QOS Policing and Marking with Catalyst 4000/4500 IOS based Supervisor Engines

Document ID: 21263

Contents

Introduction

Policing function determines if the traffic level is within the specified profile (contract). Policing function allows either dropping out−of−profile traffic or marking the traffic down to a different Differential Services Code Point (DSCP) value to enforce contracted service level. DSCP is a measure of the Quality of Service (QoS) level of the packet. Along with DSCP, IP precedence and Class of Service (CoS) are also used to convey the QoS level of the packet.

Policing should not be confused with traffic shaping, although both ensure that traffic stays within the profile (contract). Policing does not buffer the traffic, so the transmission delay is not affected. Instead of buffering out−of−profile packets, policing will drop them or mark them with a different QoS level (DSCP mark down). Traffic shaping buffers out−of−profile traffic and smoothes traffic bursts, but affects the delay and delay variation. Shaping can only be applied on an outgoing interface, while policing can be applied on both incoming and outgoing interfaces.

Catalyst 4000/4500 with Supervisor Engine 3, 4 and 2+ (SE3, SE4, SE2+ from now on in this document) supports policing in incoming and outgoing directions. Traffic shaping is also supported, however this document will only deal with policing and marking. Marking is a process of changing the packet QoS level according to a policy.

Prerequisites

Requirements

There are no specific requirements for this document.

Components Used

This document is not restricted to specific software and hardware versions.

QoS Policing and Marking Parameters

Policing is set up by defining the QoS policy maps and applying them to ports (port−based QoS) or to VLANs (VLAN−based QoS). The policer is defined by rate and burst parameters, as well as actions for in−profile and out−of−profile traffic.

There are two types of policers supported: aggregate and per−interface. Each policer can be applied to several ports or VLANs.

The aggregate policer acts upon the traffic across all of the applied ports/VLANs. For example, we apply the aggregate policer to limit the Trivial File Transfer Protocol (TFTP) traffic to 1 Mbps on VLANs 1 and 3. Such a policer will allow 1 Mbps of TFTP traffic in VLANs 1 and 3 together. If we apply a per−interface policer, it will limit the TFTP traffic on VLANs 1 and 3 to 1 Mbps each.

Note: If both ingress and egress policing is applied to a packet, the most severe decision will be taken. That is, if the ingress policer specifies to drop the packet and the egress policer specifies to mark the packet down, the packet will be dropped. Table 1 summarizes the QoS action upon the packet when it is treated by both ingress and egress policies.

Table 1: QoS Action Depending on Ingress and Egress Policy

	Ingress policy			
Egress policy	**Transmit**	**Drop**	**Markdown$_i$**	**Mark$_i$**
Transmit	Transmit	Drop	Markdown$_i$	Mark$_i$
Drop	Drop	Drop	Drop	Drop
Markdown$_e$	Markdown$_e$	Drop	Markdown$_e$	Markdown$_e$
Mark$_e$	Mark$_e$	Drop	Mark$_e$	Mark$_e$

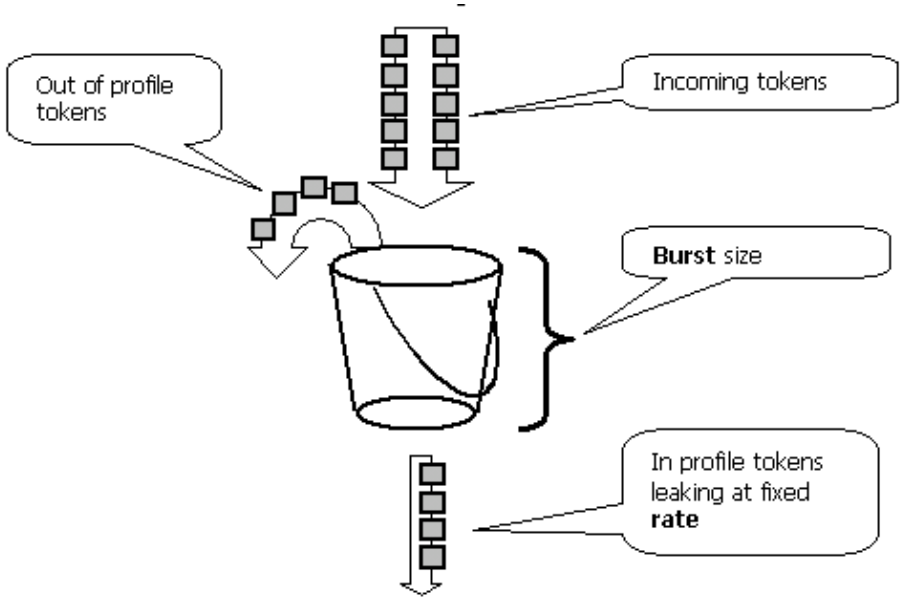
Catalyst 4000 SE3, SE4, SE2+ QoS hardware is implemented in such a way that real marking of the packet occurs after the egress policer. This means that even if the ingress policy remarks the packet (by policer mark down or normal marking), the egress policy will still see packets marked with the original QoS level. The egress policy will see the packet as if they had not been marked by the ingress policy. This means the following:

- Egress marking overrides ingress marking.
- Egress policy cannot match new QoS levels that are changed by ingress marking.

Other important implications are as follows:

- It is not possible to do marking and mark down within the same traffic class within the same policy.
- Aggregate policers are per−direction. That is, if an aggregate policer is applied to both ingress and egress, there will be two aggregate policers, one on input and one on output.
- When an aggregate policer is applied within the policy to the VLANs and to the physical interface, there effectively will be two aggregate policers − one for the VLAN interfaces and another for physical interfaces. Currently, it is not possible to police VLAN interfaces and physical interfaces together on aggregate.

Policing in Catalyst 4000 SE3, SE4, SE2+ complies with leaky bucket concept, as the model below illustrates. Tokens corresponding to incoming traffic packets are placed into a bucket (# of tokens = size of the packet). At regular intervals, a defined number of tokens (derived from the configured rate) is removed from the bucket. If there is no place in the bucket to accommodate an incoming packet, the packet is considered out−of−profile and dropped or marked down, according to the configured policing action.



It should be noted that the traffic is not buffered in the bucket, as it might appear from the above model. The actual traffic is not flowing via the bucket at all. The bucket is only used to decide whether the packet is in−profile or out−of−profile.

Note that exact hardware implementation of policing could be different, functionally it complies to the above model.

The following parameters control the operation of policing:

- Rate defines how many tokens are removed at each interval. This effectively sets the policing rate. All traffic below the rate is considered in−profile.
- Interval defines how often tokens are removed from the bucket. Interval is fixed at 16 nanoseconds (16 sec *10−9). Interval cannot be changed.
- Burst defines the maximum amount of tokens the bucket can hold at any time.

Refer to the Comparing Policing and Marking on Catalyst 6000 and Catalyst 4000/4500 IOS Based Supervisor Engines section at the end of this document for differences in burst between Catalyst 6000 and Catalyst 4000 SE3, SE4, SE2+.

The policer ensures that if you examine any period of time (from zero to infinity) that the policer will never allow more than

```
<rate> * <period> + <burst−bytes> + <1 packet> bytes
```

of traffic through the policer during that period.

Catalyst 4000 SE3, SE4, SE2+ QoS hardware has certain granularity for the policing. Depending on the rate configured, the maximum deviation from the rate is 1.5% of the rate.

When configuring burst rate, you need to take into account that some protocols (such as TCP) implement flow−control mechanisms that react on packet loss. For example, TCP reduces the window by half for each lost packet. When policed to a certain rate, the effective link utilization will be lower than the configured rate. One can increase the burst in order to achieve better utilization. A good start for such traffic would be to set the burst to be equal to twice the amount of traffic sent with desired rate during Round−Trip Time (RTT). For the same reason, it is not recommended to benchmark the policer operation by connection−oriented traffic, as it will generally show lower performance than permitted by the policer.

Note: Connectionless traffic might also react to policing differently. For example, Network File System (NFS) uses blocks, which might consist of more than one User Datagram Protocol (UDP) packet. One packet dropped might trigger many packets (entire block) to be retransmitted.

For example, the following is a calculation of the burst for a TCP session, with a policing rate of 64 Kbps and a TCP RTT of 0.05 seconds:

```
<burst> = 2* <RTT> * <rate> = 2 * 0.05 [sec] * 64000/8 [bytes/sec] = 800 [bytes]
```

Note: <burst> is for one TCP session, so it should be scaled to average the expected number of sessions going via the policer. This is an example only, so in each case, one needs to evaluate the traffic/application requirements and behavior versus the available resources in order to choose policing parameters.

The policing action is either to drop the packet (drop) or change the DSCP of the packet (mark down). In order to mark down the packet, the policed DSCP map must be modified. The default policed DSCP remarks the packet to the same DSCP, that is, no mark down occurs.

Note: Packets might be sent out−of−order when an out−of−profile packet is marked down to a DSCP to a different output queue than the original DSCP. For this reason, if ordering of packets is important, it is recommended to mark down out−of−profile packets to DSCP mapped to the same output queue as in−profile packets.

Policing and Marking Features Supported by Catalyst 4000/4500 IOS−Based Supervisor Engines

Both ingress (incoming interface) and egress (outgoing interface) policing are supported on Catalyst 4000 SE3, SE4, SE2+. The switch supports 1024 ingress and 1024 egress policers. Two ingress and two egress policers are used by the system for default no−policing behavior.

Note that when the aggregate policer is applied within the policy to a VLAN and a physical interface, an additional hardware policer entry is used. Currently, it is not possible to police VLAN interfaces and physical interfaces together on aggregate. This might be changed in future software releases.

All software versions include support for policing. The Catalyst 4000 supports up to 8 valid match statement per class, and up to 8 classes are supported per policy−map. Valid match statements are as follows:

- match access−group
- match ip dscp
- match ip precedence
- match any

Note: For non−IP V4 packets, the **match ip dscp** statement is the only way of classification, provided the packets are coming into trunking ports trusting CoS. Do not be misled by the keyword ip in the command **match ip dscp**, because internal DSCP is matched this applies to all packets, not just IP. When a port is configured to trust CoS, the latter is extracted from the L2 (802.1Q or ISL tagged) frame and converted to internal DSCP using a CoS to DSCP QoS map. This internal DSCP value can then be matched in the policy using **match ip dscp**.

Valid policy actions are as follows:

- police
- set ip dscp
- set ip precedence

- trust dscp
- trust cos

Marking allows changing the QoS level of the packet based upon classification or policing. Classification splits traffic into different classes for QoS processing based on defined criteria. In order to match IP precedence or DSCP, corresponding incoming interface should be set to the trusted mode. The switch supports trusting CoS, trusting DSCP, and untrusted interfaces. Trust specifies the field from which QoS level of the packet will be derived.

When trusting CoS, the QoS level will be derived from the L2 header of the ISL or 802.1Q encapsulated packet. When trusting DSCP, the switch will derive the QoS level from the DSCP field of the packet. Trusting CoS is only meaningful on trunking interfaces, and trusting DSCP is valid for IP V4 packets only.

When an interface is not trusted (this is default state when QoS is enabled), internal DSCP will be derived from the configurable default CoS or DSCP for the corresponding interface. If no default CoS or DSCP is configured, the default value will be zero (0). Once the original QoS level of the packet is determined, it is mapped into internal DSCP. Internal DSCP can be retained or changed by marking or policing.

After the packet undergoes QoS processing, the QoS level fields (within the IP DSCP field for IP and within the ISL/802.1Q header, if any) will be updated from internal DSCP.

There are special maps used to convert the trusted QoS metrics of the packet to internal DSCP and vice versa. These maps are as follows:

- DSCP to Policed DSCP; used to derive policed DSCP when marking down the packet.
- DSCP to CoS: used to derive the CoS level from internal DSCP to update the outgoing packet ISL/802.1Q header.
- CoS to DSCP: used to derive internal DSCP from incoming CoS (ISL/802.1Q header) when the interface is in trust CoS mode.

Note when an interface is in trust CoS mode, the outgoing CoS will always be the same as incoming CoS. This is specific to QoS implementation in Catalyst 4000 SE3, SE4, SE2+.

Configuring and Monitoring Policing

Configuring policing in IOS involves the following steps:

1. Defining a policer.
2. Defining criteria to select traffic for policing.
3. Defining service−policy using class and applying a policer to a specified class.
4. Applying a service−policy to a port or VLAN.

Consider the following example. There is a traffic generator attached to port 5/14 sending ~17 Mbps of UDP traffic with a destination of port 111. We want this traffic to be policed down to 1 Mbps and excessive traffic should be dropped.

```
! enable qos
qos
! define policer, for rate and burst values, see 'policing parameters
qos aggregate−policer pol_1mbps 1mbps 1000 conform−action transmit
exceed−action
drop
! define ACL to select traffic
access−list 111 permit udp any any eq 111
! define traffic class to be policed
```

```
class-map match-all cl_test
match access-group 111
! define QoS policy, attach policer to traffic class
policy-map po_test
class cl_test
police aggregate pol_1mbps
! apply QoS policy to an interface
interface FastEthernet5/14
switchport access vlan 2
! switch qos to vlan-based mode on this port
qos vlan-based
! apply QoS policy to an interface
interface Vlan 2
service-policy output po_test
!
```

Note that when a port is in VLAN based QoS mode, but no service policy is applied to the corresponding VLAN, the switch will follow the service policy (if any) applied on a physical port. This allows additional flexibility in combining port−based and VLAN−based QoS.

There are two types of policers supported: named aggregate and per−interface. A named aggregate policer will police the traffic combined from all interfaces to which it is applied. The example above used a named policer. A per−interface policer, unlike a named policer, will police traffic separately on each interface where it is applied. A per−interface policer is defined within the policy map configuration. Consider the following example with a per−interface aggregate policer:

```
! enable qos
qos
! define traffic class to be policed
class-map match-all cl_test2
match ip precedence 3 4
! define QoS policy, attach policer to traffic class
policy-map po_test2
class cl_test2
! per-interface policer is defined inside the policy map
police 512k 1000 conform-action transmit exceed-action drop
interface FastEthernet5/14
switchport
! set port to trust DSCP − need this to be able to match to incoming IP precedence
qos trust dscp
! switch to port-based qos mode
no qos vlan-based
! apply QoS policy to an interface
service-policy input po_test2
```

The following command is used to monitor policing operation:

```
Yoda#show policy-map interface FastEthernet5/14
FastEthernet5/14
service-policy input: po_test2
class-map: cl_test2 (match-all)
7400026 packets
match: ip precedence 3 4
police: Per-interface
Conform: 1166067574 bytes Exceed: 5268602114 bytes
class-map: class-default (match-any)
13312 packets
match: any
13312 packets
Yoda#show policy-map interface FastEthernet5/14
FastEthernet5/14
service-policy input: po_test2
class-map: cl_test2 (match-all)
```

```
7400138 packets
match: ip precedence 3 4
police: Per-interface
Conform: 1166088574 bytes Exceed: 5268693114 bytes
class-map: class-default (match-any)
13312 packets
match: any
13312 packets
```

The counter near class−map is counting the number of packets matching to corresponding class.

Be aware of the following implementation specific considerations:

- The per−class packet counter is not per−interface. That is, it counts all packets matching the class among all the interfaces where this class is applied within the service policy.
- Policers do not maintain packet counters, only byte counters are supported.
- There is no specific command to verify the offered or outgoing traffic rate per−policer.
- Counters are updated on a periodical basis. If executing the above command repeatedly in fast succession, counters might still appear at some time.

Configuring and Monitoring Marking

Configuring marking involves the following steps:

1. Define the criteria for classifying the traffic − access−list, DSCP, IP precedence, and so on.
2. Define the traffic classes to be classified using criteria previously defined.
3. Create a policy map attaching marking actions and/or policing actions to the defined classes.
4. Configuring trust mode on the corresponding interface(s).
5. Apply the policy map to an interface.

Consider the following example where we want incoming traffic with IP precedence 3 to host 192.168.196.3 UDP port 777 mapped to IP precedence 6. All other IP precedence 3 traffic is policed down to 1 Mbps, and excess traffic should be marked down to IP precedence 2.

```
! enable QoS globally
qos
! define ACL to select UDP traffic to 192.168.196.3 port 777
ip access-list extended acl_test4
permit udp any host 192.168.196.3 eq 777
! define class of traffic using ACL and ip precedence matching
class-map match-all cl_test10
match ip precedence 3
match access-group name acl_test4
! all the remaining ip precedence 3 traffic will match this class
class-map match-all cl_test11
match ip precedence 3
! define policy with above classes
policy-map po_test10
class cl_test10
! mark traffic belonging to class with ip precedence 6
set ip precedence 6
class cl_test11
! police and mark down all other ip precedence 3 traffic
police 1 mbps 1000 exceed-action policed-dscp-transmit
!
! adjust DSCP to policed DSCP map so to map DSCP 24 to DSCP 16
qos map dscp policed 24 to dscp 16
!
interface FastEthernet5/14
! set interface to trust IP DSCP
```

```
qos trust dscp
! apply policy to interface
service-policy input po_test10
!
```

The **sh policy interface** command is used to monitor the marking. The sample output and implications are documented in the policing configuration above.

Comparing Policing and Marking on Catalyst 6000 and Catalyst 4000/4500 IOS Based Supervisor Engines

Feature	Catalyst6000	Catalyst4000 SE3
Egress QoS policies	Not supported by Supervisor 1A and Superviso r2 hardware.	Supported.
Burst policing parameter calculation	Burst should be at least the same size as maximum frame supposed to pass via policer and no less then rate/interval, with the interval being 250 microseconds	No such restriction.
QoS policing L2 & L3	By default, microflow policing is only enabled for L3 on the sup1a and is not enabled at all for Supervisor 2. A CLI command is available to enable it for L2 on sup1a and L2 & L3 for sup2. Aggregate policing for sup1a & Supervisor 2 is enabled by default for L2 & L3.	Always.
Egress CoS	Always derived from internal DSCP using DSCP to CoS QoS map.	If the ingress port is in trust CoS mode, the egress CoS will be the same as the ingress CoS. Otherwise, it will be derived from the internal DSCP.
Microflow policing	Supported.	Not supported.
QoS behavior when port is in VLAN-based QoS mode, but no policy is applied to the VLAN.	No policy applied.	Fallback to port-based QoS. Will apply policy attached to port.

Related Information

- **Understanding and Configuring QoS**
- **Catalyst 4000 Family IOS Commands**
- **Technical Support − Cisco Systems**

Updated: Mar 24, 2005 Document ID: 21263