QoS Configuration

Contents

Chapter 1 QoS Configuration

1.1 Brief Introduction to QoS

In traditional IP networks, packets are treated equally. That is, the FIFO (first in first out) policy is adopted for packet processing. Network resources required for packet forwarding is determined by the order in which packets arrive. All the packets share the resources of the network. Network resources available to the packets completely depend on the time they arrive. This service policy is known as Best-effort, which delivers the packets to their destination with the best effort, with no assurance and guarantee for delivery delay, jitter, packet loss ratio, reliability, and so on.

With the fast development of computer networks, more and more networks are connected into Internet. Users hope to get better services, such as dedicated bandwidth, transfer delay, jitter voice, image, important data which enrich network service resources and always face network congestion. Internet users bring forward higher requirements for QoS. Ethernet technology is the widest network technology in the world recently. 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 execute 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.

1.1.1 Traffic

Traffic means all packets through switch.

1.1.2 Traffic Classification

Traffic classification is to identify packets conforming to certain characters according to certain rules. It is the basis and prerequisite for proving differentiated services. A traffic classification rule can use the precedence bits in the type of service (ToS) field of the IP packet header to identify traffic with different precedence characteristics. A traffic classification rule can also classify traffic according to the traffic classification policy set by the network administrator, such as the combination of source address, destination address, MAC address, IP protocol, or the port numbers of the application. Traffic classification is generally based on the information in the packet header and rarely based on the content of the packet.

1.1.3 Priority

(1) 802.1p priority lies in Layer 2 packet headers and is applicable to occasions where the Layer 3 packet header does not need analysis but QoS must be assured at Layer 2. As shown in the chapter of VLAN configuration. Each host supported 802.1Q protocol forwards packets which are from Ethernet frame source address add a 4-byte tag header. Figure as 1-1.

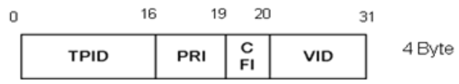

Figure 1-1 802.1Q tag

As shown in the figure above, PRI segment is 802.1p priority. It consists of 3bits whose range from 0~7. The three bits point the frame priority. The tag including 8 formats gives the precedence to forward the packets.

Table 1-1 Description on 802.1Q values

cos (decimal)	cos (binary)	Description
0	000	spare
1	001	background
2	010	best-effort
3	011	excellent-effort
4	100	controlled-load
5	101	video
6	110	voice
7	111	network-management

(2)IP precedence, TOS precedence, and DSCP values

The TOS field in the IP header contains eight bits: the first three bits represent IP precedence; the subsequent four bits represent a ToS value and 1 bit with currently unused defaults 0. The four bits of TOS packets are grouped into four classes: the smallest time delay, maximum rate, highly reliability, minimum cost. Only 1 bit can be set, if the DSCP values equal 0, that means normal service.

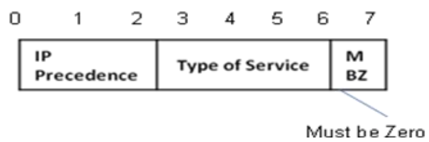

Figure 1-2 IP precedence and TOS precedence

IP precedence contains 8 formats.

Table 1-2 Description on IP Precedence

IP Precedence (decimal)	IP Precedence (binary)	Description
0	000	routine
1	001	priority
2	010	immediate
3	011	flash
4	100	flash-override
5	101	critical
6	110	internet

7	111	network

TOS precedence contains 5 formats.

Table 1-3 Description on TOS Precedence

TOS (decimal)	TOS (binary)	Description
0	0000	normal
1	0001	min-monetary-cost
2	0010	max-reliability
4	0100	max-throughput
8	1000	min-delay

According to RFC 2474, the ToS field is redefined as the differentiated services (DS) field, where a DSCP value is represented by the first six bits (0 to 5) and ranges from 0 to 63. The remaining two bits (6 and 7) are reserved.

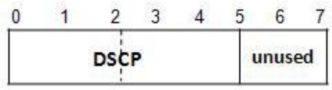

Figure 1-3 DSCP values

In a network in the Diff-Serve model, traffic is grouped into the following classes, and packets are processed according to their DSCP values

- Expedited forwarding (EF) class: In this class, packets are forwarded regardless of link share of other traffic. The class is suitable for preferential services requiring low delay, low packet loss, low jitter, and high bandwidth.
- Assured forwarding (AF) class: This class is divided into four subclasses (AF 1 to AF 4), each containing three drop priorities for more granular classification. The QoS level of the AF class is lower than that of the EF class.
- Class selector (CS) class: This class is derived from the IP ToS field and includes eight subclasses.
- Best effort (BE) class: This class is a special CS class that does not provide any assurance. AF traffic exceeding the limit is degraded to the BE class. All IP network traffic belongs to this class by default.

Table 1-4 Description on DSCP values

DSCP (decimal)	DSCP (binary)	keys
0	000000	be
46	101110	ef
10	001010	af1
18	010010	af2
26	011010	af3
34	100010	af4
8	001000	cs1
16	010000	cs2
24	011000	cs3
32	100000	cs4
40	101000	cs5
48	110000	cs6
56	111000	cs7

1.1.4 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.

1.1.5 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 doubtable 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.

1.1.6 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.

1.1.7 Interface Speed Limitation

Interface speed limitation is the speed limit based on interface which limits the total speed rate of interface outputting packet.

1.1.8 Redirection

User can re-specify the packet transmission interface based on the need of its own QoS strategies.

1.1.9 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.

1.1.10 Choose Interface Outputting Queue for Packet

Ethernet switch can choose corresponding outputting queue for specified packets.

1.1.11 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.

(1)PQ

PQ (Priority Queuing) is designed for key service application. Key service possesses an important feature, that is, require the precedent 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 scheduler, 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.

(2)WRR

WRR queue scheduler divides a port into 4 or 8 outputting queues (S2926V-O has 4 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 shortage 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) SP+ WRR

Superior priority or less priority use SP algorithm, others use WRR algorithm.

1.1.12 Cos-map 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.

1.1.13 Flow Mirror

Flow mirror means coping specified data packet to monitor interface to detect network and exclude failure.

1.1.14 Statistics Based on Flow

Statistics based on flow can statistic and analyze the packets customer interested in.

1.1.15 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.

1.2 QOS Configuration

1.2.1 Configuring Flow Monitor

Flow monitor is restriction to flow rate which can monitor the speed of a flow entering switch. If the flow is beyond specified specification, it will take actions, such as dropping packet or reconfigure their priority.

Table 1-5 Configure flow rate

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Enter port configuration mode	**interface ethernet** device/slot/port	optional, perform either of the globally and port mode
Configure flow rate	**rate-limit** { **input** | **output** } { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] } *target-rate*	optional

1.2.2 Configure Two Rate Three Color Marker

Two Rate Three Color Marker is defined in RFC 2698. There is 4 parameter for it: CIR, CBS, PIR and PBS.

Table 1-1 Configure Two Rate Three Color Marker

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Configure Two Rate Three Color Marker	**two-rate-policer** *policer-id* **cir** *cir* **cbs** *cbs* **pir** *pir* **pbs** *pbs* [**color-aware**] [**drop-red**]	optional
Enter port configuration mode	**interface ethernet** device/slot/port	optional, perform either of the globally and

		port mode
Configure Two Rate Three Color Marker	**rate-limit** { **input** | **output** } { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] } **two-rate-policer** *policer-id*	optional

1.2.3 Configuring Interface Line Rate

Line-limit is the speed limit based on interface which restricts the total speed of packet outputting.

Table 1-6 Configure interface line rate

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Enter port configuration mode	**interface ethernet** *device/slot/port*	-
Configure egress rate	**bandwidth egress** *target-rate*	optional
Configure ingress rate	**bandwidth ingress** *target-rate*	optional

1.2.4 Configuring Packet Redirection

Packet redirection configuration is redirecting packet to be transmitted to some egress.

Table 1-7 Configure interface line rate

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Configure packet redirection	**traffic-redirect** { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] } { [**interface** *interface-num* | **cpu**] }	optional

1.2.5 Configuring Traffic Copy to CPU

Switch automatically copies to CPU after configuring traffic copy to CPU.

Table 1-8 Configure traffic copy to CPU

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Configure traffic copy to CPU	**traffic-copy-to-cpu** { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] }	optional

1.2.6 Configuring Traffic Priority

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 reflects priority in packet head.

Table 1-9 Configure traffic priority

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Configure traffic priority	**traffic-priority** { [**ip-group** { *num* | *name*} [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] } { [**dscp** *dscp-value*] [**cos** { *pre-value* | **from-ipprec** }] [**local-precedence** *pre-value*] }	optional

1.2.7 Configuring Queue-Scheduler

When network congestion, it must use queue-scheduler to solve the problem of resource competition. System supports 3 kinds of queue-scheduler, that is SP, WRR and full SP+WRR.

By default is SP in system.

Table 1-10 Configure queue-scheduler

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Configure SP	**queue-scheduler strict-priority**	optional
Configure WRR	**queue-scheduler wrr** *queue1-weight queue2-weight queue3-weight queue4-weight*	optional
Configure SR+WRR	**queue-scheduler sp-wrr** *queue1-weight queue2-weight queue3-weight*	optional

1.2.8 Configuring Cos-map Relationship of Hardware Priority Queue and Priority of IEEE802.1p Protocol

The cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol is one - to - one correspondence. Administrators change the cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol timely when the one-to-one correspondence shifting.

By default, the cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol as below:

Table 1-11 802.1p and he cos-map relationship of hardware priority queue

802.1p	hardware priority queue
0	0

1	0
2	1
3	1
4	2
5	2
6	3
7	3

Administrators also change the cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol according to the actual network.

Table 1-12 Modifying 802.1p and he cos-map relationship of hardware priority queue

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Modify 802.1p and he cos-map relationship of hardware priority queue	**queue-scheduler cos-map** *queue-number packed-priority*	optional

1.2.9 Configuring Mapping Relationship between DSCP and 8 Priority in IEEE 802.1p

The same situation as 1.2.7, by default, the relation between DSCP and 8 priority in IEEE 802.1p as below;

Table 1-13 Relation between DSCP and 8 priority in IEEE 802.1p

SCP	hardware priority queue	DSCP	hardware priority queue	DSCP	hardware priority queue	DSCP	hardware priority queue
0	0	16	1	32	2	48	3
1	0	17	1	33	2	49	3
2	0	18	1	34	2	50	3
3	0	19	1	35	2	51	3
4	0	20	1	36	2	52	3
5	0	21	1	37	2	53	3
6	0	22	1	38	2	54	3
7	0	23	1	39	2	55	3
8	0	24	1	40	2	56	3
9	0	25	1	41	2	57	3
10	0	26	1	42	2	58	3
11	0	27	1	43	2	59	3
12	0	28	1	44	2	60	3
13	0	29	1	45	2	61	3
14	0	30	1	46	2	62	3
15	0	31	1	47	2	63	3

Administrators also change the mapping relationship between DSCP and 8 priority in IEEE 802.1p according to the actual network.

Table 1-14 Configuring the relation between DSCP and 8 priority in IEEE 802.1p

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Startup the relation between DSCP and 8 priority in IEEE 802.1p	**queue-scheduler dscp-map**	required by default, it is disable.
Modify the relation between DSCP and 8 priority in IEEE 802.1p	**queue-scheduler dscp-map** *dscp-value 802.1p-priority*	optional

1.2.10 Configuring Flow Statistic

Flow statistic configuration is used to statistic specified service flow packet. The statistic is accumulated value and reset to zero when re-configuring.

Table 1-16 Configure flow statistic

Operation	Command	remark
Enter globally configuration mode	**configure terminal**	-
Configure flow staticstic	**traffic-statistic** { [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]] }	optional
reset to Zero	**clear traffic-statistic** { [**all** | [**ip-group** { *num* | *name* } [**subitem** *subitem*]] [**link-group** { *num* | *name* } [**subitem** *subitem*]]] }	optional

1.2.11 Configuring Flow Mirror

Flow mirror is copying the service flow which matches ACL rules to specified monitor interface to analyze and monitor packet.

1.2.12 Displaying and Maintain QoS

After finishing above configuration, please use below commands to show the configuration.

Table 1-17 Display and maintain QoS

Operation	Command	remark
Display all the informaion of QoS	**show qos-info all**	perform either of the commands
Display QoS statistic	**show qos-info statistic**	
Display quue-scheduler mode and parameters	**show queue-scheduler**	
Display the cos-map relationship of hardware priority queue and priority of IEEE802.1p protocol	**show queue-scheduler cos-map**	
Display the dscp-map relationship of hardware priority queue and priority of IEEE802.1p protocol	show queue-scheduler dscp-map	
Display all QoS port configuration	**show qos-interface [***interface-num* **] all**	
Display rate-limit parameters	**show qos-interface [***interface-num* **] rate-limit**	
Display interface line rate parameters	**show bandwidth-control interface Ethernet** [interface-num]	
Display QoS interface statistic parameters	**show qos-interface statistic**	
Display traffic-priority parameters	**show qos-info traffic-priority**	
Display traffic-redirect parameters	**show qos-info traffic-redirect**	
Display packet redirection	**show qos-info traffic-statistic**	
Display information of traffic copy to CPU	**show qos-info traffic-copy-to-cpu**	