*] | { [**ip-group** { *access-list-number* | *access-list-name* } [**subitem** *subitem*]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem** *subitem*]] } }

Cancel activating ACL

no access-group { all | **user-group** { *access-list-number* | *access-list-name* } [**subitem** *subitem*] | { [**ip-group** { *access-list-number* | *access-list-name* } [**subitem** *subitem*]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem** *subitem*]] } }

Instruction :

This command supports activating accessing control list of layer 2 and layer 3 at the same time, but the action of each accessing control list should not be conflict, if there is conflict (such as one is permit, the other is deny), the activation fails. Switch uses straight through to activate layer 2 and layer 3 ACL, that is, subitem 1 of layer 2 ACL and layer 3 ACL combine together, and the rest may be deduced by analogy; if the number of two groups of ACL is not the same, the rest subitem can activate separately.

15.3.7 Monitor and maintenance of ACL

Configure followings in any configuration mode except user mode.

Display time information

show time-range [all | statistic | name *time-range-name*]

Display detail information of ACL

show access-list config { all | *access-list-number* | **name** *access-list-name* }

Display statistic information of ACL

show access-list config statistic

Display runtime information of ACL

show access-list runtime { all | *access-list-number* | **name** *access-list-name* }

Display runtime statistic information of ACL

show access-list runtime statistic

Concrete configuration refers to command line configuration.

15.4 Configuration example of QACL

15.4.1 Use QACL to realize user isolation

15.4.1.1 Brief introduction of isolation

Use user isolation to bind some interface and some IP address. Only the packet with the source IP address being this one can be transmitted, or it will be dropped. This can fix specified user to some interface to realize user management.

There are two types of mode: one is transmitting all ARP packet, the other is not transmitting all ARP packet. In transmitting all ARP mode, after enabling user isolation, all ARP packet can be transmitted. In not transmitting all ARP mode, after enabling user isolation, only after configuring user binding rules (such as ip +port+mac), corresponded ARP packet can be transmitted.

Followings are the configuring examples of two user isolation. Example 1 can use QACL to realize user isolation of all ARP packet; example 2 uses QACL to realize user isolation of not transmitting ARP packet.

15.4.1.2 Example 1

Example 1 uses QACL to realize user isolation of transmitting all ARP packet. This example can realize following function:

- 1) Enable user isolation (prevent all packet and permit ARP packet with VLAN id being 4016);
- 2) Configure Ethernet interface 1 to be uplink interface (permit all packet from uplink interface 1)
- 3) Configure binding rules of three users:

1> ip+port : ip is 192.168.0.1 and port to be Ethernet interface 2

2> ip+port+vid : ip is 192.168.0.2 , port is Ethernet interface 2 and vid is 2

3> ip+port+mac ; ip is 192.168.0.3 , port is Ethernet interface 2 and mac is 00:00:00:00:00:03

The configuration is as following:

(1) Define needed ACL

!Define to deny all packet ACL

QTECH(config)#access-list 200 deny ingress any egress any

!Define to transmit ACL to transmit packet from uplink interface 1

QTECH(config)#access-list 200 permit ingress interface ethernet 0/1 egress any

!Define ACL to transmit packet with VLAN ID being 4016 and from non-uplinkinterface 2

QTECH(config)#access-list 200 permit ingress 4016 interface ethernet 0/2 egress any

!Define ACL to transmit all ARP packet

QTECH(config)#access-list 200 permit arp ingress any egress any

!Define ip+port user to bind ACL with ip being 192.168.0.1 , port being Ethernet interface 2. This ip+port user bound rule can be divided into 2 ACLs :one is ACL to transmit packet with source address being 192.168.0.1, the other is ACL to transmit packet from Ethernet interface 2


```
QTECH(config)#access-list 1 permit 192.168.0.1 0
QTECH(config)#access-list 201 permit ingress interface ethernet 0/0/2 egress any
```

!Define ip+port+vid user to bind ACL with ip being 192.168.0.2 , port being Ethernet interface 2 and vid being 2. This ip+port+vid user bound rule can be divided into 2 ACLs : one is ACL to transmit packet with source address being 192.168.0.2, the other is ACL to transmit packet with vid being 2 from Ethernet interface 2

```
QTECH(config)#access-list 1 permit 192.168.0.2 0
QTECH(config)#access-list 201 permit ingress 2 interface ethernet 0/0/2 egress any
```

!Define ip+port+mac user to bind ACL with ip being 192.168.0.3 , port being Ethernet interface 2 and mac being 00:00:00:00:00:03. This ip+port+mac user bound rule can be divided into 2 ACLs : one is ACL to transmit packet with source address being 192.168.0.3, the other is ACL to transmit packet with mac being 00:00:00:00:00:03 from Ethernet interface 2

```
QTECH(config)#access-list 1 permit 192.168.0.3 0
QTECH(config)#access-list 201 permit ingress 00:00:00:00:00:03 0:0:0:0:0:0 interface ethernet 0/0/2 egress any
```

(2) Activate ACL

```
QTECH(config)#access-group link-group 200
QTECH(config)#access-group ip-group 1 link-group 201
```

15.4.1.3 Example 2

Example 2 uses QACL to realize user isolation of not transmitting all ARP packet. This example can realize following function:

- 1) Enable user isolation (prevent all packet and permit packet with VLAN id being 4016);
- 2) Configure Ethernet interface 1 to be uplink interface (permit all packet from uplink interface 1)
- 3) Configure binding rules of three users:

1> ip+port : ip is 192.168.0.1 and port to be Ethernet interface 2

2> ip+port+vid : ip is 192.168.0.2 , port is Ethernet interface 2 and vid is 2

3> ip+port+mac ; ip is 192.168.0.3 , port is Ethernet interface 2 and mac is 00:00:00:00:00:03

The configuration is as following:

(1) Define needed ACL

```
!Define to deny all packet ACL
QTECH(config)#access-list 200 deny ingress any egress any
!Define to transmit ACL to transmit packet from uplink interface 1
QTECH(config)#access-list 200 permit ingress interface ethernet 0/1 egress any
!Define ACL to transmit packet with VLAN ID being 4016 and from non-uplinkinterface 2
QTECH(config)#access-list 200 permit ingress 4016 interface ethernet 0/2 egress any
!Define ACL to transmit all ARP packet
```

!Define ip+port user to bind ACL with ip being 192.168.0.1 , port being Ethernet interface 2. This ip+port user bound rule can be divided into 2 ACLs :one is ACL to transmit packet with source address being 192.168.0.1, the other is ACL to transmit packet from Ethernet interface 2

```
QTECH(config)#access-list 1 permit 192.168.0.1 0
QTECH(config)#access-list 201 permit ingress interface ethernet 0/0/2 egress any
```

!Define ip+port+vid user to bind ACL with ip being 192.168.0.2 , port being Ethernet interface 2 and vid being 2. This ip+port+vid user bound rule can be divided into 3 ACLs : one is ACL to transmit packet with source address being 192.168.0.2, the other is ACL to transmit packet with vid being 2 from Ethernet interface 2 and the last is ACL transferring ARP packet from Ethernet interface 2 with sending protocol being 192.168.0.1 and vid being 2.

```
QTECH(config)#access-list 1 permit 192.168.0.2 0
QTECH(config)#access-list 201 permit ingress 2 interface ethernet 0/0/2 egress any
```

!Define ip+port+mac user to bind ACL with ip being 192.168.0.3 , port being Ethernet interface 2 and mac being 00:00:00:00:00:03. This ip+port+mac user bound rule can be divided into 2 ACLs : one is ACL to transmit packet with source address being 192.168.0.3, the other is ACL to transmit packet with mac being 00:00:00:00:00:03 from Ethernet interface 2 and the last is ACL transferring ARP packet from Ethernet interface 2 with sending protocol being 192.168.0.1 and mac being 00:00:00:00:00:03.

```
QTECH(config)#access-list 1 permit 192.168.0.3 0
QTECH(config)#access-list 201 permit ingress 00:00:00:00:00:03 0:0:0:0:0:0 interface ethernet 0/0/2 egress
```

any

```
( 2 ) Activate ACL
QTECH(config)#access-group link-group 200
QTECH(config)#access-group ip-group 1 link-group 201
QTECH(config)#access-group user-group 300
```

15.4.2 Use QACL to realize bandwidth control

15.4.2.1 Brief introduction

Bandwidth control means restricting the uplink and downlink speed rate of special flow. Using QACL to realize this function.

15.4.2.2 Configuration example

Use QACL to realize the flow bandwidth control with source mac address being 00:01:02:03:04:05, uplink interface being 1, downlink interface being 8, uplink speed being 3 Mbps and downlink speed being 5 Mbps.

Configuration is as following:

(1) Define needed ACL

!Define ACL transmitting packet with source interface to be Ethernet interface 8 , destination interface to be wthernet interface 1 , source MAC address to be 00:01:02:03:04:05

```
QTECH(config)#accesss-list 200 permit ingress 00:01:02:03:04:05 0:0:0:0:0:0 interface ethernet 0/0/8 egress
egress interface ethernet 0/0/1
```

!Define ACL transmitting packet with source interface being Ethernet interface 1 , destination interface being Ethernet interface 8 , source MAC address being 00:01:02:03:04:05

```
QTECH(config)#accesss-list 201 permit ingress 00:01:02:03:04:05 0:0:0:0:0:0 interface fast-ethenet 1 egress
egress interface ethernet 0/0/8
```

(2) Configure flow monitor of uplink and downlink interface

!Enter interface configuration mode of uplink interface 1

```
QTECH(config)#interface ethernet 0/1
```

!Configure corresponded flow monitor of uplink interface 1

```
QTECH(config-if-ethernet-0/0/1)##rate-limit input link-group 201 3
```

!Enter interface configuration mode of downlink interface 8

```
QTECH(config)#interface ethernet 0/0/8
```

!Configure corresponded flow monitor of downlink interface 8

```
QTECH(config-if-ethernet-0/0/8)##rate-limit input link-group 200 5
```

15.4.3 Use QACL to realize deny all packet expect

15.4.3.1 Brief introduction of deny all packet expect

deny all packet expect is used to drop all packet except needing transmitting. This function can be realized by configuring QACL.

15.4.3.2 Configuration example

Configuring deny all packet expect PPPoE, the protocol number of PPPoE is 0x8863 (decimal is 34915) and 0x8864 (decimal is 34916)

1) Drop all packets

2) Transmit PPPoE packet

Configuration is as following:

```
( 1 ) Define needed ACL
!Configure deny ACL of all packet
QTECH(config)#access-list 200 deny ingress any egress any
!Configure ACL of transmitting PPPoE packet
QTECH(config)#access-list 200 permit 34915 ingress any egress any
QTECH(config)#access-list 200 permit 34916 ingress any egress any
( 2 ) Activate ACL
QTECH(config)#access-group link-group 200
```

15.4.4 Use QACL to prevent virus

15.4.4.1 Brief introduction of QACL anti-virus

Reasonable configured QACL can be used as firewall to prevent virus to be spread through network to reduce the influence to the network. Different virus has different attacking (such as attack different interface). Configure different QACL rules for different virus, which can do effective protection. For all kinds of virus attacking, it can be obtained from professional anti-virus company (Kingsoft Company).

15.4.4.2 Configuration example

Use QACL to prevent bow wave virus

Bow wave virus will attack TCP 135 interface and infect through UDP 69 INTERFACE, TCP 4444 interface. Configuring switch to prevent QACL of this packet can effectively prevent this virus.

The configuring is as following:

```
( 1 ) Define needed ACL
!Configure ACL to prevent TCP packet of interface 135
QTECH(config)#access-list 100 deny tcp any any eq 135
!Configure ACL to prevent UDP packet of interface 69
QTECH(config)#access-list 100 deny udp any any eq 69
!Configure ACL to prevent TCP packet of interface 4444
QTECH(config)#access-list 100 deny tcp any any eq 4444
( 2 ) Activate ACL
QTECH(config)#access-group ip-group 100
```

Chapter 16 QoS Configuration

16.1 Brief introduction of QoS

In traditional packet network, all packets are equal to be handled. Each switch and router handles packet by FIFO to make best effort to send packets to the destination and not to guarantee the transmission delay and delay variation.

With the fast development of computer network, the requirement of network is higher. More and more voice, image and important data which are sensitive about bandwidth, delay and jittering transferred through network, which greatly enrich network service resources and the requirement of quality of service is higher for the network congestion. Now, Ethernet becomes the leading technology in every independent LAN, and many LAN in the form of Ethernet have become a part of internet. With the development of Ethernet technology, Ethernet connecting will become one of main connecting for internet users. To realize end-to-end QoS solution has to consider the service guarantee of Ethernet QoS, which needs Ethernet device applies to Ethernet technology to provide different levels of QoS guarantee for different types of service flow, especially the service flow highly requiring delay and jitter.

16.1.1 Flow

Flow is traffic which means all packets through switch.

16.1.2 Traffic classification

Traffic classification means adopting certain regulation to recognize packet with some features. Classification rule means the filtration regulation configured by the administrator according to managing need which can be simple, such as realizing flow with the feature of different priority according to the ToS field of IP packet head and can be complicated, such as information of integrated link layer (layer 2), network layer (layer 3), transmission layer (layer 4), such as MAC address, IP protocol, source address, destination address or application program interface number to classify packet. General classification is limited in the head of encapsulation packet. Use packet content to be classification standard is singular.

16.1.3 Access control list

To classify flow is to provide service distinctively which must be connected resource distributing. To adopt which kind of flow control is related to the stage it is in and the current load of the network. For example : monitor packet according to the promised average speed rate when the packet is in the network and queue scheduling manage the packet before it is out of the node.

16.1.4 Packet filtration

Packet filtration is to filtrate service flow, such as deny, that is, deny the service flow which is matching the traffic classification and permit other flows to pass. System adopts complicated flow classification to filtrate all kinds of information of service layer 2 packets to deny useless, unreliable, and doubtful service flow to strengthen network security.

Two key points of realizing packet filtration :

Step 1 : Classify ingress flows according to some regulation;

Step 2 : Filtrate distinct flow by denying. Deny is default accessing control.

16.1.5 Flow monitor

In order to serve customers better with the limited network resources, QoS can monitor service flow of specified user in ingress interface, which can adapt to the distributed network resources.

16.1.6 Interface speed limitation

Interface speed limitation is the speed limit based on interface which limits the total speed rate of interface outputting packet.

16.1.7 Redirection

User can re-specify the packet transmission interface based on the need of its own QoS strategies.

16.1.8 Priority mark

Ethernet switch can provide priority mark service for specified packet, which includes : TOS, DSCP, 802.1p. These priority marks can adapt different QoS model and can be defined in these different models.

16.1.9 Choose interface outputting queue for packet

Ethernet switch can choose corresponding outputting queue for specified packets.

16.1.10 Queue scheduler

It adopts queue scheduler to solve the problem of resource contention of many packets when network congestion. There are three queue scheduler matchings : Strict-Priority Queue (PQ), Weighted Round Robin (WRR) and WRR with maximum delay.

16.1.10.1 PQ

PQ(Priority Queueing)is designed for key service application. Key service possesses an important feature, that is, require the precesent service to reduce the response delay when network congestion. Priority queue divides all packets into 4 levels, that is, superior priority, middle priority, normal priority and inferior priority (3, 2, 1, 0), and their priority levels reduce in turn.

When queue schedulerimg, PQ precedently transmits the packets in superior priority according to the priority level. Transmit packet in inferior priority when the superior one is empty. Put the key service in the superior one, and non-key service (such as email)in inferior one to guarantee the packets in superior group can be first transmitted and non-key service can be transmitted in the spare time.

The shortage of PQ is : when there is network congestion, there are more packets in superior group for a long time, the packets in inferior priority will wait longer.

16.1.10.2 WRR

WRR queue scheduler divides a port into 4 or 8 outputting queues (QSW-3900 has 8 queues, that is, 3, 2, 1, 0) and each scheduler is in turn to guarantee the service time for each queue. WRR can configure a weighted value (that is, w3, w2, w1, w0 in turn) which means the percentage of obtaining the resources. For example : There is a port of 100M. Configure its WRR queue scheduler value to be 50, 30, 10, 10 (corresponding w3, w2, w1, w0 in turn) to guarantee the inferior priority queue to gain at least 10Mbit/s bandwidth, to avoid the shartage of PQ queue

scheduler in which packets may not gain the service.

WRR possesses another advantage. The scheduler of many queues is in turn, but the time for service is not fixed—if some queue is free, it will change to the next queue scheduler to make full use of bandwidth resources.

(3) WRR with maximum delay

Compared with WRR, WRR with maximum delay can guarantee the maximum time from packets entering superior queue to leaving it will not beyond the configured maximum delay.

16.1.11 cos-map

This is a relationship of hardware priority queue and priority of IEEE802.1p protocol

System will map between 802.1p protocol priority of packet and hardware queue priority. For each packet, system will map it to specified hardware queue priority according to 802.1p protocol priority of packet.

16.1.12 Flow mirror

Flow mirror means coping specified data packet to monitor interface to detect network and exclude failure.

16.1.13 Statistics based on flow

Statistics based on flow can statistic and analyse the packets customer interested in.

16.1.14 Copy packet to CPU

User can copy specified packet to CPU according to the need of its QoS strategies.

System realizes QoS function according to accessing control list, which includes : flow monitor, interface speed limit, packet redirection, priority mark, queue scheduler, flow mirror, flow statistics and coping packet to CPU.

16.2 QOS Configuration

16.2.1 QoS Configuration list

QOS Configuration includes :

- Packet redirection configuration
- Priority configuration
- Queue-scheduler configuration
- The cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol
- Flow mirror configuration
- Flow statistic configuration
- Define corresponded ACL before configuring QoS.

16.2.2 Packet redirection configuration

Packet redirection configuration is redirecting packet to be transmitted to some egress.

Use following command to configure it.

Configure it in interface configuration mode.

Redirection

```
traffic-redirect { [ ip-group { access-list-number | access-list-name } [ subitem subitem ] ]
[ link-group { access-list-number | access-list-name } [ subitem subitem ] ] } } { interface
interface-num }
```

Cancel redirection

```
no traffic-redirect { [ ip-group { access-list-number | access-list-name } [ subitem subitem ] ]
[ link-group { access-list-number | access-list-name } [ subitem subitem ] ] } }
```

Instruction :

Use this command to redirect the data packet which matched specified accessing list regulations (it is only be effective for permit rules of accessing list).

Details of this command refers to corresponded command.

16.2.3 Priority configuration

Traffic priority configuration is the strategy of remark priority for matching packet in ACL, and the marked priority can be filled in the domain which reflect priority in packet head.

Use following command to configure priority mark configuration.

Configure it in global configuration mode.

Mark packet priority

```
traffic-priority { [ ip-group { access-list-number | access-list-name } [ subitem subitem ] ]
[ link-group { access-list-number | access-list-name } [ subitem subitem ] ] } } { [ dscp dscp-value |
precedence { pre-value | from-cos } ] [ cos { pre-value | from-ipprec } ] [ local-precedence
pre-value ] }
```

Cancel packet priority configuration

```
no traffic-priority { [ ip-group { access-list-number | access-list-name } [ subitem subitem ] ]
[ link-group { access-list-number | access-list-name } [ subitem subitem ] ] } }
```

System will mark IP priority (precedence specified value of traffic-priority command), DSCP(dscp specified value of traffic-priority command), 802.1p priority(that is cos value of traffic-priority command). User can mark different priority for packet according to real QoS strategy. Switch can locate packet to interface outputting queue according to the 802.1p priority and also can locate packet to corresponding outputting queue according to the specified local priority in traffic-priority command (local-precedence specified value). If both 802.1p priority and local priority are configured, 802.1p priority will be precedent to use.

Details of this command refers to corresponded command.

16.2.4 Queue-scheduler configuration

When network congestion, it must use queue-scheduler to solve the problem of resource competition.

Use following command to configure queue-scheduler.

Configure it in global configuration mode.

Configure queue-scheduler

```
queue-scheduler { strict-priority | wrr queue1-weight queue2-weight queue3-weight queue4 }
```

Disable queue-scheduler

```
no queue-scheduler
```

System supports two types of queue-scheduler mode : Strict-Priority Queue, and Weighted Round Robin (WRR).

By default, switch uses Strict-Priority Queue.

The detailed command refers to the corresponding command line reference.

16.2.5 The cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol

There are 4 hardware priority queues which are from 0 to 3, of which 3 is the

The default mapping is the mapping defined by 802.1p :

802.1p: 0 1 2 3 4 5 6 7

packed-priority: 0 0 1 1 2 2 3 3

Use queue-scheduler cos-map command to configure 4 cos-map relationship of hardware priority queue and 8 priority of IEEE802.1p protocol

Use following command in global configuration mode.

queue-scheduler cos-map [*queue-number*] [*packed-priority*]

Use following command to display the priority cos-map.

show queue-scheduler cos-map

For example :

! Configure packed-priority 1 to mapped priority 6 of IEEE 802.1p

QTECH(config)#queue-scheduler cos-map 1 6

16.2.6 Flow mirror configuration

Flow mirror is copying the service flow which matches ACL rules to specified monitor interface to analyse and monitor packet.

Use following command to configure flow mirror.

Configure it in interface configuration mode.

Flow mirror configuration

mirrored-to { [**ip-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] } [**interface interface-num**]

Cancel flow mirror configuration

no mirrored-to { [**ip-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] }

Details of this command refers to corresponded command.

16.2.7 Flow statistic configuration

Flow statistic configuration is used to statistic specified service flow packet.

Use following command to configure it.

Configure it in global configuration mode.

Flow statistic configuration

traffic-statistic { [**ip-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] }

Clear statistic information

clear traffic-statistic { **all** | [**ip-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] }

Cancel flow statistic configuration

no traffic-statistic { [**ip-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] [**link-group** { *access-list-number* | *access-list-name* } [**subitem subitem**]] }

If reconfiguring flow statistics, the corresponded information will be cleared.

Details of this command refers to corresponded command.

16.3 Monitor and maintenance of QoS

Configure it in corresponded configuration mode. Show command can be used in any configured mode except user mode.

Display all QoS information :

show qos-info all

Display all QoS statistic information

show qos-info statistic

Display flow mirror configuration

show qos-info mirrored-to

Display queue scheduler and parameter

show queue-scheduler

Display the cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol

show queue-scheduler cos-map

Display QOS configuration of all interface

show qos-interface [interface-num] all

Display parameter configuration of flow limit

show qos-interface [interface-num] rate-limit

Display line limit configuration

show qos-interface [interface-num] line-rate

Display QOS statistic information of all interface

show qos-interface statistic

Display priority configuration

show qos-info traffic-priority

Display redirection configuration

show qos-info traffic-redirect

Display flow statistic configuration

show qos-info traffic-statistic

Display configuration of copying to CPU.

show qos-info traffic-copy-to-cpu

Details of this command refers to corresponded command.

Chapter 17 STP Configuration

17.1 Brief introduction of STP Configuration

STP(Spanning Tree Protocol) is a part of IEEE 802.1D network bridge. The realization of standard STP can eliminate network broadcast storm caused by network circle connection and the circle connection caused by misplaying and accident, and it also can provide the possibility of network backup connection.

STP protocol with IEEE 802.1D standard provides network dynamic redundancy transferring mechanism and prevents circle connection in bridge network. It determines which interface of the network bridge can transmit data packet. After executing STP matching, switch in the LAN will form a STP dynamic topology which prevents the loop existing between any two working station to prevent broadcast storm in LAN. At the same time, STP matching is responsible to detect the change of physical topology to establish new spanning tree after the changes of topology. For example : when there is a break in the switch or a channel, it can provide certain error tolerance to re-configure a new STP topology.

17.1.1 Introduction to STP

17.1.1.1 Why STP?

The Spanning Tree Protocol (STP) was established based on the 802.1D standard of IEEE to eliminate physical loops at the data link layer in a local area network (LAN). Devices running this protocol detect loops in the network by exchanging information with one another and eliminate loops by selectively blocking certain ports until the loop structure is pruned into a loop-free network structure. This avoids proliferation and infinite recycling of packets that would occur in a loop network and prevents deterioration of the packet processing capability of network devices cause by duplicate packets received.

17.1.1.2 Protocol Packets of STP

STP uses bridge protocol data units (BPDUs), also known as configuration messages, as its protocol packets.

STP identifies the network topology by transmitting BPDUs between STP compliant network devices. BPDUs contain sufficient information for the network devices to complete the spanning tree computing.

In STP, BPDUs come in two types :

- Configuration BPDUs, used to maintain the spanning tree topology.
- Topology change notification (TCN) BPDUs, used to notify concerned devices of network topology changes, if any.

17.1.1.3 Basic concepts in STP

1) Root bridge

A tree network must have a root; hence the concept of “root bridge” has been introduced in STP.

There is one and only one root bridge in the entire network, and the root bridge can change alone with changes of the network topology. Therefore, the root bridge is not fixed.

Upon network convergence, the root bridge generates and sends out at a certain interval a BPDU and other devices just forward this BPDU. This mechanism ensures the topological stability.

2) Root port

On a non-root bridge device, the root port is the port with the lowest path cost to the root bridge. The root port is used for communicating with the root bridge. A non-root-bridge device has one and only one root port. The root bridge has no root port.

3) Designated bridge and designated port

The following table describes a designated bridge and a designated port.

Description of designated bridge and designated port

Classification	Designated bridge	Designated port
For a device	The device directly connected with this device and responsible for forwarding BPDUs	The port through which the designated bridge forwards BPDUs to this device
For a LAN	The device responsible for forwarding BPDUs to this LAN segment	The port through which the designated forwards BPDUs to this LAN segment

Figure 1 shows designated bridges and designated ports. In the figure, AP1 and AP2, BP1 and BP2, and CP1 and CP2 are ports on Device A, Device B, and Device C respectively.

- If Device A forwards BPDUs to Device B through AP1, the designated bridge for Device B is Device A, and the designated port is the port AP1 on Device A.
- Two devices are connected to the LAN : Device B and Device C. If Device B forwards BPDUs to the LAN, the designated bridge for the LAN is Device B, and the designated port is the port BP2 on Device B.

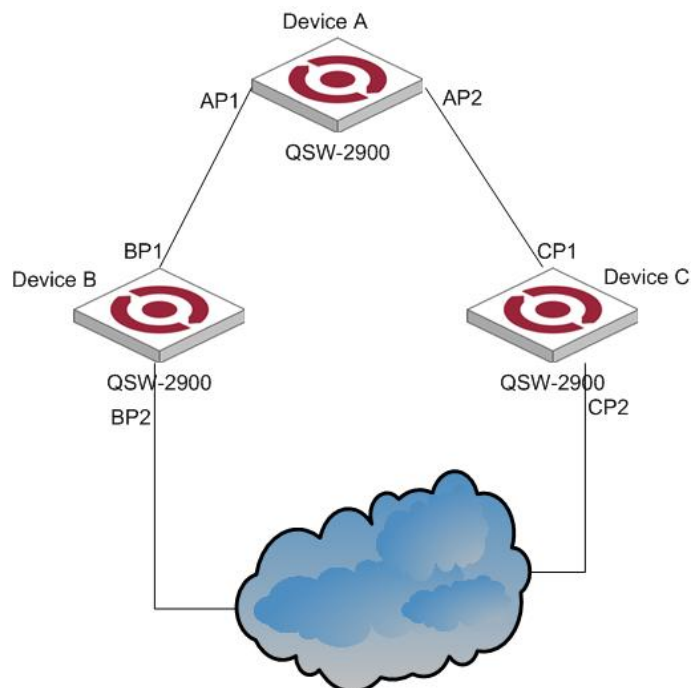


Figure 1 A schematic diagram of designated bridges and designated ports

Note :

All the ports on the root bridge are designated ports.

17.1.1.4 How STP works

STP identifies the network topology by transmitting configuration BPDUs between network devices. Configuration BPDUs contain sufficient information for network devices to complete the spanning tree computing. Important fields in a configuration BPDU include :

- Root bridge ID : consisting of root bridge priority and MAC address.
- Root path cost : the cost of the shortest path to the root bridge.
- Designated bridge ID : designated bridge priority plus MAC address.
- Designated port ID, designated port priority plus port name.
- Message age : age of the configuration BPDU
- Max age : maximum age of the configuration BPDU.
- Hello time : configuration BPDU interval.
- Forward delay : forward delay of the port.

 Note :

For the convenience of description, the description and examples below involve only four parts of a configuration BPDU :

- Root bridge ID (in the form of device priority)
- Root path cost
- Designated bridge ID (in the form of device priority)
- Designated port ID (in the form of port name)

1) Specific computing process of the STP algorithm

- Initial state

Upon initialization of a device, each port generates a BPDU with itself as the root, in which the root path cost is 0, designated bridge ID is the device ID, and the designated port is the local port.

- Selection of the optimum configuration BPDU

Each device sends out its configuration BPDU and receives configuration BPDUs from other devices. The process of selecting the optimum configuration BPDU is as follows :

Table 2 Selection of the optimum configuration BPDU

Step	Description
1	<p>Upon receiving a configuration BPDU on a port, the device performs the following processing :</p> <ul style="list-style-type: none"> • If the received configuration BPDU has a lower priority than that of the configuration BPDU generated by the port, the device will discard the received configuration BPDU without doing any processing on the configuration BPDU of this port. • If the received configuration BPDU has a higher priority than that of the configuration BPDU generated by the port, the device will replace the content of the configuration BPDU generated by the port with the content of the received configuration BPDU.
2	The device compares the configuration BPDUs of all the ports and chooses the optimum configuration BPDU.

 Note :

Principle for configuration BPDU comparison :

- The configuration BPDU that has the lowest root bridge ID has the highest priority.
- If all the configuration BPDUs have the same root bridge ID, they will be compared for their root path costs. If the root path cost in a configuration BPDU plus the path cost corresponding to this port is S, the configuration BPDU with the smallest S value has the highest priority.
- If all configuration BPDU have the same root path cost, they will be compared for their designated bridge IDs, then their designated port IDs, and then the IDs of the ports on which they are received. The smaller the ID, the higher message priority.

- Selection of the root bridge

At network initialization, each STP-compliant device on the network assumes itself to be the root bridge, with the root bridge ID being its own device ID. By exchanging configuration BPDUs, the devices compare one another's root bridge ID. The device with the smallest root bridge ID is elected as the root bridge.

- Selection of the root port and designated ports

The process of selecting the root port and designated ports is as follows :

Table 3 Selection of the root port and designated ports

Step	Description
1	A non-root-ridge device regards the port on which it received the optimum configuration BPDU as the root port.
2	<p>Based on the configuration BPDU and the path cost of the root port, the device calculates a designated port configuration BPDU for each of the rest ports.</p> <ul style="list-style-type: none"> • The root bridge ID is replaced with that of the configuration BPDU of the root port. • The root path cost is replaced with that of the configuration BPDU of the root port plus the path cost corresponding to the root port. • The designated bridge ID is replaced with the ID of this device. • The designated port ID is replaced with the ID of this port.

Step	Description
3	<p>The device compares the calculated configuration BPDU with the configuration BPDU on the port of which the port role is to be defined, and does different things according to the comparison result :</p> <ul style="list-style-type: none"> • If the calculated configuration BPDU is superior, the device will consider this port as the designated port, and the configuration BPDU on the port will be replaced with the calculated configuration BPDU, which will be sent out periodically. • If the configuration BPDU on the port is superior, the device will block this port without updating its configuration BPDU, so that the port will only receive BPDUs, but not send any, and will not forward data.

 Note :

When the network topology is stable, only the root port and designated ports forward traffic, while other ports are all in the blocked state – they only receive STP packets but do not forward user traffic.

Once the root bridge, the root port on each non-root bridge and designated ports have been successfully elected, the entire tree-shaped topology has been constructed.

The following is an example of how the STP algorithm works. The specific network diagram is shown in Figure 2. In the feature, the priority of Device A is 0, the priority of Device B is 1, the priority of Device C is 2, and the path costs of these links are 5, 10 and 4 respectively.

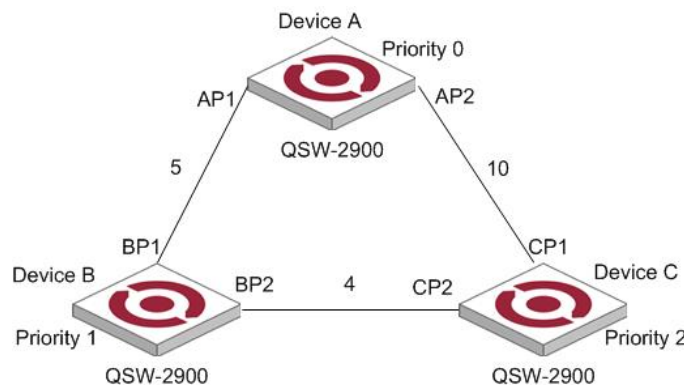


Figure 2 Network diagram for STP algorithm

- Initial state of each device

The following table shows the initial state of each device.

Table 4 Initial state of each device

Device	Port name	BPDU of port
Device A	AP1	{0, 0, 0, AP1}
	AP2	{0, 0, 0, AP2}
Device B	BP1	{1, 0, 1, BP1}
	BP2	{1, 0, 1, BP2}
Device C	CP1	{2, 0, 2, CP1}
	CP2	{2, 0, 2, CP2}

- Comparison process and result on each device

The following table shows the comparison process and result on each device.

Table 5 Comparison process and result on each device

Device	Comparison process	BPDU of port after comparison
Device A	<ul style="list-style-type: none"> Port AP1 receives the configuration BPDU of Device B {1, 0, 1, BP1}. Device A finds that the configuration BPDU of the local port {0, 0, 0, AP1} is superior to the configuration received message, and discards the received configuration BPDU. Port AP2 receives the configuration BPDU of Device C {2, 0, 2, CP1}. Device A finds that the BPDU of the local port {0, 0, 0, AP2} is superior to the received configuration BPDU, and discards the received configuration BPDU. Device A finds that both the root bridge and designated bridge in the configuration BPDUs of all its ports are Device A itself, so it assumes itself to be the root bridge. In this case, it does not make any change to the configuration BPDU of each port, and starts sending out configuration BPDUs periodically. 	AP1 : {0, 0, 0, AP1} AP2 : {0, 0, 0, AP2}
Device B	<ul style="list-style-type: none"> Port BP1 receives the configuration BPDU of Device A {0, 0, 0, AP1}. Device B finds that the received configuration BPDU is superior to the configuration BPDU of the local port {1, 0, 1, BP1}, and updates the configuration BPDU of BP1. Port BP2 receives the configuration BPDU of Device C {2, 0, 2, CP2}. Device B finds that the configuration BPDU of the local port {1, 0, 1, BP2} is superior to the received configuration BPDU, and discards the received configuration BPDU. 	BP1 : {0, 0, 0, AP1} BP2 : {1, 0, 1, BP2}
	<ul style="list-style-type: none"> Device B compares the configuration BPDUs of all its ports, and determines that the configuration BPDU of BP1 is the optimum configuration BPDU. Then, it uses BP1 as the root port, the configuration BPDUs of which will not be changed. Based on the configuration BPDU of BP1 and the path cost of the root port (5), Device B calculates a designated port configuration BPDU for BP2 {0, 5, 1, BP2}. Device B compares the computed configuration BPDU {0, 5, 1, BP2} with the configuration BPDU of BP2. If the computed BPDU is superior, BP2 will act as the designated port, and the configuration BPDU on this port will be replaced with the computed configuration BPDU, which will be sent out periodically. 	Root port BP1 : {0, 0, 0, AP1} Designated port BP2 : {0, 5, 1, BP2}
Device C	<ul style="list-style-type: none"> Port CP1 receives the configuration BPDU of Device A {0, 0, 0, AP2}. Device C finds that the received configuration BPDU is superior to the configuration BPDU of the local port {2, 0, 2, CP1}, and updates the configuration BPDU of CP1. 	CP1 : {0, 0, 0, AP2} CP2 : {1, 0, 1, BP2}

- Port CP2 receives the configuration BPDU of port BP2 of Device B {1, 0, 1, BP2} before the message was updated. Device C finds that the received configuration BPDU is superior to the configuration BPDU of the local port {2, 0, 2, CP2}, and updates

Device	Comparison process	BPDU of port after comparison
	the configuration BPDU of CP2.	
	<p>By comparison :</p> <ul style="list-style-type: none"> The configuration BPDUs of CP1 is elected as the optimum configuration BPDU, so CP1 is identified as the root port, the configuration BPDUs of which will not be changed. Device C compares the computed designated port configuration BPDU {0, 10, 2, CP2} with the configuration BPDU of CP2, and CP2 becomes the designated port, and the configuration BPDU of this port will be replaced with the computed configuration BPDU. 	<p>Root port CP1 : {0, 0, 0, AP2}</p> <p>Designated port CP2 : {0, 10, 2, CP2}</p>
	<ul style="list-style-type: none"> Next, port CP2 receives the updated configuration BPDU of Device B {0, 5, 1, BP2}. Because the received configuration BPDU is superior to its old one, Device C launches a BPDU update process. At the same time, port CP1 receives configuration BPDUs periodically from Device A. Device C does not launch an update process after comparison. 	<p>CP1 : {0, 0, 0, AP2}</p> <p>CP2 : {0, 5, 1, BP2}</p>
	<p>By comparison :</p> <ul style="list-style-type: none"> Because the root path cost of CP2 (9) (root path cost of the BPDU (5) + path cost corresponding to CP2 (4)) is smaller than the root path cost of CP1 (10) (root path cost of the BPDU (0) + path cost corresponding to CP2 (10)), the BPDU of CP2 is elected as the optimum BPDU, and CP2 is elected as the root port, the messages of which will not be changed. After comparison between the configuration BPDU of CP1 and the computed designated port configuration BPDU, port CP1 is blocked, with the configuration BPDU of the port remaining unchanged, and the port will not receive data from Device A until a spanning tree computing process is triggered by a new condition, for example, the link from Device B to Device C becomes down. 	<p>Blocked port CP2 : {0, 0, 0, AP2}</p> <p>Root port CP2 : {0, 5, 1, BP2}</p>

After the comparison processes described in the table above, a spanning tree with Device A as the root bridge is stabilized, as shown in Figure 3.

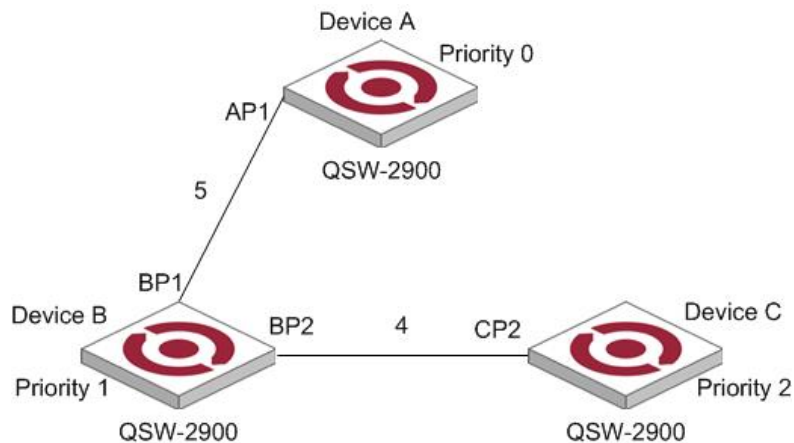


Figure 3 The final computed spanning tree

 Note :

To facilitate description, the spanning tree computing process in this example is simplified, while the actual process is more complicated.

2) The BPDU forwarding mechanism in STP

- Upon network initiation, every switch regards itself as the root bridge, generates configuration BPDUs with itself as the root, and sends the configuration BPDUs at a regular interval of hello time.
- If it is the root port that received the configuration BPDU and the received configuration BPDU is superior to the configuration BPDU of the port, the device will increase message age carried in the configuration BPDU by a certain rule and start a timer to time the configuration BPDU while it sends out this configuration BPDU through the designated port.
- If the configuration BPDU received on the designated port has a lower priority than the configuration BPDU of the local port, the port will immediately sends out its better configuration BPDU in response.
- If a path becomes faulty, the root port on this path will no longer receive new configuration BPDUs and the old configuration BPDUs will be discarded due to timeout. In this case, the device will generate a configuration BPDU with itself as the root and sends out the BPDU. This triggers a new spanning tree computing process so that a new path is established to restore the network connectivity.

However, the newly computed configuration BPDU will not be propagated throughout the network immediately, so the old root ports and designated ports that have not detected the topology change continue forwarding data through the old path. If the new root port and designated port begin to forward data as soon as they are elected, a temporary loop may occur. For this reason, STP uses a state transition mechanism. Namely, a newly elected root port or designated port requires twice the forward delay time before transitioning to the forwarding state, when the new configuration BPDU has been propagated throughout the network.

17.1.2 Introduction to MSTP

17.1.2.1 Why MSTP

1) Disadvantages of STP and RSTP

STP does not support rapid state transition of ports. A newly elected root port or designated port must wait twice the forward delay time before transitioning to the forwarding state, even if it is a port on a point-to-point link or it is an edge port, which directly connects to a user terminal rather than to another device or a shared LAN segment.

The rapid spanning tree protocol (RSTP) is an optimized version of STP. RSTP allows a newly elected root port or designated port to enter the forwarding state much quicker under certain conditions than in STP. As a result, it takes a shorter time for the network to reach the final topology stability.

 Note :

- In RSTP, a newly elected root port can enter the forwarding state rapidly if this condition is met : The old root port on the device has stopped forwarding data and the upstream designated port has started forwarding data.
- In RSTP, a newly elected designated port can enter the forwarding state rapidly if this condition is met : The designated port is an edge port or a port connected with a point-to-point link. If the designated port is an edge port, it can enter the forwarding state directly; if the designated port is connected with a point-to-point link, it can enter the forwarding state immediately after the device undergoes handshake with the downstream device and gets a response.

Although RSTP support rapid network convergence, it has the same drawback as STP does : All bridges within a LAN share the same spanning tree, so redundant links cannot be blocked based on VLANs, and the packets of all VLANs are forwarded along the same spanning tree.

2) Features of MSTP

The multiple spanning tree protocol (MSTP) overcomes the shortcomings of STP and RSTP. In addition to

support for rapid network convergence, it also allows data flows of different VLANs to be forwarded along their own paths, thus providing a better load sharing mechanism for redundant links. For description about VLANs, refer to *VLAN Configuration* in the *Access Volume*.

MSTP features the following :

- MSTP supports mapping VLANs to MST instances by means of a VLAN-to-instance mapping table.
- MSTP divides a switched network into multiple regions, each containing multiple spanning trees that are independent of one another.
- MSTP prunes loop networks into a loop-free tree, thus avoiding proliferation and endless recycling of packets in a loop network. In addition, it provides multiple redundant paths for data forwarding, thus supporting load balancing of VLAN data in the data forwarding process.
- MSTP is compatible with STP and RSTP.

17.1.2.2 Some concepts in MSTP

As shown in Figure 4, there are four multiple spanning tree (MST) regions, each made up of four switches running MSTP. In light with the diagram, the following paragraphs will present some concepts of MSTP.

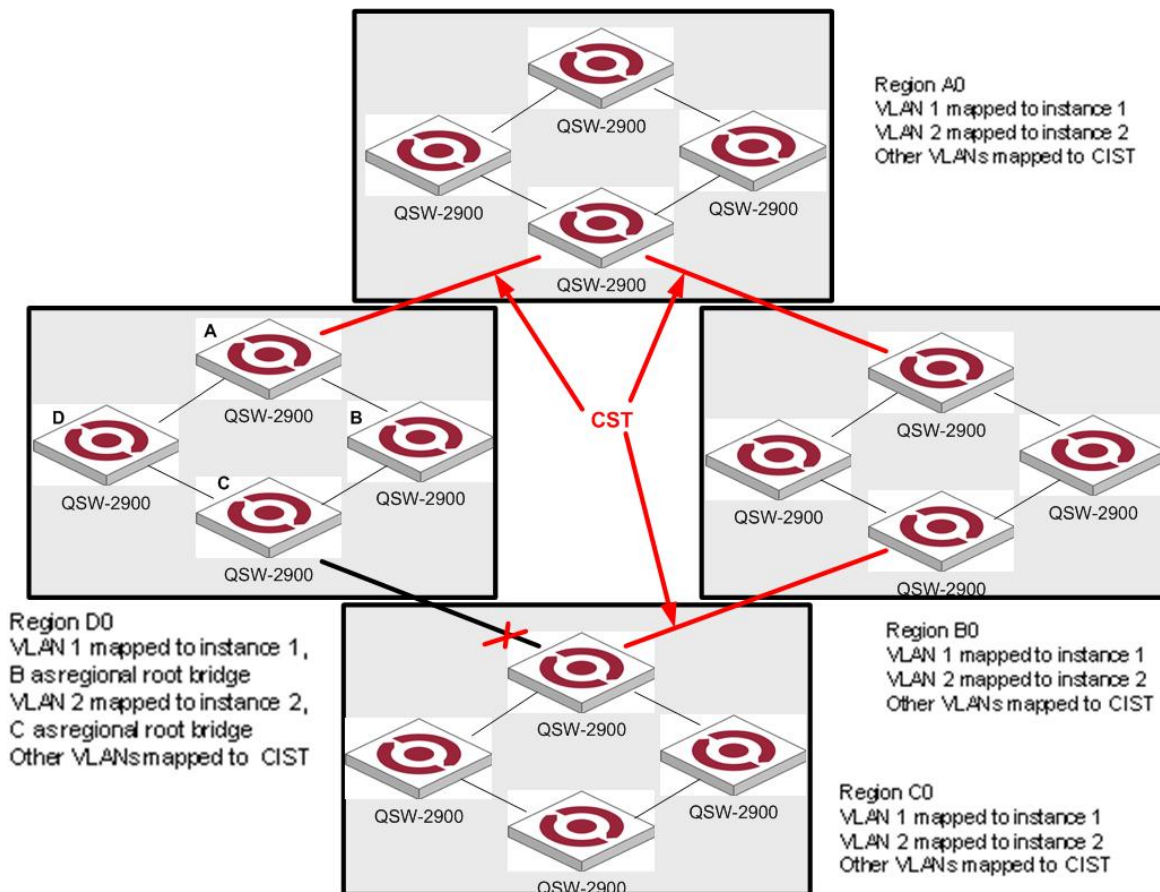


Figure 4 Basic concepts in MSTP

1) MST region

An MST region is composed of multiple devices in a switched network and network segments among them. These devices have the following characteristics :

- All are MSTP-enabled,
- They have the same region name,
- They have the same VLAN-to-instance mapping configuration,
- They have the same MSTP revision level configuration, and
- They are physically linked with one another.

In region A0 in Figure 4, for example, all the device have the same MST region configuration : the same region name, the same VLAN-to-instance mapping (VLAN1 is mapped to MST instance 1, VLAN2 to MST instance 2, and the rest to the command and internal spanning tree (CIST). CIST refers to MST instance 0), and the same

MSTP revision level (not shown in the figure).

Multiple MST regions can exist in a switched network. You can use an MSTP command to group multiple devices to the same MST region.

2) VLAN-to-instance mapping table

As an attribute of an MST region, the VLAN-to-instance mapping table describes the mapping relationships between VLANs and MST instances. In Figure 4, for example, the VLAN-to-instance mapping table of region A0 describes that the same region name, the same VLAN-to-instance mapping (VLAN1 is mapped to MST instance 1, VLAN2 to MST instance 2, and the rest to CIST).

3) IST

Internal spanning tree (IST) is a spanning tree that runs in an MST region, with the instance number of 0. ISTs in all MST regions and the common spanning tree (CST) jointly constitute the common and internal spanning tree (CIST) of the entire network. An IST is a section of the CIST in an MST region. In Figure 4, for example, the CIST has a section in each MST region, and this section is the IST in each MST region.

4) CST

The CST is a single spanning tree that connects all MST regions in a switched network. If you regard each MST region as a “device”, the CST is a spanning tree computed by these devices through MSTP. For example, the red lines in Figure 4 describe the CST.

5) CIST

Jointly constituted by ISTs and the CST, the CIST is a single spanning tree that connects all devices in a switched network. In Figure 4, for example, the ISTs in all MST regions plus the inter-region CST constitute the CIST of the entire network.

6) MSTI

Multiple spanning trees can be generated in an MST region through MSTP, one spanning tree being independent of another. Each spanning tree is referred to as a multiple spanning tree instance (MSTI). In Figure 4, for example, multiple spanning trees can exist in each MST region, each spanning tree corresponding to a VLAN. These spanning trees are called MSTIs.

7) Regional root bridge

The root bridge of the IST or an MSTI within an MST region is the regional root bridge of the MST or that MSTI. Based on the topology, different spanning trees in an MST region may have different regional roots. For example, in region D0 in Figure 4, the regional root of instance 1 is device B, while that of instance 2 is device C.

8) Common root bridge

The root bridge of the CIST is the common root bridge. In Figure 4, for example, the common root bridge is a device in region A0.

9) Boundary port

A boundary port is a port that connects an MST region to another MST configuration, or to a single spanning-tree region running STP, or to a single spanning-tree region running RSTP.

During MSTP computing, a boundary port assumes the same role on the CIST and on MST instances. Namely, if a boundary port is master port on the CIST, it is also the master port on all MST instances within this region. In Figure 4, for example, if a device in region A0 is interconnected with the first port of a device in region D0 and the common root bridge of the entire switched network is located in region A0, the first port of that device in region D0 is the boundary port of region D0.

Note :

Currently, the device is not capable of recognizing boundary ports. When the device interworks with a third party's device that supports boundary port recognition, the third party's device may malfunction in recognizing a boundary port.

10) Roles of ports

In the MSTP computing process, port roles include designated port, root port, master port, alternate port, backup port, and so on.

- Root port : a port responsible for forwarding data to the root bridge.
- Designated port : a port responsible for forwarding data to the downstream network segment or device.
- Master port : A port on the shortest path from the entire region to the common root bridge, connecting the MST region to the common root bridge.
- Alternate port : The standby port for the root port or master port. If a root port or master port is blocked, the alternate port becomes the new root port or master port.

- Backup port : The backup port of designated ports. When a designated port is blocked, the backup port becomes a new designated port and starts forward data without delay. When a loop occurs while two ports of the same MSTP device are interconnected, the device will block either of the two ports, and the backup port is that port to be blocked.
- A port can assume different roles in different MST instances.

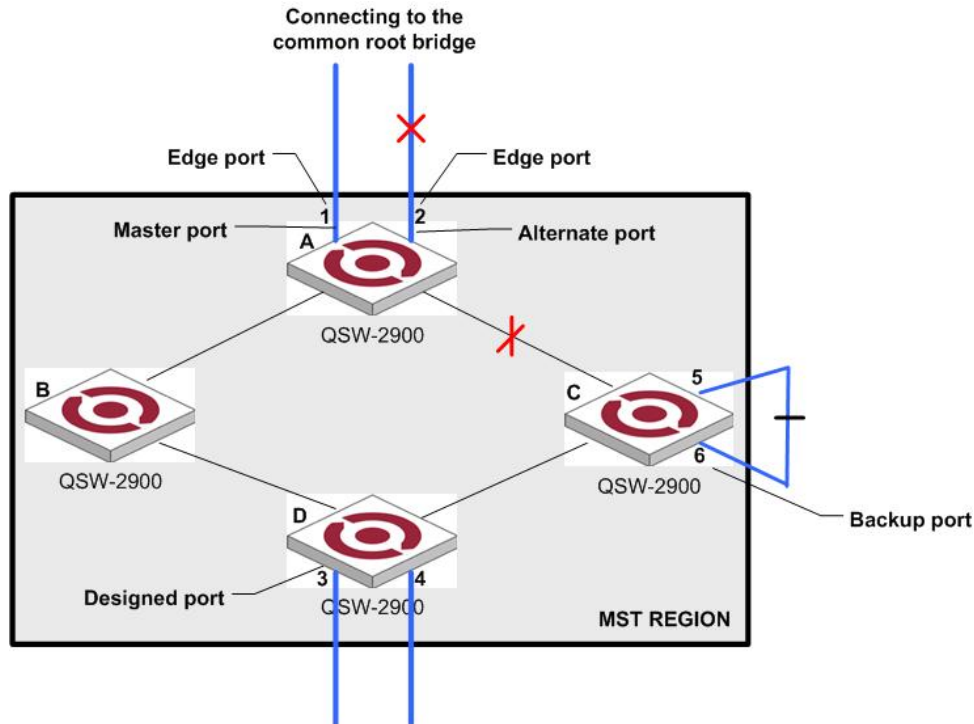


Figure 5 Port roles

Figure 5 helps understand these concepts. Where,

- Devices A, B, C, and D constitute an MST region.
- Port 1 and port 2 of device A connect to the common root bridge.
- Port 5 and port 6 of device C form a loop.
- Port 3 and port 4 of device D connect downstream to other MST regions.

11) Port states

In MSTP, port states fall into the following tree :

- Forwarding : the port learns MAC addresses and forwards user traffic;
- Learning : the port learns MAC addresses but does not forwards user traffic;
- Discarding : the port neither learns MAC addresses nor forwards user traffic.

Note :

When in different MST instances, a port can be in different states.

A port state is not exclusively associated with a port role. Table 6 lists the port state(s) supported by each port role (“√” indicates that the port supports this state, while “-“ indicates that the port does not support this state).

Table 6 Ports states supported by different port roles

Role State	Root port/Master port	Designated port	Alternate port	Backup port
Forwarding	√	√	—	—
Learning	√	√	—	—
Discarding	√	√	√	√

17.1.2.3 How MSTP works

MSTP divides an entire Layer 2 network into multiple MST regions, which are interconnected by a computed CST. Inside an MST region, multiple spanning trees are generated through computing, each spanning tree called an MST instance. Among these MST instances, instance 0 is the IST, while all the others are MSTIs. Similar to STP, MSTP uses configuration BPDUs to compute spanning trees. The only difference between the two protocols being in that what is carried in an MSTP BPDU is the MSTP configuration on the device from which this BPDU is sent.

1) CIST computing

By comparison of “configuration BPDUs”, the device with the highest priority is elected as the root bridge of the CIST. MSTP generates an IST within each MST region through computing, and, at the same time, MSTP regards each MST region as a single device and generates a CST among these MST regions through computing. The CST and ISTs constitute the CIST of the entire network.

2) MSTI computing

Within an MST region, MSTP generates different MSTIs for different VLANs based on the VLAN-to-instance mappings.

MSTP performs a separate computing process, which is similar to spanning tree computing in STP, for each spanning tree.

In MSTP, a VLAN packet is forwarded along the following paths :

- Within an MST region, the packet is forwarded along the corresponding MSTI.
- Between two MST regions, the packet is forwarded along the CST.

17.1.2.4 Implementation of MSTP on devices

MSTP is compatible with STP and RSTP. STP and RSTP protocol packets can be recognized by devices running MSTP and used for spanning tree computing.

In addition to basic MSTP functions, many management-facilitating special functions are provided, as follows :

- Root bridge hold
- Root bridge backup
- Root guard
- BPDU guard
- Loop guard
- Support for hot swapping of interface cards and active/standby changeover.

17.1.3 Protocols and Standards

MSTP is documented in :

- IEEE 802.1D : Spanning Tree Protocol
- IEEE 802.1w : Rapid Spanning Tree Protocol
- IEEE 802.1s : Multiple Spanning Tree Protocol

17.2 STP Configuration

17.2.1 STP Configuration list

The configuration can be effective only after STP enables. Configure related parameter of devices or Ethernet interface before enabling STP and these configurations will be saved after disabling STP. And the parameter will be effective after re-enabling STP. STP configuration list is as following :

- Enable/disable interface STP
- Configure STP mode
- Configure STP priority
- Configure Forward Delay
- Configure Hello Time
- Configure Max Age
- Configure path cost of specified interfaces
- Configure STP priority of specified port
- Configure interface to force to send rstp packet
- Configure link type of specified interface
- Configure the current port as an edge port
- Configure the speed limit of sending BPDU of specified interface
- STP monitor and maintenance

17.2.2 Enable/disable STP

Configure it in global configuration mode :

Enable/disable STP of the devices

spanning-tree

Disable STP of the devices

no spanning-tree

By default, switch STP disables.

For example :

! Enable STP

QTECH(config)#spanning-tree

17.2.3 Enable/disable interface STP

Disable STP of specified interface to make the interface not to attend STP calculating. Use following command in interface configuration mode :

Enable STP on specified interface

spanning-tree

Disable STP on specified interface

no spanning-tree

By default, interface STP enables.

For example :

! Disable STP on Ethernet 01

QTECH(config-if-ethernet-0/0/1)#no spanning-tree

17.2.4 Configure STP priority

Configure STP priority when STP enables, and the inferior priority of the switch can be the root bridge. Use following command in global configuration mode :

Configure STP priority

spanning-tree priority *bridge-priority*

Restore default STP priority

no spanning-tree priority

For example :

! Configure the priority of the switch in spanning tree to 30000

QTECH(config)#spanning-tree priority 30000



Caution : If the priorities of all network bridge in switching network are the same, choose the one with the smallest MAC address to be the root. If STP enables, configuring network bridge may cause the re-accounting of the STP. By default, the network bridge priority is 32768 and ranges from 0 to 65535.

17.2.5 Configure switch Forward Delay

When this switch is the root bridge, port state transition period is the Forward Delay time, which is determined by the diameter of the switched network. The longer the diameter is, the longer the time is. Configure it in global configuration mode :

Configure Forward Delay

spanning-tree forward-time *seconds*

Restore default Forward Delay

no spanning-tree forward-time

For example :

! Configure forward delay to 20 seconds

QTECH(config)#spanning-tree forward-time 20



Caution : If Forward Delay is configured too small, temporary redundancy will be caused; if Forward Delay is configured too large, network will not be restored linking for a long time. Forward Delay ranges from 4 to 30 seconds. The default forward delay time, 15 seconds is suggested to use. Forward Delay \geq Hello Time + 2.

17.2.6 Configure Hello Time

Suitable Hello Time can guarantee network bridge noticing link failure in time without occupying too much resources. Configure it in global configuration mode :

Configure Hello Time

spanning-tree hello-time *seconds*

Restore default Hello Time

no spanning-tree hello-time

For example :

! Configure Hello Time to 5 seconds

QTECH(config)#spanning-tree hello-time 5



Caution : Too large Hello Time may cause link failure thought by network bridge for losing packets

of the link to restart accounting STP; too smaller Hello Time may cause network bridge frequently to send configuration packet to strengthen the load of network and CPU. Hello Time ranges from 1 to 10 seconds. It is suggested to use the default time of 2 seconds. $\text{Hello Time} \leq \text{Forward Delay} - 2$

17.2.7 Configure Max Age

Max Age is used to judge whether the packet is outdate. User can configure it according to the real situation of the network in global configuration mode :

Configure Max Age

spanning-tree max-age *seconds*

Restore the default Max Age

no spanning-tree max-age

For example :

! Configure the Max Age to 10 seconds

```
QTECH(config)#spanning-tree max-age 10
```



Caution : Max Age is used to configure the longest aging interval of STP. Lose packet when overtiming. The STP will be frequently accounts and take crowded network to be link fault, if the value is too small. If the value is too large, the link fault cannot be known timely. Max Age is determined by diameter of network, and the default time of 20 seconds is suggested. $2 * (\text{Hello Time} + 1) \leq \text{Max Age} \leq 2 * (\text{ForwardDelay} - 1)$

17.2.8 Configure path cost of specified interfaces

Configure interface STP path cost and choose the path with the smallest path cost to be the effective path. The path cost is related to the link speed rate. The larger the speed rate is, the less the cost is. STP can auto-detect the link speed rate of current interface and converse it to be the cost. Configure it in interface configuration mode :

Configure path cost of specified interface

spanning-tree cost *cost*

Restore the default path cost of specified interface

no spanning-tree cost

Configure path cost will cause the re-accounting of the STP. Interface path cost ranges from 1 to 65535. It is suggested to use the default cost to make STP calculate the path cost of the current interface. By default, the path cost is determined by the current speed.

In IEEE 802.1D, the default path cost is determined by the speed of the interface. The port with the speed 10M have the cost of 100, 100M, 19; 1000M, 4.

17.2.9 Configure STP priority of specified port

Specify specified port in STP by configuring port priority. Generally, the smaller the value is, the superior the priority is, and the port will be more possible to be included in STP. If the priorities are the same, the port number is considered. Configure it in interface configuration mode :

Configure port priority

spanning-tree port-priority *port-priority*

Restore the default port priority

no spanning-tree port-priority

The smaller the value is, the superior the priority is, and the port is easier to be the root interface. Change the port priority may cause the re-calculating of the STP. The port priority ranges from 0 to 255. the default port priority is 128.

For example :

! Configure the port priority of Ethernet 0/0/1 in STP to 120

```
QTECH(config-if-ethernet-0/0/1)#spanning-tree port-priority 120
```

17.2.10 Configure spanning-tree root-guard

Configure spanning-tree root-guard can avoid interface to be root which is used for preventing bone network topology destroying by outer BPDU packet. Configure it in interface configuration mode :

```
configure spanning-tree root-guard
```

spanning-tree root-guard

```
restore to default root-guard
```

no spanning-tree root-guard

Example :

! Enable mst root-guard of e0/0/1

```
QTECH(config-if-ethernet-0/0/1)#spanning-tree root-guard
```

17.2.11 Configure interface to force to send rstp packet

This configuration is used to check whether there is traditional network bridge running STP.

Configure it in interface configuration mode :

Configure interface to force to send rstp packet

spanning-tree mcheck

For example :

! Configure Ethernet 0/0/1 to send RSTP packet

```
QTECH(config-if-ethernet-0/0/1)#spanning-tree mcheck
```

17.2.12 Configure link type of specified interface

In rstp, the requirement of interface quickly in transmission status is that the interface must be point to point link not media sharing link. It can specified interface link mode manually and can also judge it by network bridge.

Configure it in interface configuration mode :

Configure interface to be point-to-point link

spanning-tree point-to-point forcetrue

Configure interface not to be point-to-point link

spanning-tree point-to-point forcefalse

Configure switch auto-detect whether the interface is point-to-point link

spanning-tree point-to-point auto

For example :

! Configure the link connected to Ethernet 0/0/1 as a point-to-point link

```
QTECH(config-if-ethernet-0/0/1)#spanning-tree point-to-point forcetrue
```

17.2.13 Configure the current port as an edge port

Edge port is the port connecting to the host which can be in transmission status in very short time after linkup, but once the port receiving STP packet, it will shift to be non-edge port.

Configure it in interface configuration mode :

Configutr the port to be edge port

spanning-tree portfast

Configutr the port to be non-edge port

no spanning-tree portfast

For example :

! Configure Ethernet 0/0/1 as a non-edge port.

QTECH(config-if-ethernet-0/0/1)#spanning-tree portfast

17.2.14 Configure the speed limit of sending BPDU of specified interface

Restrict STP occupying bandwidth by restricting the speed of sending BPDU packet. The speed is determined by the number of BPDU sent in each hello time.

Configure it in interface configuration mode :

Configure the maximum number of configuration BPDUs sent by interface in each Hello time to be number

spanning-tree transit-limit *number*

For example :

! Configure the maximum number of configuration BPDUs that can be transmitted by the Ethernet 0/0/1 in each Hello time to 2

QTECH(config-if-ethernet-0/0/1)#spanning-tree transit-limit 2

17.2.15 STP monitor and maintenance

17.2.15.1 Display STP status

The displaying information is as following :

- STP status
- BridgeID
- Root BridgeID
- All kinds of configuration parameter of STP

Use following command in any configuration mode to display STP status globally or on a port :

show spanning-tree interface

For example :

! Display STP configuration

QTECH(config)#show spanning-tree interface ethernet 0/0/1

The bridge is executing the IEEE Rapid Spanning Tree protocol

The bridge has priority 32768, MAC address : 001f.ce10.14f1

Configured Hello Time 2 second(s), Max Age 20 second(s),

Forward Delay 15 second(s)

Root Bridge has priority 32768, MAC address 001f.ce10.14f1

Path cost to root bridge is 0

Stp top change 42 times

Port 1 (Ethernet0/0/1) of bridge is disabled

Spanning tree protocol is enabled

```

remote loop detect is enabled
The port is a DesignatedPort
Port path cost 200000
Port priority 128
Designated bridge has priority 32768, MAC address 001f.ce10.14f1
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), Packet Age 6
sent BPDU :    9
        TCN : 0, RST : 9, Config BPDU : 0
received BPDU : 4040
        TCN : 0, RST : 4040, Config BPDU : 0

```

17.2.16 Enable/disable STP remote-loop-detect

When multi-layer cascading, if switch in media layer shut down STP, the BPDU packet sent by upper switch will be cut by switch in media layer. When there is loop in the network below the media layer, upper switch cannot detect the loop. Remote loop detect is the complementary for this situation.

17.2.16.1 Enable STP remote-loop-detect

In interface configuration mode

spanning-tree remote-loop-detect

In global configuration mode

spanning-tree remote-loop-detect interface

Use no command to disable this function.

For example :

```
! Enable spanning-tree remote-loop-detect interface of Ethernet 0/0/1
```

```
QTECH(config)#spanning-tree remote-loop-detect interface ethernet 0/0/1
```

```
! Disable remote-loop-detect of Ethernet 0/0/1
```

```
QTECH(config-if-ethernet-0/0/1)#no spanning-tree remote-loop-detect
```

17.3 Brief Introduction of MSTP

Multiple spanning tree(IEEE802.1S, MSTP) is the upgrade for SST(Simple spanning tree, IEEE802.1D/802.1W). SST can realize link redundancy and loopback but cause the waste of effective bandwidth and overload of some link but backup of others because all vlans share a tree. MSTP makes up these flaw and realize overload balance as SST by mapping different vlan to different STP example, that is, different STP example can generate different topology and different vlan data can choose different transmitting channel according to different STP example.

17.4 MSTP Configuration

17.4.1 MSTP configuration list

The parameter in MSTP configuration only can be effective after STP enables and the mode is MSTP. The parameter configuration is reserved after MSTP disables and enables next time. The MSTP configuration list is as following :

- Configure MSTP timer parameter
- Configure MSTP configuration mark
- Configure MSTP netbridge priority
- Configure MSTP interface edge interface status
- Configure MSTP interface link type
- Configure MSTP interface path cost
- Configure MSTP interface priority
- Display MSTP configuration information
- Enable/disable digest snooping
- Configure Ignore of VLAN

17.4.2 Configure MSTP timer parameter

MSTP timer parameter includes : forward delay, hello time, max age and max hops.
Configure it in global configuration mode :

Configure forward delay

spanning-tree mst forward-time *forward-time*

Configure hello time

spanning-tree mst hello-time *hello-time*

Configure max age

spanning-tree mst max-age *max-age*

Configure max hops

spanning-tree mst max-hops *max-hops*

Example :

! Configure max hop to be 10

QTECH(config)#spanning-tree mst max-hops 10

17.4.3 Configure MSTP configuration mark

MSTP configuring mark includes : MSTP name, MSTP revision level and the mapping relationship between MSTP and VLAN. MSTP possesses the same configuring mark and interconnected network can be treated as a virtual network logically.

Configure it in global configuration mode :

Configure MSTP name

spanning-tree mst name *name*

Configure MSTP revision level

spanning-tree mst revision *revision-level*

Configure mapping relationship between MSTP and VLAN

spanning-tree mst instance *instance-num* *vlan* *vlan-list*

Example :

! Configure MSTP name to be QTECH

```
QTECH(config)#spanning-tree mst name QTECH
```

! Configure MSTP revision level to be 10

```
QTECH(config)#spanning-tree mst revision 10
```

! Configure VLAN2~7 mapping to STP instance 5

```
QTECH(config)#spanning-tree mst instance 5 vlan 2-7
```

17.4.4 Configure MSTP netbridge priority

In MSTP, netbridge priority is the parameter based on each STP instance. Netbridge priority and interface path cost determine the topology of each STP instance which construct the base of link load balance.

Configure it in global configuration mode :

Configure netbridge priority in MSTP instance

spanning-tree mst instance *instance-num* **priority** *priority*

Example :

! Configure netbridge priority in MSTP instance 4 to be 4096

```
QTECH(config)#spanning-tree mst instance 4 priority 4096
```

17.4.5 Configure MSTP interface edge interface status

As SST, interface with edge interface attribution will turn to forwarding if it hasn't received STP packet after 2 sending periods when link up.

Configure it in interface configuration mode :

Configure port to be edge port

spanning-tree mst portfast

! Example :

Configure e0/0/2 to be edge port

```
QTECH(config-if-ethernet-0/0/2)#spanning-tree mst portfast
```

17.4.6 Configure MSTP interface link type

There are two types of link : one is sharing media (through hub), the other is point-to-point. Link type is used in suggest-agree mechanism of interface fast shifting. Only point-to-point allows fast shift. Link type can be manual configured or self-detected by STP protocol.

Configure it in interface configuration mode :

Configure detection of link type

spanning-tree mst link-type point-to-point { forcetrue | forcefalse | auto }

! Example

Configure link type of e0/0/2 to be point-to-point

```
QTECH(config-if-ethernet-0/0/2)#spanning-tree mst link-type point-to-point forcefalse
```

17.4.7 Configure MSTP interface path cost

Port path cost can be divided into internal cost and external cost. The former is the configuration parameter based on each MSTP instance to determine topology of different instance in each MSTP region. The latter is parameter which has nothing to do with the instance to determine CST topology consisted by each region.

Configure it in interface configuration mode :

Configure the path cost in some instance

spanning-tree mst instance instance-num cost *cost*

Configure external path cost

spanning-tree mst external cost *cost*

Example :

! Configure the path cost in instance 2 to be 10

```
QTECH(config-if-ethernet-0/0/2)#spanning-tree mst instance 1 cost 10
```

! Configure external path cost of e0/0/2 to be 10

```
QTECH(config-if-ethernet-0/0/2)#spanning-tree mst external cost 10
```

17.4.8 Configure MSTP interface priority

In MSTP, interface priority is based on each STP instance.

Configure it in interface configuration mode :

Configure interface priority in some instance

spanning-tree mst instance instance-num port-priority *priority*

! Configure priority of e0/0/2 in instance 1 to be 16

```
QTECH(config-if-ethernet-0/0/2)#spanning-tree mst instance 1 port-priority 16
```

17.4.9 Configure spanning-tree mst root-guard

Configure spanning-tree root-guard can avoid interface to be root which is used for preventing bone network topology destroying by outer BPDU packet. Configure it in interface configuration mode :

configure spanning-tree mst root-guard

spanning-tree mst root-guard

restore to default root-guard

no spanning-tree mst root-guard

Example :

! Enable mst root-guard of e0/0/1

```
QTECH(config-if-ethernet-0/0/1)#spanning-tree mst root-guard
```

17.4.10 Display MSTP configuration information

Basic information of MSTP includes : one is MSTP configuring mark(including MSTP name, MSTP revision level and the mapping relationship between MSTP and VLAN), the other is STP instance and interface configuration.

Configure it in any mode :

Display MSTP configuring mark

show spanning-tree mst *config-id*

Display interface information of some instance

show spanning-tree mst instance *instance-num* interface [*interface-list*]

! Example :

Display MSTP configuring mark

```
QTECH(config)#show spanning-tree mst config-id
```

Display interface 0/0/2 information of instance1

```
QTECH(config)#show spanning-tree mst instance 1 interface ethernet 0/0/2
```

17.4.11 Enable/disable digest snooping

When interface of switch connects to switch which has its own private STP, switch cannot connect to each other because of the private STP protocol. Digest snooping can avoid it. Enable digest snooping, switch will think the BPDU packet from other switch is from the same MST region and it will keep the configuring notes and add the notes to the BPDU packet to be sent. Switch realizes interconnection with others in MSTP.

Configure it in interface configuration mode :

Enable interface digest-snooping

spanning-tree mst config-digest-snooping

Disable interface digest-snooping

no spanning-tree mst config-digest-snooping

Example :

! Enable digest-snooping of interface 0/0/1

```
QTECH(config-if-ethernet-0/0/1)# spanning-tree mst config-digest-snooping
```

17.4.12 Configure Ignore of VLAN

In order to control MSTP, Ignore of VLAN can be enabled and the corresponded interface will not calculate. Configure it in global configuration mode :

Enable Ignore of VLAN

spanning-tree mst ignored vlan *vlan-list*

Disable Ignore of VLAN

no spanning-tree mst ignored vlan *vlan-list*

Display Ignore of VLAN

show spanning-tree mst *ignored-vlan*

Example :

! Enable Ignore of VLAN 10 and 20-30

```
QTECH(config)# spanning-tree mst ignored vlan 10, 20-30
```

Chapter 18 Flex links Configuration

18.1 Brief introduction of Flex links

Flex links is layer 2 links backup protocol which provides for STP option scheme. Choose Flex links to realize link backup when the STP is not wanted in customer network. If STP enables, flex links is disabled. Flex links consists of a pair of interfaces(can be ports or convergent interface). One interface is transmitting data, the other is standby. The backup interface starts transmitting data when there is default in master link. The failure interface will be standby when it turns well and it will be transmitting data in 60 seconds when preempt mechanism is set. Flex links interface should disable STP and Flex links interface can configure bandwidth and delay being preempt mechanism and the superior one will be the master interface. There must be trap alarm when master or backup link default.

18.2 Flex links Configuration

18.2.1 Flex links Configuration list

- Enable or disable Flex links of interface(or convergent interface)
- Configure Flex links preemption mode
- Configure Flex links preemption mode delay
- Display Flex links information

18.2.2 Enable or disable Flex links of interface(or convergent interface)

Configure interface Flex links in interface configuration mode
 switchport backup { interface interface-num | channel-group channel-group-number}
 Configure channel-group Flex links in global configuration mode :
 channel-group channel-group-number backup { interface interface-num | channel-group channel-group-number}

For example :

```
!Configure flex links backup interface of e0/0/1 to be e0/0/2
QTECH(config-if-ethernet-0/0/1)#switchport backup interface Ethernet 0/0/2
!Configure flex links backup interface of channel-group 1 to be e0/0/2
QTECH(config)#channel group 1 backup interface Ethernet 0/0/2
```

18.2.3 Configure Flex links preemption mode

Configure interface Flex links in interface configuration mode
 switchport backup { interface interface-num | channel-group channel-group-number} preemption mode {Forced|Bandwidth|Off}
 Configure channel-group Flex links in global configuration mode :
 channel-group channel-group-number backup { interface interface-num | channel-group channel-group-number} preemption mode {Forced|Bandwidth|Off}

For example :

```
!Configure flex links preemption mode of e0/0/1 to be Forced
QTECH(config-if-ethernet-0/0/1)#switchport backup interface Ethernet 0/0/2 preemption mode Forced
!Configure flex links preemption mode of channel-group 1 to be Forced
QTECH(config)#channel group 1 backup interface Ethernet 0/0/2 preemption mode Forced
```

18.2.4 Configure Flex links preemption mode delay

Configure interface Flex links in interface configuration mode

```
switchport backup { interface interface-num | channel-group channel-group-number} preemption delay delay-time
```

Configure channel-group Flex links in global configuration mode :

```
channel-group channel-group-number backup { interface interface-num | channel-group channel-group-number}
preemption delay delay-time
```

For example :

```
!Configure flex links preemption delay of e0/0/1 to be 60 seconds
QTECH(config-if-ethernet-0/0/1)#switchport backup interface Ethernet 0/0/2 preemption delay 60
!Configure flex links preemption delay of channel-group 1 to be 60 seconds
QTECH(config)#channel group 1 backup interface Ethernet 0/0/2 preemption delay 60
```

18.2.5 Display Flex links information

In any configuration mode it will display as following:

```
Flex links master interface status
Flex links backup interface status
Flex links preemption mode
Flex links preemption delay
```

show interface switchport backup

For example :

```
!Display all Flex links information
QTECH(config)# show interface switchport backup
```

Chapter 19 Monitor Link Configuration

19.1 Brief introduction of Monitor Link

Monitor Link is for perceiving up/down status changes in indirect-connected port to achieve port linkage. Monitor Link is used for L2 topological networking to monitor up/down status of uplink and trigger up/down changes on downlink to trigger backup link switch on downlink topology.

19.1.1 Monitor Link group

Each Monitor Link group consists of both uplink and downlink and the member role is determined by user. There can be multiple port members in uplink and downlink, but each member can only belong to one Monitor Link group. Port member can be both ethernet port and aggregation port.

19.1.2 Uplink

Uplink is monitored link in Monitor Link group. When there is no uplink member or uplink member port is down, Monitor Link group is down. Only one uplink member is up, Monitor Link group is up.

19.1.3 downlink

downlink is passive link in Monitor Link group. When there is up/down changes in Monitor Link group, downlink will be changed with it.

19.2 Monitor Link Configuration

19.2.1 Monitor Link Configuration list

- Enable/disable Monitor Link
- Show Monitor Link

19.2.2 Enable/disable Monitor Link

Configure port Monitor Link in interface configuration mode
switchport monitor-link-group group-ID {uplink |downlink}
Configure Monitor Link of channel group in global configuration mode:

channel-group channel-group-number **monitor-link-group** group-ID {uplink |downlink}

Example :

!configure e0/0/1 to be uplink of Monitor Link group 2

```
QTECH(config-if-ethernet-0/0/1)# switchport monitor-link-group 2 uplink
!configure e0/0/1 to be downlink of Monitor Link group 1
QTECH(config)# channel-group 1 monitor-link-group 1 downlink
```

19.2.3 Show Monitor Link

Configure it in any configuration mode:

show monitor-link-group

Example :

!show all Monitor Link

```
QTECH(config)# show monitor-link-group
```

Chapter 20 802.1X Configuration Command

20.1 Brief introduction of 802.1X configuration

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 security hidden trouble. For application of motional 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 authentication, 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.

System realizes IEEE 802.1X authentication. Use IEEE 802.1X authentication needs : RADIUS server which system can access to make the authentication informayion to send to; IEEE 802.1X authentication client software installed in accessing user's device (such as PC).

20.2 802.1X Configuration

Configure system or interface related parameter before enabling 802.1X authentication and these configurations will be saved after disabling 802.1X. And the parameter will be effective after re-enabling 802.1X.

802.1X configuration list is as following :

- Configure RADIUS and TACACS+ project
- Configure domain
- Configure 802.1X

20.2.1 AAA configuration mode

Finish necessary configuration of domain and RADIUS project of 802.1X authentication.

Use aaa command in global configuration mode to enter AAA configuration mode.

For example :

! Enter AAA configuration mode

```
QTECH(config)#aaa
```

```
QTECH(config-aaa)#
```

20.3 RADIUS and TACACS+ Server Configuration

There are three kinds of users :

- Super-administrator
- Administrator
- Normal user

The normal users can only be in the user's mode after logging in the switch so they can only check the basic information about operation and statistics; administrator can enter each configuration mode to check and manage the system; super-administrator can both manage the system and all kinds of users.

Note :

Normal users cannot configure the switch and change their own password. Administrator can manage himself; for example, change his own privilege and password. It cannot create or delete other users and change other user's password and privilege.

This chapter contains following sections :

- [System default user](#)
- [User's authentication](#)
- [Add users](#)
- [Change password](#)
- [Modify User's Privilege Level](#)
- [Delete User](#)
- [Show users](#)
- [Configure RADIUS to be remote authentication server](#)
- [Configure TACACS+ remote authentication](#)

20.3.1 System default user

There is an internal username with password called Super-administrator. It processes the superior priority in the switch to manage both the users and the switch.

The username of Super-administrator is **admin** and its initial password is **123456**. It is suggested modifying the password after the initial-logging in. This username and its administrator privilege cannot be deleted and modified.

Note :

There must be only one super-administrator and all the configurations in the manual is setting super-administrator as example.

20.3.2 User's authentication

User's authentication can be divided into local authentication and remote authentication :

Local authentication : The users' account and password are saved in local database. All users are supported by local authentication.

Remote authentication : The users' account and password are saved in RADIUS/TACACS+ server. Super-administrator "admin" is not supported by remote authentication.

20.4 Local authentication configuration

20.4.1 Add users

At most 15 users can be added. Log in the switch first as Super-administrator and create new users as following steps :

Step	Command	Description
1	enable	Enter privileged mode
2	config terminal	Enter global configuration mode
3	username <i>username</i> privilege	Adding a new user and specified the

	<i>privilege <0,1> password <i>password</i></i>	privilege.
4	show username	Check the configuration.
5	exit	Exit to user mode
6	copy running-config startup-config	Save the configuration

 **Note :**

Username : it means the name of the user to be added which must be 1 to 32 printable characters without '!', ':', '|', '*', '?', '\\', '<', '>', '|', ''.


Level : means the priority of the user to be added which is the number between 0 and 15. 0 and 1 mean the normal user and 2 to 15 mean the administrator.

encryption-type : it can be 0 or 7. 0 means clear text and 7 means encrypted text (not supported now) .

privilege it can be 0, 1 or 2 to 15. 0 and 1 mean normal users while 2 to 15 mean administrators.

Password : the login password of new-added user which is 1 to 16 characters.

If the user's privilege level is not specified, it will default to be normal user. There is up to 8 users in the system.

 **Caution : Case-sensitive is for password but not username.**

Example :

```
!Create administrator "qtech" with its password being 1234 and privilege level is 3
QTECH(config)#username qtech privilege 3 password 0 1234
```

20.4.2 Change password

In global configuration mode, Super-administrator "admin" can use following command to change the password of all users, but other administrators can only change their own password. Normal users cannot modify their own password.

Enter global configuration mode (how to enter global configuration mode refers to the first 2 steps in Table) before following the below steps :

Table Modify password

Step	Command	Description
1	username change-password	Enter the modified password following the prompt. The new password will be effective in the next log in.
2	exit	Exit to user mode
3	copy running-config startup-config	Save the configuration

Example :

```
!Change the password of user "qtech" to be 123456
QTECH(config)#username change-password
please input you login password : *****
please input username : qtech
Please input user new password : *****
Please input user comfirm password : *****
change user qtech password success.
```

20.4.3 Modify User's Privilege Level

In global configuration mode, only Super-administrator "admin" can modify the privilege level of other users. Enter global configuration mode (how to enter global configuration mode refers to the first 2 steps in Table 4-1) before following the below steps :

Table 4-3 Modify User's Privilege Level

Step	Command	Description
1	username <i>username</i> privilege <i>privilege</i> <0-15>	Modify user's privilege.
2	show username	Check configuration.
3	exit	Exit to user mode
4	copy running-config startup-config	Save the configuration

Note :

Username : means the name of the existed user to be modified which must be 1 to 32 printable characters without '!', ':', '*', '?', '\', '<', '>', '|', '"'. If the entered username is not existed, add it to be the new one.

Level : means the priority of the existed user (except the Super-administrator) to be modified which is the number between 0 and 15. 0 and 1 mean the normal user and 2 to 15 mean the administrator.



Caution : Case-sensitive is for password but not username.

Example :

```
!Modify the privilege of the existed user "qtech" to be 1 and its password to be 1234
QTECH (config)#username qtech privilege 0 password 0 1234
```

20.4.4 Delete User

Only Super-administrator "admin" can add and delete user in global configuration mode. Enter global configuration mode (how to enter global configuration mode refers to the first 2 steps in Table 4-1) before following the below steps :

Table 4-4 Delete user

Step	Command	Description
1	no username <i>username</i>	Delete user.
2	show username	Check configuration.
3	exit	Exit to user mode
4	copy running-config startup-config	Save the configuration

Note :

Username : means the name of the user to be deleted.

When deleting a user which is used, it will be disconnected before delete it.

Example :

```
!Delete user "qtech"
QTECH(config)#no username qtech
```

20.4.5 Show users

After configuration, you can use following steps to check it. Any configuration mode is permitted.

Table 4-5 Show users

Step	Command	Description
1	show <i>username</i>	Show specific user.

2	show users	Show users' log. At most 5 users are permitted on line at the same time.
---	-------------------	--

20.5 Remote authentication configuration

20.5.1 Configure RADIUS to be remote authentication server

Configure RADIUS remote authentication

Operation	Command	Description
Enter global configuration	configure terminal	-
Enable RADIUS remote authentication	muser radius name {chap pap} [local]	Selected If "local" is configured, it means local authentication is used if remote authentication failed. By default, it is local authentication
Enter AAA configuration mode	aaa	-
Create RADIUS server name and enter RADIUS configuration mode	radius host name	-
Configure IP of authentication/accounting RADIUS server	{ primary-acct-ip primary-auth-ip } A.B.C.D { accounting port authentication port }	Selected Authentication and accounting port should be the same as that of RADIUS server. Generally, they are : Accounting port : 1813 Authentication port : 1812
Configure shared-key of authentication/accounting RADIUS server	{ acct-secret-key auth-secret-key } key	Selected Shared-key should be the same as that of RADIUS server.
Show configuration	show muser	-

20.5.2 Configure TACACS+ remote authentication

Configuring user's login through TACACS+ server authentication, accounting and authorization through TACACS+ server can be chosen. When configuring TACACS+ authorization, configure corresponded priority to users first. There are 16 levels (0-16) priorities but there are only 2 levels (0-1 means normal users and 2-15 means administrators) for QTECH switches. When configuring TACACS+ unauthorization, the priority is determined by priv_lvl replied from remote server (no reply means administrator). Authorization failure means normal user. When configuring TACACS+ accounting, it begins with the pass of authentication and ends with user's exit.

Configure TACACS+ remote authentication

Operation	Command	Description
Enable TACACS+ authorization/accounting	muser tacacs+ {account [local] author [local]] local}	Selected If "local" is configured, it means local authentication is used if remote authentication failed. By default, it is local authentication

Configure IP/shared-key/TCP port/timeout of TACACS+ remote server	tacacs+ { primary secondary } server <i>ipaddress [key keyvalue] [port portnum]</i> [timeout timevalue]	Selected By default, TCP port is 49 and timeout is 5 seconds.
Show TACACS+ configuration	show tacacs+	-
Show current authentication	show muser	-

20.5.3 802.1X Configuration

Related command of 802.1X configuration is as following :

- dot1x
- dot1x daemon
- dot1x eap-finish
- dot1x eap-transfer
- dot1x re-authenticate
- dot1x re-authentication
- dot1x timeout re-authperiod
- dot1x timeout re-authperiod interface
- dot1x port-control
- dot1x max-user
- dot1x user cut

(1) Use **dot1x** command to enable 802.1x. Domain and RADIUS server configurations can be effective after this function enabling. Use **no dot1x** command to disable 802.1x. Use **show dot1x** command to display 802.1x authentication information.

After enabling 802.1X, user accessed to system can access VLAN resources after authentication. By default, 802.1X disables.

For example :

! Enable 802.1X

```
QTECH(config)#dot1x
```

! Display 802.1x authentication information

```
QTECH(config)#show dot1x
```

(2) When 802.1x enables, use this command to configure whether a port send and sending period :

dot1x 802.1x daemon

By default, 802.1x daemon is not sent by default. When 802.1x enables, default interval to send daemon is 60seconds.

For example :

! Enable dot1x daemon on ethernet 0/0/5 with the period time of 20 seconds

```
QTECH(config-if-ethernet-0/0/5)#dot1x daemon time 20
```

(3) Use **dot1x eap-finish** and **dot1x eap-transfer** command to configure protocol type between system and RADIUS server :

After using dot1x eap-transfer command, 802.1 authentication packet encapsulated by EAP frame from user is sent to RADIUS server after transferring to data frame encapsulated by other high level protocol. After using dot1x eap-transfer command, 802.1 authentication packet encapsulated by EAP frame from user is sent to RADIUS server without any changes.

For example :

! Configure authentication packet transitting to be eap-finish

```
QTECH(config)#dot1x eap-finish
```

(4) Use dot1x re-authenticate command to re-authenticate current interface. Use dot1x re-authentication

command to enable 802.1x re-authentication. Use no dot1x re-authentication command to disable 802.1x re-authentication. Use dot1x timeout re-authperiod command to configure 802.1x re-authperiod. Use dot1x timeout re-authperiod interface command to configure 802.1x re-authperiod of a specified interface. Please refer to command line configuration to see the details.

(5) Use dot1x port-control command to configure port control mode.

After 802.1X authentication enables, all interfaces of the system default to be needing authentication, but interfaces of uplink and connecting to server need not authentication. Use dot1x port-control command to configure port control mode. Use no dot1x port-control command to restore the default port control. Use show dot1x interface command to display configuration of interface.

Configure it in interface configuration mode :

dot1x port-control { auto | forceauthorized | forceunauthorized }

For example :

! Ethernet 0/0/5 is RADIUS server port. Configure port-control mode of ethernet 0/0/5 to be forceauthorized in interface configuration mode

```
QTECH(config-if-ethernet-0/0/5)#dot1x port-control forceauthorized
```

! Display 802.1X configuration of ethernet 0/0/5

```
QTECH(config)#show dot1x interface ethernet 0/0/5
```

```
port ctrlmode      Reauth  ReauthPeriod(s) MaxHosts
```

```
e0/0/5  forceauthorized disabled  3600      160
```

```
Total [26] item(s), printed [1] item(s).
```

(6) Use dot1x max-user command to configure the maximum number of supplicant systems an ethernet port can accommodate. Use no dot1x max-user command to configure the maximum number to be 1.

Configure it by using following command :

dot1x max-user *user-num*

For example :

! Configure the max-user of ethernet 0/0/5 is 10 in interface configuration mode

```
QTECH(config-if-ethernet-0/0/5)#dot1x max-user 10
```

(7) Use dot1x user cut command to remove specified online user.

Remove specified online user by specified username and MAC address.

For example :

! Remove user with username of aaa@QTECH.com

```
QTECH(config)#dot1x user cut username aaa@QTECH.com
```

Chapter 21 SNTP Client Configuration

21.1 Brief introduction of SNTP protocol

The working theory of SNTP is as following :

SNTPv4 can be worked in three modes : unicast, broadcast (multicast) and anycast.

In unicast mode, client actively sends requirement to server, and server sends response packet to client according to the local time structure after receiving requirement.

In broadcast and multicast modes, server sends broadcast and multicast packets to client timing, and client receives packet from server passively.

In anycast mode, client actively uses local broadcast or multicast address to send requirement, and all servers in the network will response to the client. Client will choose the server whose response packet is first received to be the server, and drops packets from others. After choosing the server, working mode is the same as that of the unicast.

In all modes, after receiving the response packet, client resolves this packet to obtain current standard time, and calculates network transmit delay and local time complementary, and then adjusts current time according them.

21.2 SNTP client configuration

SNTP client configuration command includes :

- Enable/disable SNTP client
- SNTP client working mode configuration
- SNTP client unicast server configuration
- SNTP client broadcast delay configuration
- SNTP client multicast TTL configuration
- SNTP client poll interval configuration
- SNTP client retransmit configuration
- SNTP client valid server configuration
- SNTP client MD5 authentication configuration

21.2.1 Enable/disable SNTP client

Use `sntp client` command in global configuration mode to enable SNTP client. Use `no sntp client` command to disable SNTP client. After SNTP enabling, switch can obtain standard time through internet by SNTP protocol to adjust local system time.

Enable SNTP client using following command :

sntp client

no sntp client

For example :

! Enable SNTP client

```
QTECH(config)#sntp client
```

21.2.2 SNTP client working mode configuration

SNTPv4 can work in three modes : unicast, broadcast (multicast), anycast. In unicast and anycast, client

sends requirement and gets the response to adjust system time. In broadcast and multicast, client waits for the broadcast packet sent by server to adjust system time.

sntp client mode { broadcast | unicast | anycast [key number] | multicast }

no sntp client mode

For example :

! Configure SNTP client to operate in anycast

QTECH(config)#sntp client mode anycast

21.2.3 SNTP client unicast server configuration

In unicast mode, SNTP client must configure server address. The related command is as following :

sntp server *ip-address* [*key number*]

no sntp server

Only in unicast, configured server address can be effective.

For example :

! Configure unicast server ip-address to be 192.168.0.100

QTECH(config)#sntp server 192.168.0.100

21.2.4 SNTP client broadcast delay configuration

SNTP client broadcast delay configuration is as following :

sntp client broadcastdelay *milliseconds*

no sntp client broadcastdelay

Only in broadcast (multicast), configured transmit delay can be effective. After configuration, SNTP client can add transmit delay after obtaining time from server to adjust current system time.

For example :

! Configure broadcastdelay to be 1 second

QTECH(config)#sntp client broadcastdelay 1000

21.2.5 SNTP client multicast TTL configuration

Use following command to configure ttl-value of multicast packet :

sntp client multicast ttl *ttl-value*

no sntp client multicast ttl

This command should be effective by sending packet through multicast address in anycast operation mode. In order to restrict the range of sending multicast packet, TTL-value setting is suggested. The default ttl-value is 255.

For example :

! Configure TTL-value of sending multicast packet to be 5

QTECH(config)#sntp client multicast ttl 5

21.2.6 SNTP client poll interval configuration

Use following command to configure poll-interval of SNTP client in unicast or anycast. :

sntp client poll-interval *seconds*

no sntp client poll-interval

Only in unicast and anycast mode, configured poll interval can be effective. SNTP client sends requirement in a poll interval to the server to adjust current time.

For example :

! Configure poll-interval to be 100 seconds

QTECH(config)#sntp client poll-interval 100

21.2.7 SNTP client retransmit configuration

Uses following command to configure retransmit times inunicast and anycast operation mode. :

sntp client retransmit *times*

no sntp client retransmit

sntp client retransmit-interval *seconds*

no sntp client retransmit-interval

This command is effective in unicast and anycast operation mode. SNTP requirement packet is UDP packet, overtime retransmission system is adopted because the requirement packet cannot be guaranteed to send to the destination. Use above commands to configure retransmit times and the interval.

For example :

! Configure overtime retransmission to be twice and the interval to be 5

```
QTECH(config)#sntp client retransmit-interval 5
```

```
QTECH(config)#sntp client retransmit 2
```

21.2.8 SNTP client valid server configuration

In broadcast and multicast mode, SNTP client receives protocol packets from all servers without distinction. When there is malice attacking server (it will not provide correct time), local time cannot be the standard time. To solve this problem, a series of valid servers can be listed to filtrate source address of the packet.

Corresponded command is as following :

sntp client valid-server *ipaddress*

no sntp client valid-server

For example :

! Configure servers in network interface 10.1.0.0/16 to be valid servers

```
QTECH(config)#sntp client valid-server 10.1.0.0 0.0.255.255
```

21.2.9 SNTP client MD5 authentication configuration

SNTP client can use valid server list to filtrate server, but when some malice attackers using valid server address to forge server packet and attack switch, switch can use MD5 authentication to filtrate packet, and authenticated packet can be accepted by client.

Configuration command is as following :

sntp client authenticate

no sntp client authenticate

sntp client authentication-key number md5 *value*

no sntp client authentication-key *number*

sntp trusted-key number

no sntp trusted-key number

For example :

! Configure SNTP client MD5 authentication-key, with the key ID being 12, and the key being abc and trusted-key being 12

```
QTECH(config)#sntp client authenticate
```

```
QTECH(config)#sntp client authentication-key 12 md5 abc
```

```
QTECH(config)#sntp trusted-key 12
```

Chapter 22 Syslog Configuration

22.1 Brief introduction of Syslog

Syslog is system information center, which handles and outputs information uniformly.

Other modules send the information to be outputted to Syslog, and Syslog confirms the form of the outputting of the information according to user's configuration, and outputs the information to specified displaying devices according to the information switch and filtration rules of all outputting directions.

Because of Syslog, information producer all modules of outputting information need not care where the information should be send at last, console, telnet terminal or logging host (Syslog server). They only need send information to Syslog. The information consumer console, Telnet terminal, logging buffer, logging host and SNMP Agent can choose the information they need and drop what they needn't for suitable filtration rules.

Syslog information level reference :

several level	Description	corresponded explanation
0 : emergencies	the most emergent error	need reboot
1 : alerts	need correct immediately	self-loop, hardware error
2 : critical	key error	memory, resources distribution error
3 : errors	non-key errors need cautions	general error; invalid parameter which is hard to restore
4 : warnings	Warning for some error which may exist	alarm; losing packet which is not important; disconnect with the exterior server
5 : notifications	information needs cautions	Trap backup outputting
6 : informational	general prompt information	command line operation log; set operation for MIB node
7 : debugging	debug information	debugging outputting; process, data of service protocol

22.2 Syslog Configuration

Syslog configuration command includes :

- Enable/disable Syslog
- Syslog sequence number configuration
- Syslog time stamps configuration
- Syslog logging language configuration
- Syslog terminal outputting configuration
- Syslog logging buffered outputting configuration
- Syslog Flash storage outputting configuration

- Syslog logging host outputting configuration
- Syslog SNMP Agent outputting configuration
- Module debug configuration

22.2.1 Enable/disable Syslog

Use logging command in global configuration mode to enable Syslog. Use no logging command to disable Syslog and no information will be displayed.

Configuration command is as following :

logging

no logging

For example :

! Enable Syslog

QTECH(config)#logging

22.2.2 Syslog sequence number configuration

Use logging sequence-numbers command to configure global sequence number to be displayed in Syslog. Use no logging sequence-numbers command to configure global sequence number not to be displayed in Syslog.

logging sequence-numbers

no logging sequence-numbers

For example :

! Configure global sequence number to be displayed in Syslog outputting information.

QTECH(config)#logging sequence-numbers

22.2.3 Syslog time stamps configuration

Use following command to configure the type of timestamps in Syslog. There 3 types of timestamps : timestamps are not displayed, uptime is the timestamps, and datatime is the timestamps.

Configure command is as following :

logging timestamps { notime | uptime | datetime }

no logging timestamps

For example :

! Configure datetime to be the timestamps

QTECH(config)#logging timestamps datetime

22.2.4 Syslog terminal outputting configuration

Use following command in global configuration mode to enable monitor logging and configure filter regulation.

(1) Logging monitor configuration command is as following :

logging monitor { all | *monitor-no* }

no logging monitor { all | *monitor-no* }

monitor-no : 0 means console, and 1 to 2 means Telnet terminal.

For example :

! Enable monitor logging

QTECH(config)#logging monitor 0

(2) Terminal monitor configuration command is as following :

terminal monitor

no terminal monitor

This command has influence on current terminal and current log in.

For example :

! Enable current terminal information displaying

```
QTECH(config)#terminal monitor
```

(3) Logging monitor configuration command is as following :

logging monitor { all | *monitor-no* } { level | none | level-list { level [to level] } &<1-8> } [module { xxx | ... } *]

no logging monitor { all | monitor-no } **filter**

xxx : means the name of the module. ... means other modules are omitted

For example :

! Configure filter regulations of all terminals to allow all modules of levels 0 to 7 to output information

```
QTECH(config)#logging monitor 0 7
```

22.2.5 Syslog logging buffered outputting configuration

Use logging buffered command in global configuration mode to enable buffered logging and configure filter regulations. Use no logging buffered command to disable buffered logging and restore to default filter regulations.

(1) Logging buffered configuration command is as following :

logging buffered**no logging buffered**

For example :

! Enable buffered logging

```
QTECH(config)# logging buffered
```

(2) Filtration rules configuration command is as following :

logging buffered { level | none | level-list { level [to level] } &<1-8> } [module { xxx | ... } *]

no logging buffered filter

xxx : means the name of the module. ... means other modules are omitted.

For example :

! Configure filter regulations of all terminals to allow all module of level 0 to 6 to output information

```
QTECH(config)#logging buffered 6
```

22.2.6 Syslog Flash storage outputting configuration

Use logging flash command in global configuration command to enable flash logging and configure filter regulations.

(1) Logging buffered configuration command is as following

logging flash**no logging flash**

For example :

! Enable flash logging

```
QTECH(config)# logging flash
```

(2) Filtration rules configuration command is as following :

logging flash { level | none | level-list { level [to level] } &<1-8> } [module { xxx | ... } *]

no logging flash filter

xxx : means the name of the module. ... means other modules are omitted.

For example :

! Configure filter regulations of all terminals to allow all modules to output information with the level of 0, 1, 2, 6

QTECH(config)#logging flash level-list 0 to 2 6

22.2.7 Syslog logging host outputting configuration

Use following command to configure host ip address, and enable host logging, and configure filter regulation of Syslog server.

(1) Server address configuration command is as following :

logging ip-address

no logging ip-address

At most 15 logging hosts are allowed to configure.

For example :

! Configure server address to be 1.1.1.1 :

QTECH(config)#logging 1.1.1.1

(2) Logging buffered configuration command is as following :

logging host { all | ip-address }

no logging host { all | ip-address }

For example : :

! Enable logging host 1.1.1.1

QTECH(config)#logging host 1.1.1.1

(3) Filtration rules configuration command is as following :

logging host { all | ip-address } { level | none | level-list { level [to level] } &<1-8> } [module { xxx | ... } *]

no logging host { all | ip-address } filter

xxx : means the name of the module. ... means other modules are omitted.

For example :

! Configure filter regulations of logging host 1.1.1.1 to allow module vlan of level 7 to output information

QTECH(config)#logging host 1.1.1.1 none

QTECH(config)#logging host 1.1.1.1 level-list 7 module vlan

(4) Logging facility configuration command is as following :

logging facility { xxx | ... }

no logging facility

xxx : The name of logging facilities.... means other logging facilities are omitted.

For example :

! Configure logging facility to be localuse7

QTECH(config)#logging facility localuse7

(5) Fixed source address configuration command is as following :

logging source ip-address

no logging source

ip-address must be an interface address of a device.

For example :

! Configure logging host outputting to use fixed source address 1.1.1.2 :

QTECH(config)#logging source 1.1.1.2

22.2.8 Syslog SNMP Agent outputting configuration

Use logging snmp-agent command to enable SNMP Agent logging and configure filter configuration. Use no logging snmp-agent command to disable SNMP Agent logging and restore to default filter configuration.

Configure Trap host ip address for Syslog information to send to SNMP Workstation by Trap packet. (refer to SNMP configuration)

(1) Logging buffered configuration command is as following :

logging snmp-agent

no logging snmp-agent

For example :

! Enable SNMP Agent logging

```
QTECH(config)#logging snmp-agent
```

(2) Filtration rules configuration command is as following :

logging snmp-agent { level | none | level-list { *level* [*to level*] } &<1-8> } [module { *xxx* | ... } *]

no logging snmp-agent filter

xxx : means the name of the module. ... means other modules are omitted.

For example :

! Configure SNMP Agent filtrate rules to be permitting information with the level 0 ~ 5

```
QTECH(config)#logging snmp-agent 5
```

22.2.9 Module debug configuration

Use debug command to enable debug of a module. Use no debug command to disable debug of a module :

debug { all | { *xxx* | ... } * }

no debug { all | { *xxx* | ... } * }

xxx : means the name of the module. ... means other modules are omitted.

For example :

! Enable debug of module vlan

```
QTECH(config)#debug vlan
```

Chapter 23 LLDP configuration

23.1 Brief introduction of LLDP protocol

LLDP(Link Layer Discovery Protocol)is the new protocol defined by IEEE 802.1AB. It realizes proclaiming information about itself to other neighbor devices through network and receives the bulletin information from neighbor devices and stores it to standard MIB of LLDP. It is convenient for user to check the device model and linked interfaces of downlink neighbor devices and maintains central office and manage network. Network administrator can know the link of network layer 2 by accessing MIB.

23.1.1 LLDP Overview

Link Layer Discovery Protocol (LLDP) operates on data link layer. It stores and maintains the information about the local device and the devices directly connected to it for network administrators to manage networks through NMS (network management systems). In LLDP, device information is encapsulated in LLDPDUs in the form of TLV (meaning type, length, and value) triplets and is exchanged between directly connected devices. Information in LLDPDUs received is restored in standard MIB (management information base).

23.1.1.1 LLDP operating mode

LLDP can operate in one of the following modes.

- TxRx mode. A port in this mode sends and receives LLDPDUs.
- Tx mode. A port in this mode only sends LLDPDUs.
- Rx mode. A port in this mode only receives LLDPDUs.
- Disable mode. A port in this mode does not send or receive LLDPDUs.

LLDP is initialized when an LLDP-enabled port changes to operate in another LLDP operating mode. To prevent LLDP from being initialized too frequently, LLDP undergoes a period before being initialized on an LLDP-enabled port when the port changes to operate in another LLDP operating mode. The period is known as initialization delay, which is determined by the re-initialization delay timer.

23.1.1.2 Sending LLDPDUs

A LLDP-enabled device operating in the TxRx mode or Tx mode sends LLDPDUs to its directly connected devices periodically. It also sends LLDPDUs when the local configuration changes to inform the neighboring devices of the change timely. In any of the two cases, an interval exists between two successive operations of sending LLDPDUs. This prevents the network from being overwhelmed by LLDPDUs even if the LLDP operating mode changes frequently.

To enable the neighboring devices to be informed of the existence of a device or an LLDP operating mode change (from the disable mode to TxRx mode, or from the Rx mode to Tx mode) timely, a device can invoke the fast sending mechanism. In this case, the interval to send LLDPDUs changes to one second. After the device sends specific number of LLDPDUs, the interval restores to the normal. (A neighbor is discovered when a device receives an LLDPDU and no information about the sender is locally available.)

23.1.1.3 Receiving LLDPDUs

An LLDP-enabled device operating in the TxRx mode or Rx mode validates the TLVs carried in the LLDPDUs it receives and stores the valid neighboring information. An LLDPDU also carries a TTL (time to live) setting with it. The information about a neighboring device maintained locally ages out when the corresponding TTL expires.

The TTL of the information about a neighboring device is determined by the following expression :

$$\text{TTL multiplier} \times \text{LLDPDU sending interval.}$$

You can set the TTL by configuring the TTL multiplier. Note that the TTL can be up to 65535 seconds. TTLs longer than it will be rounded off to 65535 seconds.

23.2 LLDP configuration

23.2.1 LLDP configuration list

The configuration can be effective only after LLDP enables. Configure related parameter of devices or Ethernet interface before enabling LLDP and these configurations will be saved after disabling LLDP. And the parameter will be effective after re-enabling LLDP. LLDP configuration list is as following :

- Enable/disable global LLDP
- Configure LLDP hello-time
- Configure LLDP hold-time
- Interface LLDP packet receiving/sending mode configuration
- Display LLDP information

23.2.2 Enable/disable global LLDP

Use following command in global configuration mode :
Enable global LLDP

lldp

Disable global LLDP

no lldp

By default, global LLDP disables.

For example :

! Enable global LLDP

QTECH(config)#lldp

23.2.3 Configure LLDP hello-time

Use following command in global configuration mode :
Configure LLDP hello-time

lldp hello-time <5-32768>

Restore default LLDP hello-time

no lldp hello-time

The default LLDP hello-time is 30 seconds

For example :

! Configure LLDP hello-time to be 10

QTECH(config)#lldp hello-time 10

23.2.4 Configure LLDP hold-time

Use following command in global configuration mode :
Configure LLDP hold-time

lldp hold-time <2-10>

Restore default LLDP hold-time

no lldp hold-time

The default LLDP hold-time is 4

For example :

! Configure LLDP hold-time to be 2

QTECH(config)#lldp hold-time 2

23.2.5 Interface LLDP packet receiving/sending mode configuration

Use following command in interface configuration mode :
Configure interface LLDP packet receiving/sending mode

lldp { rx | tx | rxtx }

Parameter :

rx : only receive LLDP packet

tx : only send LLDP packet

rxtx : receiving/sending LLDP packet

Disable interface LLDP packet receiving/sending

no lldp

By default, interface LLDP packet receiving/sending mode is rxtx

For example :

! Configure e 0/0/1 only to send LLDP packet

```
QTECH(config-if-ethernet-0/0/1)#lldp tx
```

23.2.6 Display LLDP information

Display followings in any configuration mode :

- 1) Enable/disable global LLDP
- 2) Related parameter of global LLDP
- 3) Interface packet receiving/sending mode
- 4) Interface packet receiving/sending statistics
- 5) Neighbour devices information found

show lldp interface [<interface-list>]

For example :

! Display LLDP information of interface Ethernet 0/0/1

```
QTECH(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 : 2019 Pkt Rx : 1943
```

```
Neighbor (1) :
```

```
TTL : 119(s)
```

```
Chassis ID : 00 : 1f : ce : 10 : 14 : f1
```

```
Port ID : port(7)
```

```
System Name : QSW-3900
```

```
System Description : QTECH Switch
```

Port Description : e0/0/7

Port Duplex : auto

Port Speed : FULL-100

Port Link Aggregation : support , in aggregation , aggregated port ID is 7

Chapter 24 ERRP Command Configuration

24.1 Brief introduction of ERRP

ERRP(Ethernet Ring Redundancy Protocol) is the private Ethernet ring protocol of QTECH which is used to protect real-time service (video/voice delay sensitive service). The basic working theory is many switches serial connect to be ring to provide link redundancy, and a master device detects/maintains the ring. The master device provides redundant port which can release redundant port when the ring break down to guarantee the service smooth. The calculation is less, so the convergency is faster than STP.

24.2 ERRP Overview

Ethernet Ring Redundancy Protocol (ERRP) is an Ethernet ring-specific link layer protocol. It can not only prevent data loop from causing broadcast storm efficiently when the Ethernet ring is complete, but also restore communication channels among nodes on the Ethernet ring rapidly when a link is torn down.

Compared with Spanning Tree Protocol (STP), ERRP features :

- Expedited topology convergence
- Independent of the number of nodes on the Ethernet ring

24.3 Basic Concepts in ERRP

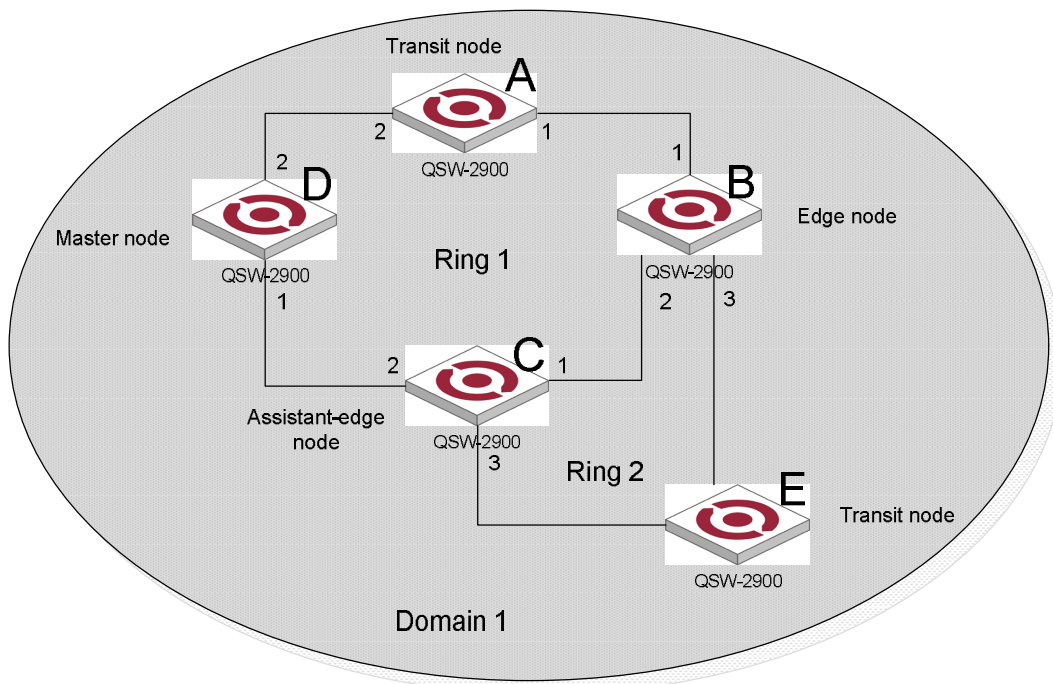


Figure 1 ERRP networking diagram

24.3.1 ERRP domain

The interconnected devices with the same domain ID and control VLANs constitute an ERRP domain. An ERRP domain contains multiple ERRP rings, in which one ring serves as the primary ring and other rings serve as subrings. You can set a ring as either the primary ring or a subring.

As shown in Figure 1, Domain 1 is an ERRP domain, including two ERRP rings : Ring 1 and Ring 2. All the nodes on the two ERRP rings belong to the ERRP domain.

24.3.2 ERRP ring

A ring-shaped Ethernet topology is called an ERRP ring. An ERRP domain is built up on an ERRP ring. An ERRP ring falls into primary ring and subring. Both levels are set to 0 and 1 respectively when configuration.

As shown in Figure 1, Domain 1 contains two ERRP rings : Ring 1 and Ring 2. Ring 1 level is set to 0, meaning the primary ring; Ring 2 level is set to 1, meaning the subring.

For a ring, there are two cases :

- Health state : All the physical links on the Ethernet ring are connected.
- Disconnect state : Some physical link on the Ethernet ring fails.

24.3.3 Control VLAN and data VLAN

- Control VLAN is a VLAN specially designed to transfer ERRP packets. The ports accessing an ERRP ring on devices belong to the control VLAN of the ring and only these ports can join this VLAN. IP address configuration is prohibited on the ports of the control VLAN. You can configure a control VLAN for the primary ring (namely the primary control VLAN). However, the control VLAN of a subring (namely the secondary control VLAN) is assigned automatically by the system and its VLAN ID is the control VLAN ID of the primary ring plus 1.
- Data VLAN is a VLAN designed to transfer data packets, including the ports accessing the Ethernet ring and other ports on devices.

24.3.4 Node

Every device on an ERRP ring is referred to as a node. Node mode includes :

- Master node : Each ring has a master node primarily used to make loop detection and loop guard.
- Transit node : All the nodes excluding the master node on the primary ring; and all the nodes on a subring except for the master node and the nodes where the primary ring intersects with the subring.
- Edge node : A node residing on the primary ring and a subring at the same time. The node is a special transit node that serves as a transit node on the primary ring and an edge node on the subring.
- Assistant-edge node : A node residing on the primary ring and a subring at the same time. The node is a special transit node that serves as a transit node on the primary ring and an assistant-edge node on the subring. This node is used in conjunction with the edge node to detect the integrity of the primary ring and perform loop guard.

As shown in Figure 1, Ring 1 is the primary ring and Ring 2 is a subring. Device A is the master node of Ring 1, Device B, Device C and Device D are the transit nodes of Ring 1; Device B and Device C reside on Ring 2 at the same time, so they are the edge nodes of Ring 2. You can specify one of them as an edge node and the other as an assistant edge node. Device E is the master node of Ring 2.

24.3.5 Primary port and secondary port

Each master node or transit node has two ports accessing an ERRP ring, in which one serves as the primary port and the other serves as the secondary port. You can determine the role of a port.

- 1) In terms of functionality, the difference between the primary port and the secondary port of a master node is :
 - The primary port and the secondary port are designed to play the role of sending and receiving loop-detect packets respectively.
 - When an ERRP ring is in health state, the secondary port of the master node will logically deny data VLANs and permit only the packets of the control VLANs.
 - When an ERRP ring is in disconnect state, the secondary port of the master node will permit data VLANs, that is, forward packets of data VLANs.
- 2) In terms of functionality, there is no difference between the primary port and the secondary port of the transit node. Both are designed for the transfer of protocol packets and data packets over an ERRP ring.

As shown in Figure 1, Device A is the master node of Ring 1. Port 1 and port 2 are the primary port and the secondary port of the master node on Ring 1 respectively. Device B, Device C and Device D are the transit nodes of Ring 1. Their port 1 and port 2 are the primary port and the secondary port on Ring 1 respectively.

24.3.6 Common port and edge port

Each edge node or assistant edge node have two ports accessing a subring, with one being a common port and the other being an edge port. Common port is a port accessing the primary ring and a subring simultaneously; and edge port is a port accessing only a subring.

As shown in Figure 1, Device B and Device C lie on Ring 1 and Ring 1. Device B's port 2 and Device C's port 1 access the primary ring and a subring at the same time, so they are common ports. Device B's port 3 and Device C's port 3 access only a subring, so they are edge ports.

24.3.7 Multi-domain intersection common port

Of the two ports on a node where rings of different domains intersect, the common port is the one on the primary ring that belongs to different domains at the same time. This port must not be on a subring. The role of the port is determined by user configuration.

24.3.8 Timers

The master node uses two timers to send and receive ERRP packets : the Hello timer and the Fail timer.

- The Hello timer is used for the primary port to send Health packets.
- The Fail timer is used for the secondary port to receive Health packets from the master node.

If the secondary port receives the Health packets before the Fail timer expires, the overall ring is in health state. Otherwise, the ring transits into disconnect state until the secondary port receives the Health packet again.

 Note :

In an ERRP domain, a transit node learns the Hello timer value and the Fail timer value on the master node through the received Health packets, guaranteeing the consistency of two timer values across a ring.

The Fail timer value must be greater than or equal to 3 times of the Hello timer value.

24.3.9 ERRP Packets

Table shows the types of ERRP packets and their functions.

ERRP packet types and their functions

Type	Description
Health	The master node initiates Health packets to detect the integrity of a ring in a network.
Link-Down	The transit node, the edge node or the assistant edge node initiates Link-Down packets to notify the master node the disappearance of a ring in case of a link failure.
Common-Flush-FDB	The master node initiates Common-Flush-FDB packets to notify the transit nodes to update their own MAC entries and ARP entries when an ERRP ring transits to disconnect state.
Complete-Flush-FDB	The master node initiates Complete-Flush-FDB packets to notify the transit nodes to update their own MAC entries and ARP entries, and release from blocking ports temporarily when an ERRP ring transits into health state.
Edge-Hello	The edge node initiates Edge-Hello packets to examine the links of the primary ring between the edge node and the assistant edge node.

Type	Description
Major-Fault	Assistant edge node initiates Major-Fault packets to notify the edge node of a failure when a link of primary ring between edge node and assistant edge node is torn down.

24.4 Typical ERRP Networking

Here are several typical networking applications.

24.4.1 Single ring

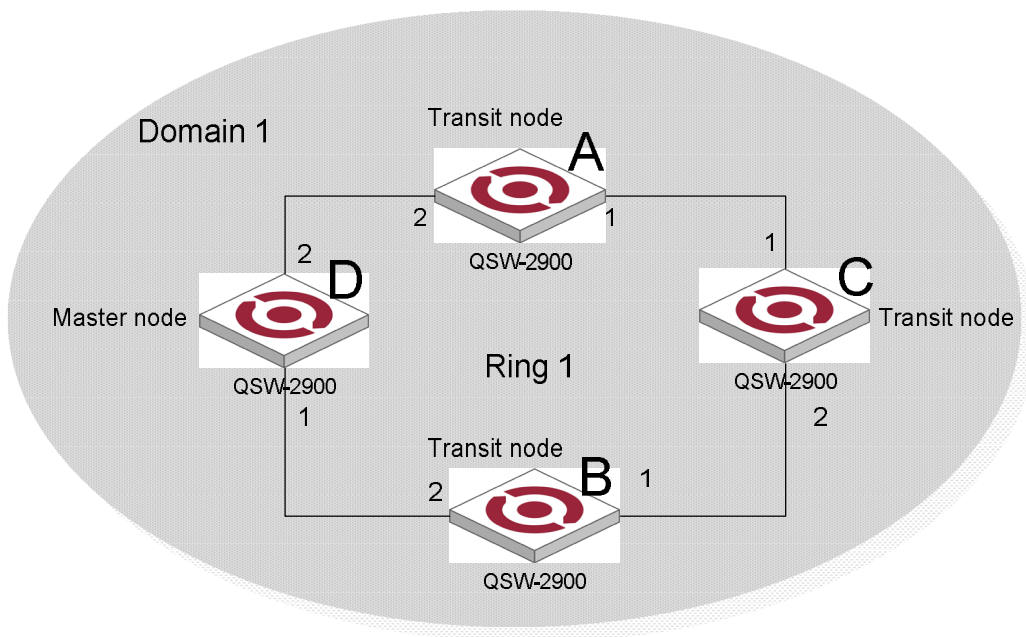


Figure 2 Single ring

There is only a single ring in the network topology. In this case, you only need to define an ERRP domain.

24.4.2 Multi-domain tangent rings

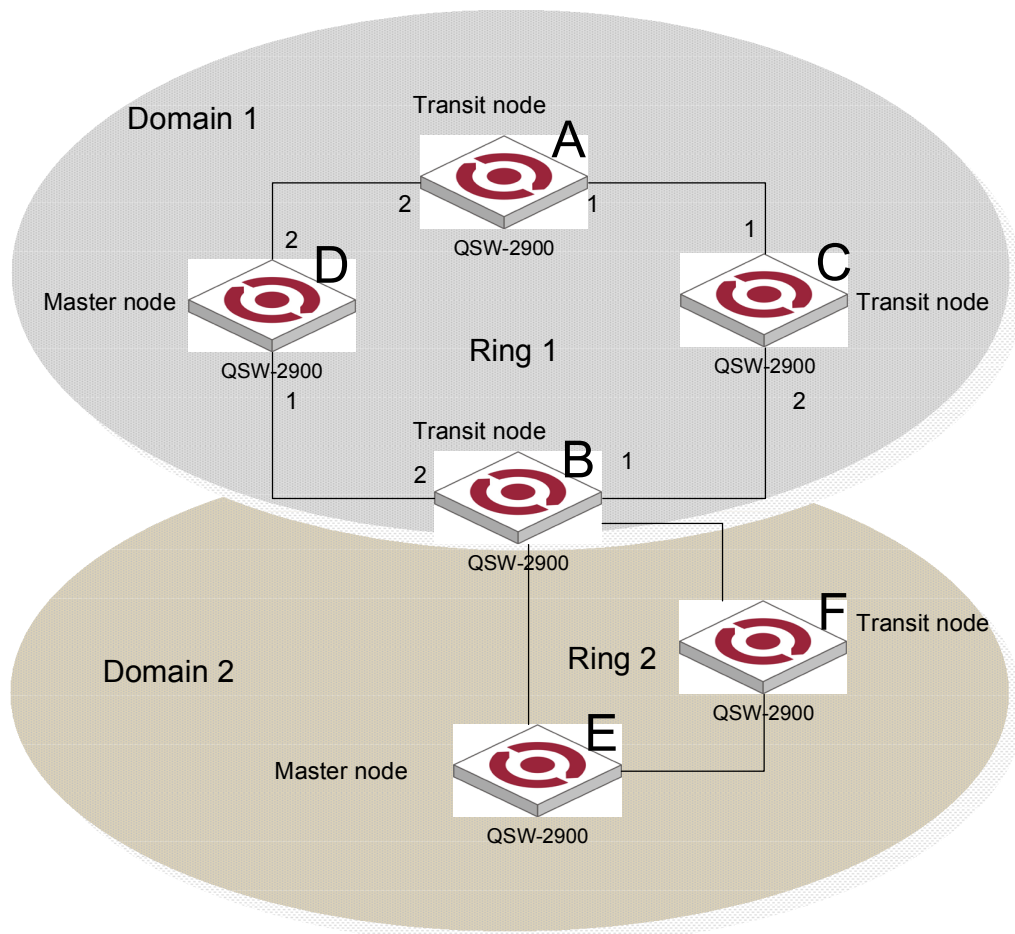


Figure 3 Multi-domain tangent rings

There are two or more rings in the network topology and only one common node between rings. In this case, you need define an ERRP domain for each ring.

24.4.3 Single-domain intersecting rings

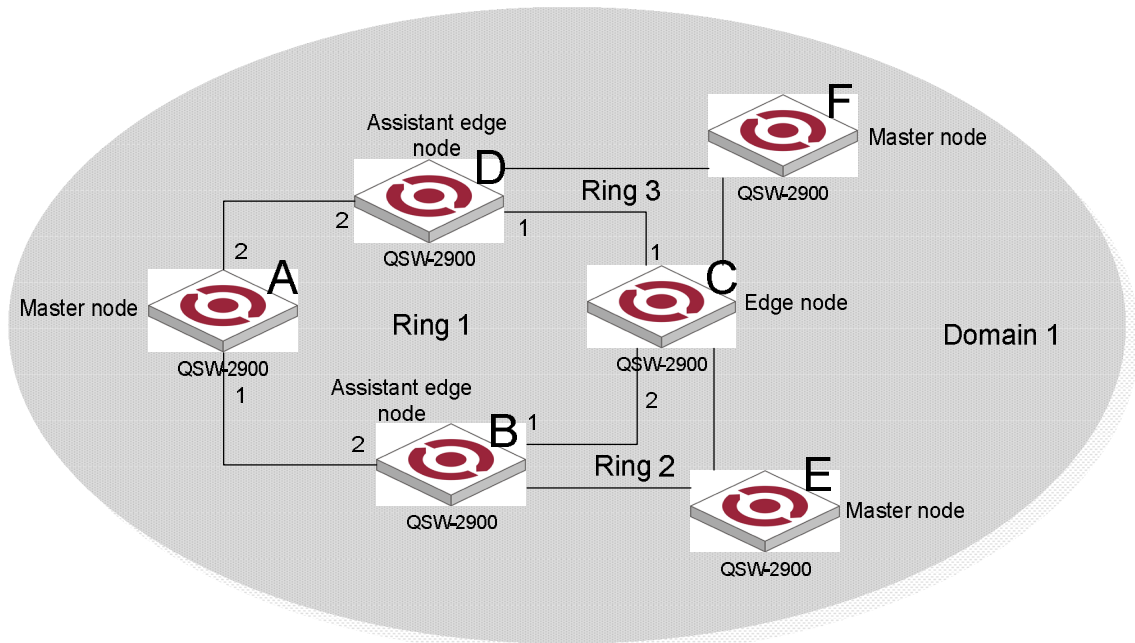


Figure 4 Single-domain intersecting rings

There are two or more rings in the network topology and two common nodes between rings. In this case, you only need to define an ERRP domain, and set one ring as the primary ring and other rings as subrings.

24.4.4 Dual homed rings

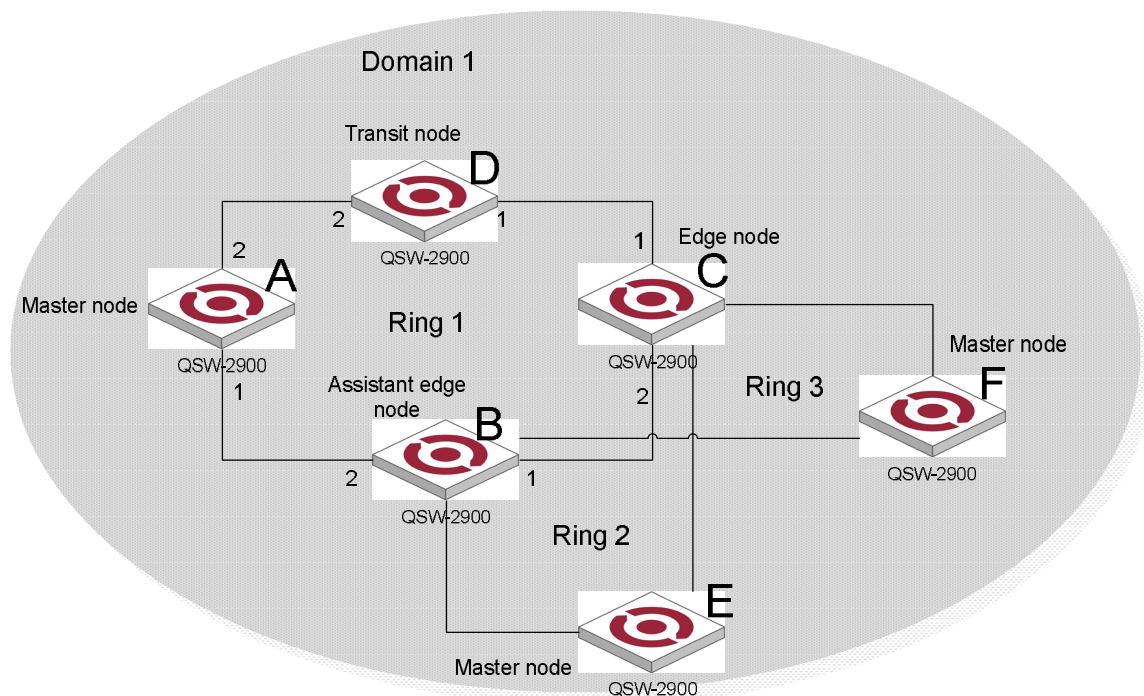


Figure 5 Dual homed rings

There are two or more rings in the network topology and two similar common nodes between rings. In this case, you only need to define an ERRP domain, and set one ring as the primary ring and other rings as subrings.

24.4.5 Multi-domain intersecting rings

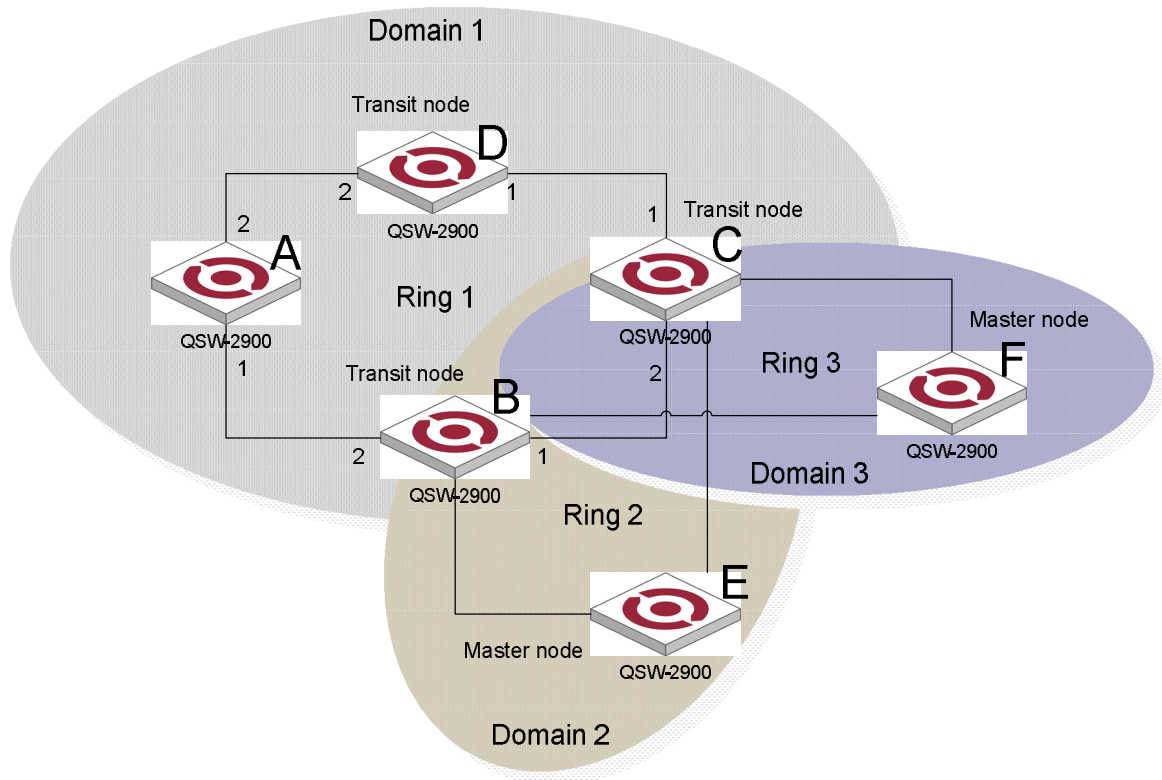


Figure 6 Multi-domain intersecting rings

There are two or more domains in a network, and there two different common nodes between any two domains. Figure 6 defines three ERRP domains, each containing one and only one ERRP primary ring. In the case of multi-domain intersection, the rings in different domains are independently configured. Each single domain can contain multiple rings, among which there must be one and only one primary ring. The data VLAN in one domain must be isolated from the data VLAN in another.

24.5 How ERRP Works

24.5.1 Polling mechanism

The primary port of the master node sends Health packets across the control VLAN periodically.

- If the ring works properly, the secondary port of the master node will receive Health packets and the master node will maintain it in block state.
- If the ring is torn down, the secondary port of the master node will not receive Health packets after the timeout timer expires. The master node will release the secondary port from blocking data VLAN while sending Common-Flush-FDB packets to notify all transit nodes to update their own MAC entries and ARP entries.

24.5.2 Link down alarm mechanism

The transit node, the edge node or the assistant edge node sends Link-Down packets to the master node immediately when they find any port belonging to an ERRP domain is down. Upon the receipt of a Link-Down packet, the master node releases the secondary port from blocking data VLAN while sending Common-Flush-FDB packet to notify all the transit nodes, the edge nodes and the assistant nodes to update their own MAC entries and

ARP entries.

24.5.3 Ring recovery

The master node may find the ring is restored after a period of time after the ports belonging to the ERRP domain on the transit node, the edge node or the assistant edge node are up again. A temporary loop may arise in the data VLAN in this period. As a result, broadcast storm occurs.

To prevent temporary loops, non-master nodes block them immediately (and permits only the packets of the control VLAN) when they find their ports accessing the ring are up again. The blocked ports are activated only when the nodes ensure that no loop will be brought forth by these ports.

24.5.4 Broadcast storm suppression mechanism in a multi-homed subring in case of primary ring link failure

As shown in Figure 5, Ring 1 is the primary ring, and Ring 2 and Ring 3 are subrings. When two links of the primary ring between the edge node and the assistant edge node are down, the master nodes of Ring 2 and Ring 3 will open their respective secondary ports, and thus a loop among B, C, E and F is generated. As a result, broadcast storm occurs.

In this case, to prevent from generating this loop, the edge node will block the edge port temporarily. The blocked edge port is activated only when the edge node ensures that no loop will be brought forth when the edge port is activated.

24.5.5 Protocols and Standards

Related standard : RFC 3619.

24.6 ERRP Configuration

24.6.1 ERRP Configuration list

Only when ERRP and ring enable, the configuration can be effective.theconfiguration will be reserved when ERRP and ring disable and it will be effective when ERRP and ring enable next time.

- ERRP configuration
- Configure ERRP timer
- Enter ERRP configuration mode
- Create ERRP ring
- Enable/disable ERRP ring

24.6.2 ERRP configuration

Configure it in global configuration mode

errp

Disable ERRP :

no errp

It is defaulted to disable ERRP.

For example :

! Enable ERRP

QTECH(config)#errp

24.6.3 Configure ERRP timer

Configure it in global configuration mode :
Configure packet overtime

errp fail-timer *timer-value*

Parameter :

timer-value : integrity in the range of 1-10
Configure packet sending interval

errp hello-timer *timer-value*

Parameter :

timer-value : integrity in the range of 1-10

For example :

! Configure ERRP packet sending interval to be 1 second

QTECH(config)#errp hello-timer 1

24.6.4 Enter ERRP configuration mode

Configure it in global configuration mode :

errp domain *domain-id*

Parameter :

domain-id : ERRP domain id

For example :

! Configure ERRP domain 0

QTECH(config)#errp domain 0

24.6.5 Configure control-vlan of ERRP domain

Configure it in ERRP domain mode :

control-vlan *vlan-id*

no control-vlan

Parameter :

vlan-id : control vlan id of ERRP domain which is the integrity in the range of 1-4093.

Note :

Control VLAN is relative to data VLAN. Data VLAN is for transmitting data packet and control VLAN is only for transmitting ERRP protocol packet. Every ERRP domain owns two control VLANs, that are master control VLAN and sub-control VLAN. Protocol packet of master ring is transmitted in master control-VLAN and protocol packet of sub-ring is transmitted in sub-control VLAN. When configuring, specify master control VLAN, and sub-control VLAN is the one whose VLAN ID is 1 bigger than that of the master control VLAN.

Port only accessing to Ethernet ring (ERRP port) of each switch belong to control VLAN. ERRP port of master ring belong to both master control VLAN and sub-control VLAN. ERRP port of sub-ring belongs to sub-control VLAN only. There can be ERRP port and non- ERRP port in data VLAN. Master ring is taken as a logical node of sub-ring. The protocol packet of sub-ring is transparently transmitted through master ring and handled as data packet in master ring. The protocol packet of master ring can only be transmitted in master ring.

Add all ERRP port to corresponded master and sub-control VLAN before or after handed down ERRP configuration and configure master and sub-control VLAN being tag vlan.

Example :

! Configure control VLAN of ERRP domain 0 being 25

QTECH(config-errp-0)#control-vlan 25

! Delete control VLAN of ERRP domain 0. if there is activated ring, the control VLAN will not allow to be deleted.

QTECH(config-errp-0)#no control-vlan

24.6.6 Create ERRP ring

Configure it in ERRP configuration mode :
Create master role

ring *ring-id* **role master primary-port** *pri-port* **secondary-port** *sec-port* **level** *level*
Create transit role

ring *ring-id* **role transit primary-port** *pri-port* **secondary-port** *sec-port* **level** *level*
Create edge role

ring *ring-id* **role edge common-port** *common-port* **edge-port** *edge-port*
Create Create assistant-edge role

ring *ring-id* **role assistant-edge common-port** *common-port* **edge-port** *edge-port*
Parameter :

ring-id : ring id which is in the range of 0-15

pri-port : port id such as ethernet 0/0/1

sec-port : port id such as ethernet 0/0/1

common-port : port id such as ethernet 0/0/1

sec-port : port id such as ethernet 0/0/1

level : ring level. 0 means primary ring and 1 means secondary.

For example :

! Configure primary ring 0 with role mode being master, primary port being 1 and secondary port being 2

QTECH(config-errp)#ring 0 role master primary-port ethernet 0/0/1 secondary-port ethernet 0/0/2 level 0

24.6.7 Enable/disable ERRP ring

Configure it in ERRP configuration mode :

ring *ring-id* { enable | disable }

Parameter :

ring-id : ring id

enable : activate a ring

diable : inactivate a ring

For example :

! Enable ring 0

QTECH(config-errp)#ring 0 enable

24.6.8 Display ERRP domain and ring information

Display in any configuration :

show errp [domain domain-id [ring ring-id]]

Parameter :

domain-id : domain id

ring-id : ring id

Example :

! Display ring 1 of ERRP domain 0

QTECH(config)#show errp domain 0 ring 1

Chapter 25 PPPoE Plus Configuration

25.1 Brief Introduction of PPPoE Plus

PPPoE+ is short for PPPoE Intermediate agent which is proposed early in DSL FORUM to define according to user line mark proportion of RFC 3046. The realization theory is similar to DHCP Option82 which makes some complement on PPPoE protocol packet. After accessing device get PPPoE protocol packet, insert user physical information for uplink direction and strip it for downlink direction before transmission.

This solution is designed for the PPPoE access method and is based on the Access Node implementing a PPPoE intermediate agent function in order to insert access loop identification. This functionality is described in the following.

The PPPoE Intermediate Agent intercepts all upstream PPPoE discovery stage packets, i.e. the PADI, PADR and upstream PADT packets, but does not modify the source or destination MAC address of these PPPoE discovery packets. Upon reception of a PADI or PADR packet sent by the PPPoE client, the Intermediate Agent adds a PPPoE TAG to the packet to be sent upstream. The TAG contains the identification of the access loop on which the PADI or PADR packet was received in the Access Node where the Intermediate Agent resides. If a PADI or PADR packet exceeds 1500 octets after adding the TAG containing the access loop identification, the Intermediate Agent must not send the packet to the Broadband Network Gateway. In response to the received PADI or PADR packet, the PPPoE Intermediate Agent should issue the corresponding PADO or PADS response with a Generic-Error TAG to the sender.

This is format of PPPoE TAG (type standard) on the QSW-3900 :

0 0/0/0 : 4096.VID Switch MAC/0/0/slot/sub-slot/port

Specially for HUAWEI BRAS connectivity has a type huawei of PPPoE TAG :

0 0/0/0 : 4096.VID Switch MAC/Hostname/0/slot/sub-slot/port

25.2 PPPoE Plus Configuration

25.2.1 PPPoE Plus Configuration list

PPPoE Plus Configuration list is as following :

- Enable/disable global PPPoE Plus
- Choose the type of PPPoE TAG

25.2.2 Enable/disable PPPoE Plus

Configure it in global configuration mode :

Enable global PPPoE Plus

pppoeplus

Disable global PPPoE Plus

no pppoeplus

By default, PPPoE Plus is disabled.

Example :

! Enable global PPPoE Plus

QTECH(config)#pppoeplus

To display PPPoE Plus, configure it in any configuration mode :

Display PPPoE Plus

show pppoeplus

25.2.3 Configure PPPoE Plus type

Configure it in global configuration mode :

Configure PPPoE Plus type

pppoeplus type { standard | huawei }

The default type is standard. The adding tag form will include hostname information when the type is huawei.

 **Note :**

All PPPoE clients must be members of IP managed VLAN. Please refer to the “[Configure and manage VLAN](#)”

Chapter 26 CFM Configuration

26.1 Brief introduction of CFM

CFM (Connectivity Fault Management) is a point-to-point OAM protocol defined by IEEE 802.1ag standard which is used to manage failure of operating network, including continuity detection, loopback, tracer, trap alarm and remote failure alarm.

26.2 Connectivity fault management overview

Connectivity fault management (CFM) is a link layer OAM (Operations, Administration and Maintenance) mechanism used for link connectivity detection and fault location.

26.3 Basic Concepts in Connectivity Fault Detection

26.3.1 Maintenance domain

A maintenance domain (MD) is the part of network where CFM plays its role. The MD boundary is defined by some maintenance points configured on the ports. MD is identified by MD name and is divided into 8 levels, represented by integer 0 to 7. The bigger the number, the higher the level. A higher level MD can contain lower level MDs, but they cannot overlap. In other words, a higher level MD covers larger area than a lower level MD.

26.3.2 Maintenance association

Maintenance association (MA) is a set of maintenance points in a maintenance domain. It is identified in the form “MD name + MA name”.

MA works within a VLAN. Packets sent by the maintenance points in a MA carry the corresponding VLAN tag. A maintenance point can receive packets sent by other maintenance points in the same MA.

26.3.3 Maintenance point

A maintenance point (MP) is configured on a port and belongs to a MA. MP can be divided into two types : maintenance association end point (MEP) and maintenance association intermediate point (MIP).

26.3.3.1 MEP

Each MEP is identified by an integer called MEP ID. The MEPs define the range of MD. The MA and MD that MEPs belong to define the VLAN attribute and level of the packets sent by the MEPs. MEPs are divided into inbound MEP and outbound MEP.

In Figure 1, outbound MEPs are configured on the ports. In Figure 2, inbound MEPs are configured on the two ports.

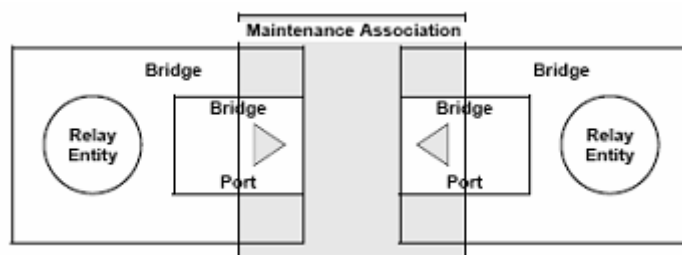


Figure 1 Outbound MEP

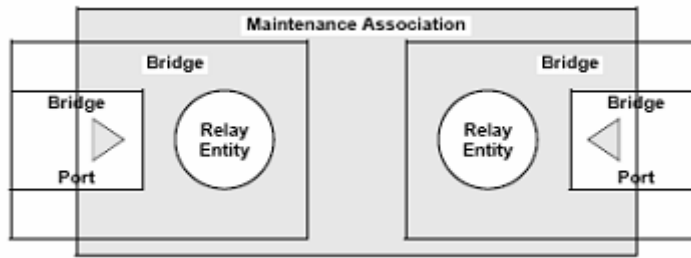


Figure 2 Inbound MEP

26.3.3.2 MIP

Maintenance association intermediate point (MIP) can handle and respond to CFM packets. The MA and MD that a MIP belongs to define the VLAN attribute and level of the packets received.

Figure 3 demonstrates a grading example of CFM module. In the figure, there are six devices, labeled as 1 to 6 respectively. Suppose each device has two ports, and MEPs and MIPs are configured on some of these ports. Four levels of MDs are designed in this example, the bigger the number, the higher the level and the larger the area covered. In this example, the X port of device 2 is configured with the following MPs : a level 5 MEP, a level 3 inbound MEP, a level 2 inbound MEP, and a level 0 outbound MEP.

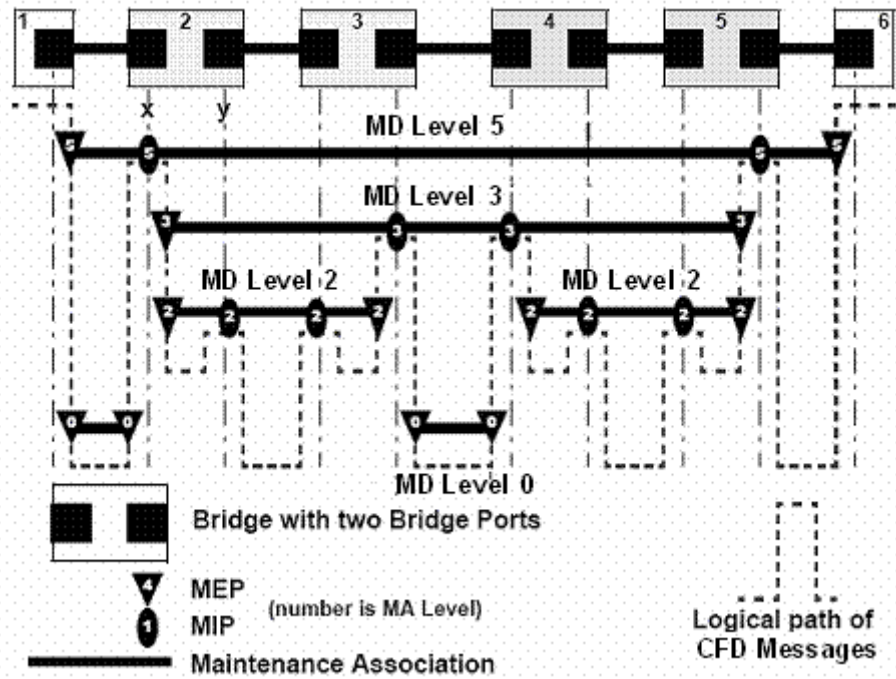


Figure 3 Levels of MPs

26.3.4 Basic Functions of Connectivity Fault Management

CFM works effectively only in well-deployed and well-configured networks. Its functions, which are implemented through the maintenance points, include :

- Continuity check (CC);
- Loopback (LB)
- Linktrace (LT)

26.3.4.1 Continuity check

Continuity check is responsible for checking connectivity between MEPs. Connectivity fault is usually caused by device fault or configuration error. This function is implemented through periodic sending of continuity

check messages (CCM) by the MEPs. As a multicast message, a CCM sent by one MEP is intended to be received by all the other MEPs in the same MA. If a MEP fails to receive the CCM within 3.5 sending periods, the link is regarded as faulty and a corresponding log is generated. When multiple MEPs send CCMs at the same time, the many-to-many link check is achieved.

26.3.4.2 Loopback

Loopback is responsible for verifying connectivity between a local device and a remote device. To implement this function, a local MEP sends a loopback message (LBM) to the remote MEP. Depending on whether the local MEP can receive loopback reply message (LBR) from the remote MEP, the link state between the two can be verified. LBM and LBR are unicast messages. They are used to verify the connectivity between two points.

26.3.4.3 Linktrace

Linktrace is responsible for identifying the path between the source MEP and the target MEP. This function is implemented in the following way : the source sends a linktrace message (LTM) to the target MEP. After receiving the message, the target MEP as well as the MIPs that the LTM passes send back linktrace reply message (LTR) to the source. Based on the replying messages, the source can identify the path to the target.

26.3.5 Protocols and Standards

The connectivity fault management function is implemented in accordance with IEEE P802.1ag.

26.4 CFM Configuration

26.4.1 CFM Configuration list

Configure domain before configuring other parameter when enabling CFM. CFM command list is as following :

- Configure cfm domain
- Configure cfm mep level
- Configure cfm mip level
- Configure remote cfm rmep level
- Configure cfm cc interval
- Enable/disable VLAN sending cfm cc enable level
- cfm ping
- cfm traceroute
- Display cfm domain
- Display cfm maintenance-points local
- Display cfm maintenance-points remote
- Display cfm cc database
- Display cfm errors

26.4.2 Configure cfm domain

Configure it in global configuration mode :
Configure cfm domain

cfm domain *domain-name* **level** *level-id*

Parameter :
domain-name : CFM domain name
level-id : the integrity from 0-7
Remove cfm domain

no cfm domain level *level-id*

It is defaulted not to configure cfm domain.

For example :

! Configure cfm domain customer level 7

QTECH(config)#cfm domain customer level 7

26.4.3 Configure cfm mep level

Configure it in interface configuration mode :

Configure cfm mep level

cfm mep level *level-id* direction {*up* | *down* } mpid *mep-id* vlan *vlan-id*

Parameter :

level-id : the integrity from 0-7

up : direction of MEP

down : direction of MEP

mep-id : MEP id

vlan-id : VLAN of MEP

Delete cfm mep level

no cfm mep level *level-id* vlan *vlan-id*

It is defaulted not to configure cfm mep level.

For example :

! Configure cfm mep level 7 direction up mpid 7110 vlan 110

QTECH(config-if-ethernet-0/0/1)#cfm mep level 7 direction up mpid 7110 vlan 110

26.4.4 Configure cfm mip level

Configure it in interface configuration mode :

Configure cfm mip level

cfm mip level *level-id*

Parameter :

level-id : the integrity from 0-7

Delete cfm mip level

no cfm mip level *level-id*

It is defaulted not to configure cfm mip level

For example :

! Configure cfm mip level 7

QTECH(config-if-ethernet-0/0/1)#cfm mip level 7

26.4.5 Configure remote cfm rmep level

Configure it in global configuration mode :

Configure remote cfm rmep level

cfm rmep level *level-id* mpid *mep-id* vlan *vlan-id*

Parameter :

level-id : the integrity from 0-7

mep-id : MEP id

vlan-id : VLAN of MEP

Delete remote cfm rmep level

no cfm rmep level *level-id* mpid *mep-id* vlan *vlan-id*

It is defaulted not to configure remote cfm rmep level.

For example :

! Configure cfm rmep level 7 mpid 7110 vlan 110

QTECH(config)#cfm rmep level 7 mpid 7110 vlan 110

26.4.6 Configure cfm cc interval

Configure it in global configuration mode :

Configure cfm cc interval

cfm cc interval { 1 | 10 | 60 }

Parameter :

1 : sending interval is 1 second

10 : sending interval is 10 seconds

60 : sending interval is 60 seconds

Restore cfm cc interval

no cfm cc interval

The default cfm cc interval is 10s

For example :

! Configure cfm cc interval to be 1s

QTECH(config)#cfm cc interval 1

26.4.7 Enable/disable VLAN sending cfm cc enable level

Configure it in global configuration mode :

Enable VLAN sending cfm cc enable level

cfm cc enable level *level-list* vlan *vlan-list*

Parameter :

level-list : level list needed enabling

vlan-list : VLAN list needed enabling

Disable VLAN sending cfm cc enable level

no cfm cc enable level *level-list* vlan *vlan-list*

It is defaulted to enable VLAN sending cfm cc enable level.

For example :

! Configure cfm cc enable level 0-7 vlan 1-10

QTECH(config)#cfm cc enable level 0-7 vlan 1-10

26.4.8 cfm ping

cfm ping command is used to check network connection and the arrival of destination mac address.

Configure it in global configuration mode :

cfm ping [-c *count*] [-s *packetsize*] [-t *timeout*] *mac level level-id* **vlan** *vlan-id*

Parameter :

-c *count* : the number of sending packet.

-s *packetsize* : the length of sending packet which is in the unit of bit.

-t *timeout* : the response timeout after sending packet which is in the unit of seconds.

mac : the destination mac address needed ping.

level-id : the integrity from 0-7

vlan-id : the VLAN needed ping.

For example :

! cfm ping 00 : 1f : ce : 10 : 14 : f1 level 7 vlan 110

QTECH#cfm ping 00 : 1f : ce : 10 : 14 : f1 level 7 vlan 110

PING 001f:ce10.14f1 :

reply from 001f : ce10 : 14f1

reply from 001f : ce10 : 14f1

```

reply from 001f : ce10 : 14f1
reply from 001f : ce10 : 14f1
reply from 001f : ce10 : 14f1
5 packets transmitted, 5 packets received, 0.0% packet loss

```

26.4.9 cfm traceroute

cfm traceroute command is used for link tracet and checking network connection. Configure it in global configuration mode :

cfm traceroute [-f *first_ttl* | -h *maximum_hops* | -w *time_out*] *target-mac level level-id vlan vlan-id*

Parameter :

first_ttl : first ttl of sending [packet which](#) is in the range of 1 to 255 and default value is 255;

maximum_hops : max ttl of sending packet which is in the range of 1 to 255 and default value is 10;

time_out : the response timeout after sending packet which is in the range of 10 to 60 with the unit of second and default value is 5 seconds;

target_mac : destination mac address

level-id : the integrity from 0-7

vlan-id : VLAN to be tracetert

For example :

```
! cfm traceroute 00 : 1f : ce : 10 : 14 : f1 level 4 vlan 110
```

```
QTECH#cfm traceroute 00 : 1f : ce : 10 : 14 : f1 level 4 vlan 110
```

26.4.10 Display cfm domain

Configure it in any configuration mode :

It will display as following :

- cfm domain name
- cfm domain level

show cfm domain

For example :

```
! Display cfm domain
```

```
QTECH(config)#show cfm domain
```

26.4.11 Display cfm maintenance-points local

Configure it in any configuration mode :

It will display as following :

- cfm maintenance-points mpid
- cfm maintenance-points type
- cfm maintenance-points vlan
- cfm maintenance-points level
- cfm maintenance-points interface
- Enable/disable cfm maintenance-points

- cfm maintenance-points mac address

show cfm maintenance-points local

For example :

! Display cfm maintenance-points local

```
QTECH(config)# show cfm maintenance-points local
```

26.4.12 Display cfm maintenance-points remote

Configure it in any configuration mode :

It will display as following :

- cfm maintenance-points remote mpid
- cfm maintenance-points remote vlan
- cfm maintenance-points remote mac address
- cfm maintenance-points remote ingress interface
- cfm maintenance-points remote aging time

show cfm maintenance-points remote

For example :

! Display cfm maintenance-points remote

```
QTECH(config)# show cfm maintenance-points remote
```

26.4.13 Display cfm cc database

Configure it in any configuration mode :

It will display as following :

- Mac address
- vlan-id
- ingress interface

show cfm cc database

For example :

! Display cfm cc database

```
QTECH(config)# show cfm cc database
```

26.4.14 Display cfm errors

Configure it in any configuration mode :

It will display as following :

- cfm errors mpid
- cfm errors vlan
- cfm errors level
- cfm maintenance-points remote mac address
- error reason

show cfm errors

For example :

! Display cfm errors

```
QTECH(config)# show cfm error
```