



Configuring QoS on the System

This chapter contains the following sections:

- [Information About System Classes, on page 1](#)
- [Configuring System QoS, on page 3](#)
- [Verifying the System QoS Configuration, on page 8](#)

Information About System Classes

System Classes

The system qos is a type of MQC target. You use a service policy to associate a policy map with the system qos target. A system qos policy applies to all interfaces on the switch unless a specific interface has an overriding service-policy configuration. The system qos policies are used to define system classes, the classes of traffic across the entire switch, and their attributes. To ensure QoS consistency (and for ease of configuration), the device distributes the system class parameter values to all its attached network adapters using the Data Center Bridging Exchange (DCBX) protocol.

If service policies are configured at the interface level, the interface-level policy always takes precedence over system class configuration or defaults.

On the Cisco Nexus device, a system class is uniquely identified by a qos-group value. A total of six system classes are supported. Two of the six system classes are defaults and are always present on the device. Up to four additional system classes can be created by the administrator.

Default System Classes

The device provides the following system classes:

- Drop system class

By default, the software classifies all unicast and multicast Ethernet traffic into the default drop system class. This class is identified by qos-group 0.

This class is created automatically when the system starts up (the class is named **class-default** in the CLI). You cannot delete this class and you cannot change the match criteria associated with the default class.



Note If congestion occurs when data traffic (class-default) and FCoE traffic (class-fcoe) is flowing at the same time, then the queuing percentage configuration starts up.

The FCoE traffic is a no-drop class and does not get policed to the bandwidth assigned as per the queuing class. FCoE traffic cannot be dropped as it expects a lossless medium. When congestion occurs PFC frames are generated at FCoE ingress interfaces and dropping only occurs on the data traffic, even if data traffic is below the assigned bandwidth.

For optimizing the throughput you can spread the data traffic load for a longer duration.

- FCoE system class (For the Cisco Nexus 5500 Series device)

For the Cisco Nexus 5500 Series device, the class-fcoe is not automatically created. Before you enable FCoE on the Cisco Nexus 5500 Series device running Cisco NX-OS Release 5.0(2)N1(1), you must enable class-fcoe in the three types of qos policies:

- type qos policy maps
- type network-qos policy map (attached to system qos)
- type queuing policy map (class-fcoe must be configured with a non-zero bandwidth percentage for input queuing policy maps.

When class-fcoe is not included in the qos policies, vFC interfaces do not come up and increased drops occur.



Note The Cisco Nexus 5500 Series device supports five user-defined classes and one default drop system class.

MTU

The Cisco Nexus device is a Layer 2 switch, and it does not support packet fragmentation. A maximum transmission unit (MTU) configuration mismatch between ingress and egress interfaces may result in packets being truncated.

When configuring MTU, follow these guidelines:

- MTU is specified per system class. The system class allows a different MTU for each class of traffic but they must be consistent on all ports across the entire switch. You cannot configure MTU on the interfaces.
- Fibre Channel and FCoE payload MTU is 2158 bytes across the switch. As a result, the rxbufsize for Fibre Channel interfaces is fixed at 2158 bytes. If the Cisco Nexus device receives an rxbufsize from a peer that is different than 2158 bytes, it will fail the exchange of link parameters (ELP) negotiation and not bring the link up.
- Enter the **system jumbomtu** command to define the upper bound of any MTU in the system. The system jumbo MTU has a default value of 9216 bytes. The minimum MTU is 2158 bytes and the maximum MTU is 9216 bytes.

- Configuring the MTU to 9216 bytes on both the Layer 3 ports and the Network QoS at the same time is not supported.
- The system class MTU sets the MTU for all packets in the class. The system class MTU cannot be configured larger than the global jumbo MTU.
- The FCoE system class (for Fibre Channel and FCoE traffic) has a default MTU of 2158 bytes. This value cannot be modified.
- The switch sends the MTU configuration to network adapters that support DCBX.



Note MTU is not supported in Converged Enhanced Ethernet (CEE) mode for DCBX.

Configuring System QoS

Attaching the System Service Policy

The `service-policy` command specifies the system class policy map as the service policy for the system.

Procedure

	Command or Action	Purpose
Step 1	switch# configure terminal	Enters global configuration mode.
Step 2	switch(config)# system qos	Enters system class configuration mode.
Step 3	switch(config-sys-qos)# service-policy type { network-qos qos queuing } [input output] <i>policy-name</i>	Specifies the policy map to use as the service policy for the system. There are three policy-map configuration modes: <ul style="list-style-type: none"> • network-qos—Network-wide (system qos) mode. • qos—Classification mode (system qos input or interface input only). • queuing—Queuing mode (input and output at system qos and interface).

	Command or Action	Purpose
		<p>Note There is no default policy-map configuration mode; you must specify the type. The input keyword specifies that this policy map should be applied to traffic received on an interface. The output keyword specifies that this policy-map should be applied to traffic transmitted from an interface. You can only apply input to a qos policy; you can apply both input and output to a queuing policy.</p>
Step 4	<p>(Optional) switch(config-sys-qos)# service-policy type {network-qos qos queuing} [input output], fcoe default policy-name</p>	<p>Specifies the default FCoE policy map to use as the service policy for the system. There are four pre-defined policy-maps for FCoE:</p> <ul style="list-style-type: none"> • service-policy type qos input fcoe-default-in-policy • service-policy type queuing input fcoe-default-in-policy • service-policy type queuing output fcoe-default-out-policy • service-policy type network-qos fcoe-default-nq-policy <p>Note Before enabling FCoE on a Cisco Nexus device, you must attach the pre-defined FCoE policy maps to the type qos, type network-qos, and type queuing policy maps.</p>

Example

This example shows how to set a no-drop Ethernet policy map as the system class:

```
switch(config)# class-map type network-qos ethCoS4
switch(config-cmap-nq)# match qos-group
switch(config-cmap-nq)# exit
switch(config)# policy-map type network-qos ethNoDrop
switch(config-pmap-nq)# class type network-qos ethCoS4
switch(config-pmap-c-nq)# pause no-drop
switch(config-pmap-c-nq)# exit
switch(config-pmap-nq)# exit
switch(config)# system qos
switch(config-sys-qos)# service-policy type network-qos ethNoDrop
```

Restoring the Default System Service Policies

If you have created and attached new policies to the system QoS configuration, enter the **no** form of the command to reapply the default policies.

Procedure

	Command or Action	Purpose
Step 1	switch# configure terminal	Enters global configuration mode.
Step 2	switch(config)# system qos	Enters system class configuration mode.
Step 3	switch(config-sys-qos)# no service-policy type qos input <i>policy-map name</i>	Resets the classification mode policy map. This policy-map configuration is for system QoS input or interface input only:
Step 4	switch(config-sys-qos)# no service-policy type network-qos <i>policy-map name</i>	Resets the network-wide policy map.
Step 5	switch(config-sys-qos)# no service-policy type queuing output <i>policy-map name</i>	Resets the output queuing mode policy map.
Step 6	switch(config-sys-qos)# no service-policy type queuing input <i>policy-map name</i>	Resets the input queuing mode policy map.

Example

The following example shows how to reset the system QoS configuration:

```
switch# configure terminal
switch(config)# system qos
switch(config-sys-qos)# no service-policy type qos input my-in-policy
switch(config-sys-qos)# no service-policy type network-qos my-nq-policy
switch(config-sys-qos)# no service-policy type queuing output my-out-policy
switch(config-sys-qos)# no service-policy type queuing input my-in-policy
```

Configuring the Queue Limit for a Specified Fabric Extender

At the Fabric Extender configuration level, you can control the queue limit for a specified Fabric Extender for egress direction (from the network to the host). You can use a lower queue limit value on the Fabric Extender to prevent one blocked receiver from affecting traffic that is sent to other noncongested receivers ("head-of-line blocking"). A higher queue limit provides better burst absorption and less head-of-line blocking protection. You can use the **no** form of this command to allow the Fabric Extender to use all available hardware space.



Note

At the system level, you can set the queue limit for Fabric Extenders by using the **fex queue-limit** command. However, configuring the queue limit for a specific Fabric Extender will override the queue limit configuration set at the system level for that Fabric Extender.

You can specify the queue limit for the following Fabric Extenders:

- Cisco Nexus 2148T Fabric Extender (48x1G 4x10G SFP+ Module)
- Cisco Nexus 2224TP Fabric Extender (24x1G 2x10G SFP+ Module)
- Cisco Nexus 2232P Fabric Extender (32x10G SFP+ 8x10G SFP+ Module)
- Cisco Nexus 2248T Fabric Extender (48x1G 4x10G SFP+ Module)
- Cisco Nexus N2248TP-E Fabric Extender (48x1G 4x10G Module)
- Cisco Nexus N2348UPQ Fabric Extender (48x10G SFP+ 6x40G QSFP Module)

Procedure

	Command or Action	Purpose
Step 1	switch# configure terminal	Enters global configuration mode.
Step 2	switch(config)# fex <i>fex-id</i>	Specifies the Fabric Extender and enters the Fabric Extender mode.
Step 3	switch(config-fex)# hardware <i>fex_card_type</i> queue-limit <i>queue-limit</i>	Configures the queue limit for the specified Fabric Extender. The queue limit is specified in bytes. The range is from 81920 to 652800 for a Cisco Nexus 2148T Fabric Extender and from 2560 to 652800 for all other supported Fabric Extenders.

Example

This example shows how to restore the default queue limit on a Cisco Nexus 2248T Fabric Extender:

```
switch# configure terminal
switch(config-if)# fex 101
switch(config-fex)# hardware N2248T queue-limit 327680
```

This example shows how to remove the queue limit that is set by default on a Cisco Nexus 2248T Fabric Extender:

```
switch# configure terminal
switch(config)# fex 101
switch(config-fex)# no hardware N2248T queue-limit 327680
```

Enabling the Jumbo MTU

You can enable the jumbo Maximum Transmission Unit (MTU) for the whole switch by setting the MTU to its maximum size (9216 bytes) in the policy map for the default Ethernet system class (class-default).

When you configure jumbo MTU on a port-channel subinterface you must first enable MTU 9216 on the base interface and then configure it again on the subinterface. If you enable the jumbo MTU on the subinterface before you enable it on the base interface then the following error will be displayed on the console:

```
switch(config)# int po 502.4
switch(config-subif)# mtu 9216
ERROR: Incompatible MTU values
```

For Layer 3 routing on Cisco Nexus devices, you need to configure the MTU on the Layer 3 interfaces (SVIs and physical interfaces with IP addresses) in addition to the global QoS configuration below.

To use FCoE on switch, add class-fcoe in the custom network-qos policy. If already using FCoE, make sure to add the below lines in the config so that the FCoE does not go down on the switch after enabling the jumbo qos policy.

```
switch# conf t
switch(config)# policy-map type network-qos jumbo
switch(config-pmap-nq)# class type network-qos class-fcoe
switch(config-pmap-nq-c)# end
```

This example shows how to change qos to enable the jumbo MTU:

```
switch# conf t
switch(config)# policy-map type network-qos jumbo
switch(config-pmap-nq)# class type network-qos class-default
switch(config-pmap-c-nq)# mtu 9216
```



Note The **system jumbomtu** command defines the maximum MTU size for the switch. However, jumbo MTU is supported only for system classes that have MTU configured.

Verifying the Jumbo MTU

On the Cisco Nexus device, traffic is classified into one of eight QoS groups. The MTU is configured at the QoS group level. By default, all Ethernet traffic is in QoS group 0. To verify the jumbo MTU for Ethernet traffic, use the **show queuing interface ethernet slot/chassis_number** command and find "HW MTU" in the command output to check the MTU for QoS group 0. The value should be 9216.

The **show interface** command always displays 1500 as the MTU. Because the Cisco Nexus device supports different MTUs for different QoS groups, it is not possible to represent the MTU as one value on a per interface level.



- Note**
- For Layer 3 routing on the Cisco Nexus device, you must verify the MTU on the Layer 3 interfaces (SVIs and physical interfaces with IP addresses) in addition to the global QoS MTU. You can verify the Layer 3 MTU by using the **show interface vlan vlan_number** or **show interface slot/chassis_number**.
 - A total of 640k port buffer is available on the Cisco Nexus 5500 Series platform and a total of 480k port buffer is available on the Cisco Nexus 5000 Series platform. When a customer queuing policy is created, the actual amount of buffer used is reduced.

This example shows how to display jumbo MTU information for Ethernet 1/19:

```
switch# show queuing interface ethernet1/19
Ethernet1/19 queuing information:
TX Queuing
  qos-group  sched-type  oper-bandwidth
  0           WRR         50
```

```

1          WRR          50

RX Queuing
qos-group 0
q-size: 243200, HW MTU: 9280 (9216 configured)
drop-type: drop, xon: 0, xoff: 1520
Statistics:
  Pkts received over the port          : 2119963420
  Ucast pkts sent to the cross-bar     : 2115648336
  Mcast pkts sent to the cross-bar     : 4315084
  Ucast pkts received from the cross-bar : 2592447431
  Pkts sent to the port                : 2672878113
  Pkts discarded on ingress            : 0
  Per-priority-pause status           : Rx (Inactive), Tx (Inactive)

qos-group 1
q-size: 76800, HW MTU: 2240 (2158 configured)
drop-type: no-drop, xon: 128, xoff: 240
Statistics:
  Pkts received over the port          : 0
  Ucast pkts sent to the cross-bar     : 0
  Mcast pkts sent to the cross-bar     : 0
  Ucast pkts received from the cross-bar : 0
  Pkts sent to the port                : 0
  Pkts discarded on ingress            : 0
  Per-priority-pause status           : Rx (Inactive), Tx (Inactive)

Total Multicast crossbar statistics:
  Mcast pkts received from the cross-bar : 80430744

```

Verifying the System QoS Configuration

Use one of the following commands to verify the configuration:

Command	Purpose
show policy-map system	Displays the policy map settings attached to the system QoS.
show policy-map [name]	Displays the policy maps defined on the switch. Optionally, you can display the named policy only.
show class-map	Displays the class maps defined on the switch.
running-config ipqos	Displays information about the running configuration for QoS.
startup-config ipqos	Displays information about the startup configuration for QoS.