Configuring MPLS QoS

This chapter describes how to configure Quality of Service for Multiprotocol Label Switching (MPLS) Layer 3 virtual private networks (VPNs).

- About MPLS Quality of Service (QoS), on page 1
- Guidelines and Limitations for MPLS QoS, on page 4
- Configuring MPLS QoS, on page 4
- About Traffic Queuing, on page 12
- Verifying MPLS QoS, on page 13

About MPLS Quality of Service (QoS)

MPLS QoS enables you to provide differentiated types of service across an MPLS network. Differentiated types of service satisfy a range of requirements by supplying the service specified for each packet. QoS allows you to classify the network traffic, police and prioritize the traffic flow, and provide congestion avoidance.

This section includes the following topics:

- MPLS QoS Terminology, on page 1
- MPLS QoS Features, on page 2

MPLS QoS Terminology

This section defines some MPLS QoS terminology:

- Classification is the process that selects the traffic to be marked. Classification matches traffic with the selection criteria into multiple priority levels or classes of service. Traffic classification is the primary component of class-based QoS provisioning. The switch makes classification decisions based on the EXP bits in the topmost label of the received MPLS packets (after a policy is installed).
- Differentiated Services Code Point (DSCP):
  - Is the first six bits of the ToS byte in the IP header.
  - Only present in an IP packet.
  - Can be present in an IPv4 or an IPv6 packet.
  - Is the first 6 bits of the 8-bit Traffic Class octet in the IPv6 header.

- E-LSP is a label switched path (LSP) on which nodes infer the QoS treatment for MPLS packets exclusively from the experimental (EXP) bits in the MPLS header. Because the QoS treatment is inferred from the EXP (both class and drop precedence), several classes of traffic can be multiplexed onto a single LSP (use the same label). A single LSP can support up to eight classes of traffic because the EXP field is a 3-bit field.
- EXP bits define the QoS treatment (per-hop behavior) that a node should give to a packet. It is the equivalent of the DiffServ Code Point (DSCP) in the IP network. A DSCP defines a class and drop precedence. The EXP bits are generally used to carry all the information encoded in the IP DSCP. In some cases, however, the EXP bits are used exclusively to encode the dropping precedence.
- Marking is the process of setting a Layer 3 DSCP value in a packet. Marking is also the process of choosing different values for the MPLS EXP field to mark packets so that they have the priority that they require during periods of congestion.
- MPLS Experimental Field: Setting the MPLS experimental (EXP) field value satisfies the requirement of operators who do not want the value of the IP precedence field modified within IP packets transported through their networks. By choosing different values for the MPLS EXP field, you can mark packets so that packets have the priority that they require during periods of congestion. By default, the three most significant bits of the DSCP are copied into the MPLS EXP field during imposition. You can mark the MPLS EXP bits with an MPLS QoS policy.

MPLS QoS Features

QoS enables a network to provide improved service to selected network traffic. This section explains the following MPLS QoS features, which are supported in an MPLS network:

MPLS Experimental Field

Setting the MPLS experimental (EXP) field value satisfies the requirement of service providers who do not want the value of the IP precedence field modified within IP packets transported through their networks.

By choosing different values for the MPLS EXP field, you can mark packets so that packets have the priority that they require during periods of congestion.

By default, the IP precedence value is copied into the MPLS EXP field during imposition. You can mark the MPLS EXP bits with an MPLS QoS policy.

Classification

Classification is the process that selects the traffic to be marked. Classification accomplishes this by partitioning traffic into multiple priority levels, or classes of service. Traffic classification is the primary component of class-based QoS provisioning.

Policing and Marking

Policing causes traffic that exceeds the configured rate to be discarded or marked down to a higher drop precedence. Marking is a way to identify packet flows to differentiate them. Packet marking allows you to partition your network into multiple priority levels or classes of service.

The MPLS QoS policing and marking features that you can implement depend on the received traffic type and the forwarding operation applied to the traffic.

DSCP Default Behavior

The DiffServ Tunneling mode set on ingress and egress decides the default behavior of DSCP field handling. The MPLS network support of DiffServ specification defines the following tunneling modes:

- **Uniform mode**

  The ingress tunnel endpoint copies the DSCP bits from the incoming IP packet into the MPLS EXP bits of the imposed labels during encapsulation. During decapsulation at the egress tunnel endpoint, the original DSCP values are not preserved. Instead, the outer header MPLS EXP is copied to the DSCP of the inner IP header.

- **Pipe mode**

  During decapsulation at the egress tunnel endpoint, the outer header MPLS EXP is discarded while retaining the DSCP of the inner IP header.

By default, all Cisco Nexus platform switches use uniform mode for encapsulation.

By default, the following mentioned switches use pipe mode for decapsulation:

- Cisco Nexus 9300-FX/FX2/FX3/GX/GX2 Platform Switches
- Cisco Nexus 9500-EX/FX/GX line cards
- Cisco Nexus 9800 Platform Switches

Note Cisco Nexus 9500-R line cards use uniform mode for decapsulation.

This distinction in mode operation from the remaining platforms is necessary to provide support for Segment Routing (SR) Traceroute. SR traceroute relies on the expiration of Time to Live (TTL) value of packet which carries MPLS echo request.

In Pipe mode of operation, TTL value retains the original value from inner IP header when the packet initially encapsulated, and removes outer header. This impacts SR traceroute to malfunction. However, in Uniform mode, the TTL value on the outer header is copied to the inner IP header during decapsulation retaining the amount of decrease value through passing the MPLS tunnel.

If necessary, to retain the DSCP value of the inner IP header, you can configure the tunnel decapsulation mode from Uniform to Pipe and Pipe to Uniform mode.

```
switch(config)#mpls qos pipe-mode
```

The no form of this command exists, and it sets switch to the default uniform mode.

Note This command is applicable for Cisco Nexus 9500-R line cards only.

You must configure this command before the MPLS interfaces/labels are created. If you configure using the pipe-mode command, you cannot use SR traceroute.

Guidelines and Limitations for MPLS QoS

MPLS Quality of Service (QoS) has the following configuration guidelines and limitations:

- When setting the QoS policy, the **topmost** keyword in the **set mpls experimental imposition** CLI is not supported.
- MPLS QoS does not support marking based on policing.
- L3 EVPN egress node - policing is not supported on a system level mpls-in-policy.
- Egress QoS classification that is based on MPLS EXP is not supported.
- EXP labels are only set for newly pushed or swapped labels. The EXP in the inner labels remains unchanged.
- When the traffic from the ingress line card takes the fabric module path to the line card, the line cards acting as the MPLS Ingress LSR node do not support ECN marking. This occurs for the Cisco Nexus 9500 platform switches with the N9K-X9700-EX and N9K-X9700-FX line cards.
- On the Label Edge Router (LER), policy match on EXP is not supported. Inner DSCP can be used to match the packets.
- Interface policy cannot be used to classify MPLS L3 EVPN packets on the Egress Label Edge Router (LER).System level MPLS-Default policy is used to classify the traffic.
- Explicit Congestion Notification (ECN) Marking is not supported on the label switching router transit node.
- Only the default QoS Service template is supported for the MPLS handoff in Cisco NX-OS Release 9.3(1). You cannot set the EXP labels on the MPLS.
- Beginning with Cisco NX-OS Release 9.3(5), MPLS QoS is supported on Cisco Nexus 9364C-GX, Cisco Nexus 9316D-GX, and Cisco Nexus 93600CD-GX switches.
- PFC is not supported for MPLS QoS and VXLAN MPLS DCI.
- Even after removing the queuing policy from an interface, previous micro-burst statistics remain. Use the clear queuing burst-detect command to clear the remaining records.
- RACL on an ingress port of egress PE (sr decap) is not supported.
- In order to write an EXP value in the label, an explicit poicy is necessary on the PE. In absence of a policy, the default EXP value is 7.

Configuring MPLS QoS

Note Be aware that the Cisco NX-OS commands for this feature may differ from those commands used in Cisco IOS.

Configuring MPLS Ingress Label Switched Router

To configure MPLS Ingress label switched router, perform the following:

MPLS Ingress LSR Classification

To match the value of the Differentiated Services Code Point (DSCP) field, use the **match dscp** command in QoS policy-map class configuration mode. To disable the setting, use the **no** form of this command.

Note Default entries are programmed to match on DSCP and mark EXP when no ingress QoS policy is configured (Uniform mode behavior at encap).

Before you begin

- You must enable MPLS configuration.
- Ensure that you are in the correct VDC (or use the switch to vdc command).

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **class-map type qos** *class-map-name*
	Example:
	`switch(config)# class-map type qos Class1`
	`switch(config-cmap-qos)#`
		Defines a class map, and enters class-map configuration mode.

Step 3	[**no**] **match** [**not**] **dscp** *dscp-list*
	Example:
	`switch(config)# switch(config-cmap-qos)#`
	`match dscp 2-4`
		List of DSCP values. Specifies that the packets should be matched (or not) on the DSCP label in the MPLS header as follows:
		• **dscp-list**—The list can contain values and ranges. Values can range from 0 to 63.

Configuring MPLS Ingress Policing and Marking

To configure a policy-map value and set the EXP value on all imposed label entries, use the **set mpls experimental imposition** command in QoS policy-map class configuration mode. To disable the setting, use the **no** form of this command.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **policy-map type qos** *policy-map-name*
	Example:
	`switch(config)# policy-map type qos pmap1`
	`switch(config-pmap-qos)#`
		Defines a policy map, and enters policy-map configuration mode.

Step 3	**class** *class-name* **Example:** `switch(config-pmap-qos)# class Class1`	Names the class-map.
Step 4	**set mpls experimental imposition** *exp_imposition_name* **Example:** `switch(config)# switch(config-pmap-qos)# set mpls experimental imposition 2`	MPLS experimental (EXP) values. Value range from 0 to 7.
Step 5	**set qos-group** *group-number* **Example:** `switch(config-cmap-qos)# set qos-group 1`	Identifies the qos-group number.
Step 6	**police cir** *burst-in-msec* **bc** *conform-burst-in-msec* **conform-action** *conform-action* **violate-action** *violate-action* **Example:** `switch(config-pmap-qos)# police cir 100 mbps bc 200 ms conform transmit violate drop`	Defines a policer for classified traffic in policy-map class configuration mode.

Step 7	**interface** *type slot/port*
	Example:
	`switch(config)# interface ethernet 2/2`
	`switch(config-if)#`
		Enters the interface configuration mode for the specified input interface, output interface, virtual circuit (VC), or a VC that will be used as the service policy for the interface or VC.

Step 8	**service-policy type qos input** *policy-map-name*
	Example:
	`switch(config-if)# service-policy type qos input pmap1`
	`switch(config-if)#`
		Attaches a policy map to an input interface, a virtual circuit (VC), an output interface, or a VC that will be used as the service policy for the interface or VC.

Configuring MPLS Transit Label Switching Router

To configure MPLS Transit Label Switching Routers, perform the following:

MPLS Transit LSR Classification

To map the value of the MPLS EXP field on all imposed label entries, use the **set mpls experimental topmost** command in QoS policy-map class configuration mode. To disable the setting, use the **no** form of this command.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **class-map type qos** *class-map-name*
	Example:
	`switch(config)# class-map type qos Class1`
	`switch(config-cmap-qos)#`
		Defines a class map, and enters class-map configuration mode.

Step 3	[**no**] **match** [**not**] **mpls experimental topmost** *exp-list*
	Example:
	`switch(config)# switch(config-cmap-qos)#`
	`match mpls experimental topmost 2, 4-7`
		List of MPLS experimental (EXP) values. Specifies that the packets should be matched (or not) on the 3-bit EXP field in the outermost (topmost) MPLS label in the MPLS header as follows:
		• **exp-list**—The list can contain values and ranges. Values can range from 0 to 7.

Configuring MPLS Transit Policing and Marking

To configure a policy-map value and set the EXP value on all imposed label entries, use the **service-policy type qos input pmap1** command in interface configuration mode. To disable the setting, use the **no** form of this command.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **policy-map type qos** *policy-map-name* **Example:**	Defines a policy map, and enters policy-map configuration mode.

	Command or Action	Purpose
	`switch(config)# policy-map type qos Class1` `switch(config-pmap-qos)#`	
Step 3	**class** *class-name* **Example:** `switch(config-pmap-qos)# class Class1`	Names the class-map.
Step 4	**set mpls experimental imposition** *exp_imposition_name* **Example:** `switch(config)# switch(config-pmap-qos)# set mpls experimental imposition 2`	MPLS experimental (EXP) values. Value range from 0 to 7.
Step 5	**set qos-group** *group-number* **Example:** `switch(config-pmap-qos)# set qos-group 1`	Identifies the qos-group number.
Step 6	**police cir** *burst-in-msec* **bc** *conform-burst-in-msec* **conform-action** *conform-action* **violate-action** *violate-action* **Example:** `switch(config-pmap-qos)# police cir 100 mbps bc 200 ms conform transmit violate drop`	Defines a policer for classified traffic in policy-map class configuration mode. • violate-action - **drop** is the only supported keyword for Transit LSR

Step 7	**interface** *type slot/port*
	Example:
	`switch(config)# interface ethernet 2/2`
	`switch(config-if)#`
		Enters the interface configuration mode for the specified input interface, output interface, virtual circuit (VC), or a VC that will be used as the service policy for the interface or VC.

Step 8	**service-policy type qos input** *policy-map-name*
	Example:
	`switch(config-if)# service-policy type qos input pmap1`
	`switch(config-if)#`
		Attaches a policy map to an input interface, a virtual circuit (VC), an output interface, or a VC that is used as the service policy for the interface or VC.

Configuring MPLS Egress Label Switching Router

To configure MPLS Egress label switched router, perform the following:

MPLS Egress LSR Classification

To classify the incoming SR MPLS traffic to egress queue, use the match on Differentiated Services Code Point (DSCP) field.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **class-map type qos** *class-map-name*
	Example:
	`switch(config)# class-map type qos Class1`
	`switch(config-cmap-qos)#`
		Defines a class map, and enters class-map configuration mode.

Step 3	[**no**] **match** [**not**] **dscp** *dscp-list*
	Example:
	`switch(config)# switch(config-cmap-qos)#`
	`match dscp 2-4`
		List of DSCP values. Specifies that the packets should be matched (or not) on the DSCP label in the MPLS header as follows:
		• **dscp-list**—The list can contain values and ranges. Values can range from 0 to 63.

MPLS Egress LSR Classification - Default Policy Template

To classify the incoming traffic to the egress queue of an EVPN tunnel, use the default **default-mpls-in-policy** command at the system level. To disable the setting, use the **no** form of this command.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **system qos**
	Example:
	`switch(config)# system qos`
	`switch(config-sys-qos)#`
		Enters system QoS configuration mode.

Step 3	[**no**] **service-policy type qos input default-mpls-in-policy** **Example:** `switch(config-sys-qos)# service-policy type qos input default-mpls-in-policy`	Specifies the “default-mpls-in-policy” at the system level to match on the incoming SR L3 EVPN MPLS traffic.

The following is the default MPLS in policy template configured with the **service-policy type qos input default-mpls-in-policy** command.

```
policy-map type qos default-mpls-in-policy
    class  c-dflt-mpls-qosgrp1
      set qos-group 1
    class  c-dflt-mpls-qosgrp2
      set qos-group 2
    class  c-dflt-mpls-qosgrp3
      set qos-group 3
    class  c-dflt-mpls-qosgrp4
      set qos-group 4
    class  c-dflt-mpls-qosgrp5
      set qos-group 5
    class  c-dflt-mpls-qosgrp6
      set qos-group 6
    class  c-dflt-mpls-qosgrp7
      set qos-group 7
    class  class-default
      set qos-group 0

class-map type qos match-any c-dflt-mpls-qosgrp1
  Description: This is an ingress default qos class-map that classify traffic with prec 1
  match precedence 1

class-map type qos match-any c-dflt-mpls-qosgrp2
  Description: This is an ingress default qos class-map that classify traffic with prec 2
  match precedence 2

class-map type qos match-any c-dflt-mpls-qosgrp3
  Description: This is an ingress default qos class-map that classify traffic with prec 3
  match precedence 3

class-map type qos match-any c-dflt-mpls-qosgrp4
  Description: This is an ingress default qos class-map that classify traffic with prec 4
  match precedence 4

class-map type qos match-any c-dflt-mpls-qosgrp5
  Description: This is an ingress default qos class-map that classify traffic with prec 5
  match precedence 5

class-map type qos match-any c-dflt-mpls-qosgrp6
  Description: This is an ingress default qos class-map that classify traffic with prec 6
  match precedence 6

class-map type qos match-any c-dflt-mpls-qosgrp7
  Description: This is an ingress default qos class-map that classify traffic with prec 7
  match precedence 7
```

Custom MPLS-in-Policy Mapping

You can override the queue mapping of incoming traffic by editing a local copy of the template provided. The system matching is always based on precedence, and requires the “mpls-in-policy” string to be part of the policy name. Marking with QoS is supported. Set can be qos-group, vlan-cos, or both.

```
class-map type qos match-all prec-1
      match precedence 1
    class-map type qos match-all prec-2
      match precedence 2

policy-map type qos test-mpls-in-policy
    class  prec-1
      set qos-group 3
    class  prec-2
      set qos-group 4
```

```
system qos
  service-policy type qos input test-mpls-in-policy
```

Note Classification based on Precedence is only supported and Marking is not supported on system level mpls-in-policy.

Configuring MPLS Egress LSR - Policing and Marking

To configure and apply a policy-map with policer config, use the **service-policy type qos input pmap1** command in interface configuration mode. To disable the setting, use the **no** form of this command.

Note Policing is not supported for SR L3 EVPN MPLS traffic

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **policy-map type qos** *class-map-name*
	Example:
	`switch(config)# policy-map type qos Class1`
	`switch(config-pmap-qos)#`
		Defines a class map, and enters class-map configuration mode.

Step 3	**policy** *policy-name* **Example:** `switch(config-pmap-qos)# class Class1`	Names the class-map.
Step 4	**set dscp** *dscp-value* **Example:** `switch(config-pmap-qos)# set dscp 4`	Identifies the dscp value.
Step 5	**set qos-group** *group-number* **Example:** `switch(config-pmap-qos)# set qos-group 1`	Identifies the qos-group number.
Step 6	[**no**] **police cir** *burst-in-msec* **bc** *conform-burst-in-msec* **conform-action** *conform-action* **violate-action** *violate-action*	Defines a policer for classified traffic in policy-map class configuration mode.

	Command or Action	Purpose
	Example: `switch(config-pmap-qos)# police cir 100 mbps bc 200 ms conform transmit violate drop`	

Step 7	**interface** *type slot/port*
	Example:
	`switch(config)# interface ethernet 2/2`
	`switch(config-if)#`
		Enters the interface configuration mode for the specified interface.

Step 8	[**no**] **service-policy type qos input** *policy-map-name*
	Example:
	`switch(config-if)# service-policy type qos input pmap1`
	`switch(config-if)#`
		Attaches a policy map to an input interface, a virtual circuit (VC), an output interface, or a VC that will be used as the service policy for the interface or VC.

About Traffic Queuing

Traffic queuing is the ordering of packets and applies to both input and output of data. Device modules can support multiple queues, which you can use to control the sequencing of packets in different traffic classes. You can also set weighted random early detection (WRED) and taildrop thresholds. The device drops packets only when the configured thresholds are exceeded.

Configuring QoS Traffic Queuing

To set the output queue, use the **set qos-group** command in policy map configuration mode. To disable the setting, use the **no** form of this command.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **policy-map type qos** *class-map-name*
	Example:
	`switch(config)# class-map type qos Class1`
	`switch(config-cmap-qos)#`
		Defines a class map, and enters class-map configuration mode.

Step 3	**class** *class-name* **Example:** `switch(config-cmap-qos)# class Class1`	Names the class-map.

	Command or Action	Purpose
Step 4	**set qos-group** *qos_group_number* **Example:** `switch(config-pmap-c-qos)# set qos-group`	Applies queueing parameters for the named QoS group in policy map. Value range from 0 to 7.

Verifying MPLS QoS

To display the MPLS QoS configuration, perform the following task:

Command	Description
show hardware internal forwarding table utilization	Displays information about the MAX label entries and Used label entries.
show class-map	Displays the interface class mapping statistics.
show policy-map system type qos input	Displays the cumulative statistics that show the packets matched for every class for all the interfaces (only for the EVPN tunnel case). For more information, see the sample output following this table.
show policy-map type qos interface interface	Displays the statistics that show the packets matched for every class on that interface in the given direction.
show policy-map type qos <pmap name>	Displays the service policy maps configured on the interfaces.
show queuing interface	Displays the queuing information of interfaces.

The following example displays the cumulative statistics that show the packets matched for every class for all the interfaces (only for the EVPN tunnel case).

```
switch# show policy-map system type qos input


  Service-policy (qos) input:   default-mpls-in-policy

    Class-map (qos):   c-dflt-mpls-qosgrp1 (match-any)

     Slot 3
        2775483 packets
     Aggregate forwarded :
        2775483 packets
```

```
  Match: precedence 1
  set qos-group 1

Class-map (qos):   c-dflt-mpls-qosgrp2 (match-any)

 Slot 3
    2775549 packets
 Aggregate forwarded :
    2775549 packets
  Match: precedence 2
  set qos-group 2

Class-map (qos):   c-dflt-mpls-qosgrp3 (match-any)

 Slot 2
    2777189 packets
 Aggregate forwarded :
    2777189 packets
  Match: precedence 3
  set qos-group 3

Class-map (qos):   c-dflt-mpls-qosgrp4 (match-any)

 Slot 3
    2775688 packets
 Aggregate forwarded :
    2775688 packets
  Match: precedence 4
  set qos-group 4

Class-map (qos):   c-dflt-mpls-qosgrp5 (match-any)

 Slot 3
    2775756 packets
 Aggregate forwarded :
    2775756 packets
  Match: precedence 5
  set qos-group 5

Class-map (qos):   c-dflt-mpls-qosgrp6 (match-any)

 Slot 3
    2775824 packets
 Aggregate forwarded :
    2775824 packets
  Match: precedence 6
  set qos-group 6

Class-map (qos):   c-dflt-mpls-qosgrp7 (match-any)

 Slot 3
    2775892 packets
 Aggregate forwarded :
    2775892 packets
  Match: precedence 7
  set qos-group 7

Class-map (qos):   class-default (match-any)

 Slot 3
    2775962 packets
 Aggregate forwarded :
    2775962 packets
  set qos-group 0
```