Configuring Queuing and Scheduling on M-Series I/O Modules

This chapter describes how to configure the QoS queuing and scheduling features on M-Series I/O modules of the Cisco NX-OS device.

Finding Feature Information

Your software release might not support all the features documented in this module. For the latest caveats and feature information, see the Bug Search Tool at https://tools.cisco.com/bugsearch/ and the release notes for your software release. To find information about the features documented in this module, and to see a list of the releases in which each feature is supported, see the "New and Changed Information"chapter or the Feature History table in this chapter.

Information About Queuing and Scheduling

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.

Traffic scheduling is the methodical output of packets at a desired frequency to accomplish a consistent flow of traffic. You can apply traffic scheduling to different traffic classes to weight the traffic by priority.

The queuing and scheduling processes allow you to control the bandwidth that is allocated to the traffic classes, so that you achieve the desired trade-off between throughput and latency for your network.

Table 1: System-Defined Queue Types

Queue Type	Direction	Description
2q4t	Input	2 queues with 4 WRED or tail drop thresholds per queue
1p3q4t	Output	1 strict priority plus 3 normal queues with 4 WRED or tail-drop thresholds per queue
8q2t	Input	8 queues with 2 tail drop thresholds per queue
1p7q4t	Output	1 strict priority queue plus 7 normal queues with 4 WRED or tail drop thresholds per queue
1p7qlt	Output	1 strict priority queue plus 7 normal queues with 1WRED or tail drop thresholds per queue
1p3q1t	Output	1 strict priority queue plus 3 normal queues with 1 WRED or tail drop thresholds per queue
2p2q1t	Output	2 strict priority queues plus 2 normal queues with 1 WRED or tail drop thresholds per queue
2p6q1t	Output	2 strict priority queues plus 6 normal queues with 1 WRED or tail drop thresholds per queue
3p1q1t	Output	3 strict priority queues plus 1 normal queue with 1 WRED or tail drop thresholds per queue
3p5qlt	Output	3 strict priority queues plus 5 normal queues with 1 WRED or tail drop thresholds per queue

The queues match on the class of service (CoS) field. The device ensures that every CoS value from 0 to 7 maps to a queue for each queue type. Only one queue for a queue type can be assigned with a specific CoS value. For more information about the system-defined queues, see Table 3 .

DSCP-to-queue mapping is disabled by default. Use the **hardware qos dscp-to-queue ingress** command to enable DSCP mapping on the modules. The table below describes the system-defined DSCP queuing class-maps with the default DSCP values.

Table 2: System-Defined Queuing Class Maps with Default DSCP Values

Class Map Queue Name	Description	Default DSCP Value
8q2t-in-q1	Ingress queue 1 of type 8q2t	40-63
8q2t-in-q2	Ingress queue 2 of type 8q2t	—
8q2t-in-q3	Ingress queue 3 of type 8q2t	—
8q2t-in-q4	Ingress queue 4 of type 8q2t	—
8q2t-in-q5	Ingress queue 5 of type 8q2t	—
8q2t-in-q6	Ingress queue 6 of type 8q2t	—
8q2t-in-q7	Ingress queue 7 of type 8q2t	—
8q2t-in-q-default	Ingress default queue of type 8q2t	0-39

Setting Ingress Port CoS

You can set the CoS field in all ingress packets for untrusted ports. By default, ports are trusted and the CoS field is not modified. You can use this method to configure the port state to trusted or untrusted.

For information about configuring ingress port CoS, see the “Configuring Ingress Port CoS” section.

Modifying Class Maps

You can modify the CoS values that are matched by system-defined queuing class maps, which modify the CoS-to-queue mapping. Table 3 lists the default system-defined CoS values. Each CoS value appears only once in the queues of the same type.

If you want to change the system-default queuing class maps, you must also change the queuing policies applied on the interfaces because any changes in the queuing class maps causes traffic disruptions and might also cause packet drops.

Caution When you modify a system-defined queuing class map, the changes occur immediately and it might disrupt traffic on all virtual device contexts (VDCs).

Note For traffic crossing Layer 3, the queue mapping CoS-to-queue occurs automatically.

For information about configuring class maps, see the “Modifying Queuing Class Maps for COS” section.

Note Starting from Cisco NX-OS Release 6.2(2), DSCP-to-queue mapping on ingress class maps is supported on M Series 10G modules. However, the DSCP-to-queue mapping on all egress class maps is not supported.

Congestion Avoidance

You can use the following methods to proactively avoid traffic congestion on the device:

- Apply WRED to a class of traffic, which allows the device to drop packets based on the CoS field. WRED is designed to work with TCP traffic.
- Apply tail drop to a class of traffic, which allows the device to drop packets based on the CoS field.
- Apply WRED to a class of traffic, which allows the device to drop packets based on the DSCP field. WRED is designed to work with TCP traffic.
- Apply tail drop to a class of traffic, which allows the device to drop packets based on the DSCP field.

For information about configuring congestion avoidance, see the "Modifying Queuing Class Maps for DSCP" section.

Congestion Management

For ingress packets, you can configure congestion management by specifying a bandwidth that allocates a minimum data rate to a queue.

For egress packets, you can choose one of the following congestion management methods:

- Specify a bandwidth that allocates a minimum data rate to a queue.
- Impose a maximum data rate on a class of traffic so that excess packets are retained in a queue to shape the output rate.
- Allocate all data for a class of traffic to a priority queue. The device distributes the remaining bandwidth among the other queues.

For information about configuring congestion management, see the "Configuring WRED by DSCP Values" section.

Virtualization Support

A VDC is a logical representation of a set of system resources. Other than configuring class maps, queuing and scheduling apply only to the VDC where the commands are entered. For information about configuring class maps, see the "Modifying Queuing Class Maps for COS" section.

For information about configuring VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

Prerequisites for Queuing and Scheduling

Queuing and scheduling have the following prerequisites:

- You must be familiar with the "Using Modular QoS CLI" chapter.
- You are logged on to the switch.
- You are in the correct VDC. A VDC is a logical representation of a set of system resources. You can use the **switchto vdc** command with a VDC number.

Guidelines and Limitations

Queuing and scheduling have the following configuration guidelines and limitations:

- If a no-drop class is paused and the IP traffic is received with the CoS value of the no-drop class, IP traffic is queued in default queue due to the dscp-to-queue mapping behaviour. This is applicable to Cisco Nexus 7700 Series switches by default. Note that the dscp-to-queue mapping can be disabled.
-
- Configure system-defined class maps with care because the changes occur immediately and traffic might be disrupted on all VDCs.
- Defining the CoS with the **match cos** command is not supported for custom configured class maps.
- Specifying DSCP values for a class map with the **match dscp** command is not supported for custom configured class maps.
- When you are working with 10-Gigabit Ethernet ports in the shared mode, the egress queuing policy applies to all the ports in the port group. With the 10-Gigabit Ethernet ports in shared mode, all the ports in the port group must be in the same VDC. For information about the shared and dedicated modes, see the *Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide*. For information about the port groups, see the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide*.
- You cannot set either the queue limit or WRED on ingress 10-Gigabit Ethernet ports except for the 8 port, 10-Gigabit Ethernet I/O module.
- To ensure accurate hardware programming on Cisco M1 Series modules, when you add a physical interface to a port-channel, which already has a queuing policy applied, you must remove the queuing policy and reapply. Without this reapplication, queuing policy will not be correctly applied on all the interfaces. Otherwise, you must ensure that the queuing policy is applied on the port-channel interface only after all the physical ports are bundled into the port-channel.
- On Cisco M1 series modules, it may not be possible to configure actual values for traffic shaping. For example, on a 1 Gigabit interface with 65% average shaping, the output rate on the interface goes only up to 450Mbps, whereas with 70%, it goes to 850Mbps.
- When changing egress Class of Service (CoS) to queue mapping, ensure that you specify 2 or 3 seconds as the minimum time limit between changes. Otherwise, continuous traffic drop might occur.
- The Cisco M3 series modules do not support per-queue counters for egress drops (multicast, unknown unicast, or broadcasts). The egress drops will be per port and per Q-Default counter.
- Only 8e templates are supported on Cisco M3 series modules.
- In Cisco Nexus 7700 switches, the Cisco M3 series module supports only network-qos policies based on the 8e-4q8q template.
- The Cisco M3 series module supports only the network-qos template. This template contains all the CoS values that match the MTU size.
- All data traffic will be enqueued to the default queue of dot1q-tunnel port because this port is untrusted by default.

Configuring Queuing and Scheduling

Queuing and scheduling are configured by creating policy maps of type queuing that you apply to either traffic direction of an interface. You can modify system-defined class maps, which are used in policy maps to define the classes of traffic to which you want to apply policies.

Additional considerations are as follows:

- Changes to system class maps take effect immediately across all VDCs.

  The specified CoS values immediately map to the new queues.

- Changes are disruptive.

  The traffic passing through ports of the specified port type experience a brief period of traffic loss. All ports of the specified type are affected. For example, if you change COS-to-queue mapping for the M1 10G egress interface type, all M1 10G ports in all VDCs experience a brief disruption.

- Performance can be impacted.

  If one or more ports of the specified type do not have a queuing policy applied that defines the behavior for the new queue, then the traffic mapping to that queue might experience performance degradation.

- If you change the CoS-to-queue mapping by modifying the queuing class maps, you must ensure that a new queuing policy was applied to all ports of that type that use the new queues.

- If you change the DSCP-to-ingress-queue mapping by modifying the queuing class maps, you must ensure that a new queuing policy is applied to all ports of that type that use the new queues.

- By default, nonused queues do not have an allocated buffer. Allocate buffers to these queues to avoid tail drop.

- Changes to system class-maps are made only on the default VDC.

For information about configuring policy maps and class maps, see “Using Modular QoS CLI.”

You can configure the congestion-avoidance features, which include tail drop and WRED, in any queue. You can configure one of the egress congestion management features, such as priority, shaping, and bandwidth, in output queues, and you can configure bandwidth in input queues.

We recommend that you modify the CoS value before you create a policy map. You can modify the CoS values that are matched by device-defined class map queues. You must assign each CoS value from 0 to 7 to one or more of the queues for each queue type. Each CoS value is used only once in each queue type.

We recommend that you modify the DSCP value before you create a policy map. You can modify the DSCP values that are matched by device-defined class map queues. You must assign each DSCP value from 0 to 63 to one or more of the queues for each queue type. Each DSCP value is used only once in each queue type.

The system-defined policy maps default-in-policy and default-out-policy are attached to all ports to which you do not apply a queuing policy map. The default policy maps cannot be configured. For more information about the default policy maps, see Table 5.

This example shows that if you downgrade from Release 4.0(3) to Release 4.0(2) and enter the **show running-configuration** command, the input default queuing policy has an unknown enum in the display:

```
switch# show running-config
version 4.0(2)
```

```
…
…
policy-map type queuing default-in-policy
class type queuing unknown enum 0
queue-limit percent 50
bandwidth percent 80
class type queuing unknown enum 0
queue-limit percent 50
bandwidth percent 20
```

If you copy and paste this configuration into any Cisco NX-OS release, the device sends errors while executing all the commands starting from the **policy-map type queuing default-in-policy** command. You can ignore these errors because they do not affect the performance of the device.

Configuring Ingress Port CoS

To make a port untrusted, set the CoS value to a static value.

Note

- By default, ports are trusted (trust CoS) and the CoS field is not modified. When you configure the ingress port CoS value, the port becomes untrusted.
- For the untagged bridged traffic, a Cisco Nexus 7000 Series device ignores the Differentiated Services Code Point (DSCP) and queues on ingress and egress directions, if the CoS value is 0.
- By default, Layer 3 ports trust DSCP and also copy the DSCP value to CoS.

You use the ingress default queues from the system-defined queue classes for the type of module to which you want to apply the policy map. For the list of system-defined class maps for each type of module, see Table 3.

The CoS values set using this procedure apply to all packets that ingress the specified interfaces, not just to the class-default packets. If you set the CoS value, the device modifies the value before ingress queuing and scheduling so the CoS-modified packets are classified differently.

Note If you want to change the system-defined queuing class maps, you must either modify the configured queuing policies or create new queuing policies and attach these policies to the affected interfaces. If you fail to do so, you can render the default queuing or the configured queuing policies invalid, which might affect the interfaces in multiple VDCs.

Note When DSCP is enabled and configured on a queue, and if the port is untrusted, and an ingress port cos is configured, DSCP is considered for queuing.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.

	Command or Action	Purpose
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode. **Note** To configure port CoS, you can use only an ingress default system-defined queue type.
Step 4	switch(config-pmap-c-que)# **set cos** *value*	Sets the CoS field in all ingress packet to the value specified. The range is from 0 to 7.
Step 5	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 6	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 7	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Modifying Queuing Class Maps for COS

You can modify the CoS values that are matched by system-defined class maps. See Table 3 which lists the default system-defined CoS values.

The system-defined class maps can be changed only from the default VDC. Changes occur immediately and are applied to all ports on all VDCs that use the modified class map.

Note When you modify a system-defined class map, the changes occur immediately and might disrupt traffic on all VDCs that use the modified class map.

Note Defining the CoS with the match cos command is not supported for custom configured class-maps.

The device automatically modifies the CoS values that you configured in other queues so that each CoS value appears only once in the queues of the same type.

Before you begin

Ensure that you are in the default VDC for the device.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **class-map type queuing match-any** {*class-queuing-name* | *WORD*}	Configures the class map of type queuing, and then enters class-map queuing mode. The match on WORD is used for defining hierarchical class-maps in a queuing policy. The argument, *WORD*, is supported only on the F-Series Modules.
Step 3	switch(config-cmap-que)# **match cos** *value-range*	Sets the CoS value range matched by this queue. You can specify a range of values by using a hyphen between the beginning and ending values and a comma between values. The range is from 0 to 7. **Note** Repeat Steps 2 and 3 to modify CoS values for additional queues.
Step 4	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 5	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 6	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Modifying Queuing Class Maps for DSCP

You can modify the DSCP values that are matched by system-defined class maps. The system-defined class maps can be changed only from the default VDC. Changes occur immediately and are applied to all ports on all VDCs that use the modified class map.

To allow the modifications to be implemented, the DSCP-to-queue mapping must be enabled. If you have not enabled the DSCP-to-queue mapping earlier, you can use the **hardware qos dscp-to-queue ingress module type** command to enable DSCP-to-queue mapping.

To disable the DSCP mapping, use the **no hardware qos dscp-to-queue ingress** command.

Note When you modify a system-defined class map, the changes occur immediately and might disrupt traffic on all VDCs that use the modified class map.

The device automatically modifies the DSCP values that you configured in other queues so that each DSCP value appears only once in the queues of the same type.

Before you begin

Ensure that you are in the default VDC for the device.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	(Optional) switch(config)# **hardware qos dscp-to-queue ingress module type** {**all** | **f-series** | **m-series**}	Enables the dscp-to-queue mapping on the specified module(s). Use the **hardware qos dscp-to-queue ingress module type** command if you have not enabled dscp-to-queue mapping.
Step 3	switch(config)# **class-map type queuing match-any** *class-queuing-name*	Configures the class map of type queuing, and then enters class-map queuing mode.
Step 4	switch(config-cmap-que)# **match dscp** *value-range*	Sets the DSCP value range matched by this queue. You can specify a range of values by using a hyphen between the beginning and ending values and a comma/space between values. The range is from 0 to 63. **Note** Repeat Steps 5 and 6 to modify DSCP values for additional queues
Step 5	switch(config-cmap-que)# **exit**	Exits class-map queue mode, and enters global configuration mode.

Configuring Congestion Avoidance

You can configure congestion avoidance with tail drop or WRED features. Both features can be used in ingress and egress policy maps.

Note WRED and tail drop cannot be configured in the same class.

Configuring Tail Drop by COS Values

You can configure tail drop on both ingress and egress queues by setting thresholds by CoS values. The device drops packets that exceed the thresholds. You can specify a threshold based on the queue size or buffer memory that is used by the queue.

Note You cannot configure the queue size on ingress 10-Gigabit Ethernet ports except for the 8-port, 10-Gigabit Ethernet I/O module.

You use the system-defined queue classes for the type of module to which you want to apply the policy map. See Table 3.

Note WRED and tail drop cannot be configured in the same class.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode.
Step 4	switch(config-pmap-c-que)# **queue-limit cos** *value* {*threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | *percent percent_of_queuelimit*}	Assigns a tail drop threshold based on the queue size or percentage of the buffer memory that is used by the queue. The device drops packets that exceed the specified threshold. You can configure the threshold by the number of packets, number of bytes, or the duration of time at the underlying interface minimum guaranteed link rate. The default threshold is in packets. The size is from 1 to 83886080. The duration is from 1 to 83886080. The percentage is from 1 to 100. **Note** Repeat Step 4 to assign tail drop thresholds for other CoS values. Repeat Steps 3 through 5 to assign tail drop thresholds for other queue classes.
Step 5	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 6	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 7	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Configuring Tail Drop by DSCP Values

You can configure tail drop on ingress queues by setting thresholds by DSCP values. The device drops packets that exceed the thresholds. You can specify a threshold based on the queue size or buffer memory that is used by the queue.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** {*queuing-policy-map-name*}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name you specify.
Step 3	switch(config)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode.
Step 4	switch(config-pmap-c-que)# **queue-limit dscp** *value* {*queue-size* | **percent** *percent_of_queuelimit*}	Assigns a tail drop threshold based on the queue size or percentage of the buffer memory that is used by the queue. The device drops packets that exceed the specified threshold. Repeat Step 4 to assign tail drop thresholds for other DSCP values. Repeat Steps 3 through 5 to assign tail drop thresholds for other DSCP queue classes.
Step 5	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 6	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name*	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 7	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Configuring WRED by COS Values

Before configuring WRED, ensure that the CoS values are there (see the “Modifying Queuing Class Maps for COS” section.

You can configure WRED on both ingress and egress queues to set minimum and maximum packet drop thresholds. The frequency of dropped packets increases as the queue size exceeds the minimum threshold. When the maximum threshold is exceeded, all packets for the CoS value are dropped.

Note You cannot configure WRED on ingress 10-Gigabit Ethernet ports except for the 8-port 10-Gigabit Ethernet I/O module.

You can configure WRED thresholds by the CoS value, and configure a single WRED threshold to use on all CoS values that you do not specifically configure.

Note WRED and tail drop cannot be configured in the same class.

You use the system-defined queue classes for the type of module to which you want to apply the policy map.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode.
Step 4	switch(config-pmap-c-que)# **random-detect cos-based** [**aggregate** [**minimum-threshold**] {*min-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*} [**maximum-threshold**] {*max-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*}]	Configures WRED for all CoS values not configured by a CoS-specific random-detect command. You can specify minimum and maximum thresholds used to drop packets from the queue. You can configure thresholds by the number of packets, number of bytes, the duration of time at the underlying interface minimum guaranteed link rate, or as the percentage of queue size. The minimum and maximum thresholds must be of the same type. If no aggregate arguments are supplied, no aggregate WRED is configured. The default threshold is in packets. The thresholds are from 1 to 83886080. The percentage range is from 1 to 100. **Note** You must enter this command, even if you enter the command with no values. **Note** You can specify only one **random-detect cos-based** command in a class.
Step 5	(Optional) switch(config-pmap-c-que)# **random-detect** {**cos** *cos-list* [**minimum-threshold**] {*min-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent**	Configures WRED for specific CoS values. You can specify minimum and maximum thresholds used to drop packets from the queue. You can configure thresholds by the number of packets,

	Command or Action	Purpose
	min-percent-of-qsize} [**maximum-threshold**] {*max-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*}}	number of bytes, the duration of time at the underlying interface minimum guaranteed link rate, or as the percentage of the queue size. The minimum and maximum thresholds must be of the same type. The default threshold is in packets. Thresholds are from 1 to 83886080. The percentage range is from 1 to 100. Optional: repeat Step 5 to configure WRED for other CoS values. Optional: repeat Steps 3 through 6 to configure WRED for other queuing classes.
Step 6	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 7	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 8	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Configuring WRED by DSCP Values

Before configuring WRED, ensure that the DSCP values are present (see the "Modifying Queuing Class Maps for DSCP" section).

You can configure WRED on ingress queues to set minimum and maximum packet drop thresholds. The frequency of dropped packets increases as the queue size exceeds the minimum threshold. When the maximum threshold is exceeded, all packets for the DSCP value are dropped.

You cannot configure WRED on ingress 10-Gigabit Ethernet ports except for the 8-port 10-Gigabit Ethernet I/O module.

You can configure WRED thresholds by the DSCP value, and configure a single WRED threshold to use on all DSCP values that you do not specifically configure.

Note WRED and tail drop cannot be configured in the same class.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore

	Command or Action	Purpose
		characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode.
Step 4	switch(config-pmap-c-que)# **random-detect dscp-based** [**aggregate** [**minimum-threshold**] {*min-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*} [**maximum-threshold**] {*max-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*}]	Configures WRED for all DSCP values not configured by a DSCP-specific random-detect command. You can specify minimum and maximum thresholds used to drop packets from the queue. You can configure thresholds by the number of packets, number of bytes, the duration of time at the underlying interface minimum guaranteed link rate, or as the percentage of queue size. The minimum and maximum thresholds must be of the same type. If no aggregate arguments are supplied, no aggregate WRED is configured. The default threshold is in packets. The thresholds are from 1 to 52428800. The percentage range is from 1 to 100. **Note** You must enter this command, even if you enter the command with no values. **Note** You can specify only one **random-detect cos-based** command in a class.
Step 5	(Optional) switch(config-pmap-c-que)# **random-detect** {**dscp** *dscp-value* [**minimum-threshold**] {*min-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*} [**maximum-threshold**] {*max-threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *min-percent-of-qsize*}}	Configures WRED for specific DSCP values. You can specify minimum and maximum thresholds used to drop packets from the queue. You can configure thresholds by the number of packets, number of bytes, the duration of time at the underlying interface minimum guaranteed link rate, or as the percentage of the queue size. The minimum and maximum thresholds must be of the same type. The default threshold is in packets. Thresholds are from 1 to 52428800. The percentage range is from 1 to 100. Optional: repeat Step 5 to configure WRED for other DSCP values. Optional: repeat Steps 3 through 6 to configure WRED for other DSCP queuing classes.
Step 6	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.

Configuring Congestion Management

You can configure only one of the following congestion management methods in a policy map:

- Allocate a minimum data rate to a queue by using the bandwidth and bandwidth remaining commands.
- Allocate all data for a class of traffic to a priority queue by using the **priority** command. You can use the **bandwidth remaining** command to distribute remaining traffic among the nonpriority queues. By default, the system evenly distributes the remaining bandwidth among the nonpriority queues.
- Allocate a maximum data rate to a queue by using the **shape** command.

In addition to the congestion management feature that you choose, you can configure one of the following queue features in each class of a policy map:

- Taildrop thresholds based on the queue size and the queue limit usage. For more information, see the "Configuring Tail Drop by COS Values" section and "Configuring Tail Drop by DSCP Values" section.
- WRED for preferential packet drops based on CoS. For more information, see the "Configuring WRED by COS Values" section and "Configuring WRED by DSCP Values" section.

Configuring Bandwidth and Bandwidth Remaining

You can configure the bandwidth and bandwidth remaining on both ingress and egress queues to allocate a minimum percentage of the interface bandwidth to a queue. You use the system-defined ingress or egress queue class for the type of module to which you want to apply the policy map. For the list of system-defined ingress or egress queue classes for each module, see Table 3.

Note When a guaranteed bandwidth is configured, the priority queue must be disabled in the same policy map.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode.
Step 4	switch(config-pmap-c-que)# **bandwidth** {**rate** [**bps** | **kbps** | **mbps** | **gbps**] | **percent**} • **bandwidth remaining percent** *percent*	Assigns a minimum rate of the interface bandwidth to an output queue. You can configure a data rate by the bit rate or as the

	Command or Action	Purpose
	Assigns the percentage of the bandwidth that remains to this queue. The range is from 0 to 100.	percentage of the underlying interface link rate. The default units are kbps. The data rate is from 1 to 10,000,000,000. The percentage range is from 1 to 100. You can use only the **percent** keyword for interfaces set to autonegotiate. Repeat Steps 3 to 4 to assign bandwidth or bandwidth remaining for other queuing classes.
Step 5	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 6	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 7	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Configuring Priority

If you do not specify the priority, the system-defined egress pq queues behave as normal queues. For information on the system-defined type queuing class maps, see "Using Modular QoS CLI."

You can configure only one level of priority on an egress priority queue. You use the system-defined priority queue class for the type of module to which you want to apply the policy map. For the list of available system-defined class maps for each module, see Table 3.

For the nonpriority queues, you can configure how much of the remaining bandwidth to assign to each queue. By default, the device evenly distributes the remaining bandwidth among the nonpriority queues.

Note When a priority queue is configured, the other queues can only use the remaining bandwidth in the same policy map.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.

	Command or Action	Purpose
Step 3	switch(config-pmap-c-que)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode. You must select one of the system-defined priority queues.
Step 4	switch(config-pmap-c-que)# **priority** [**level** *value*]	Selects this queue as a priority queue. Only one priority level is supported.
Step 5	switch(config-pmap-c-que)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode. Choose a nonpriority queue where you want to configure the remaining bandwidth. By default, the system evenly distributes the remaining bandwidth among the nonpriority queues.
Step 6	switch(config-pmap-c-que)# **bandwidth remaining percent** *percent*	Assigns the percentage of the bandwidth that remains to this queue. The range is from 0 to 100. Repeat Steps 5 to 6 to assign bandwidth remaining for the other nonpriority queues.
Step 7	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 8	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 9	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Configuring Shaping

Note The device forces the shape rate to the closest value in the following percentage intervals: 100, 50, 33, 25, 12.5, 6.25, 3.13, or 1.07.

You can configure shaping on an egress queue to impose a maximum rate on it. You use the system-defined egress queue class for the type of module to which you want to apply the policy map. For the list of available system-defined class maps for each module, see Table 3.

Note Configuring shaping for a queue is independent of priority or bandwidth in the same policy map.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config-pmap-c-que)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode. You must select one of the system-defined priority queues.
Step 4	switch(config-pmap-c-que)# **shape** [**average**] {*rate* [**bps** | **kbps** | **mbps** | **gbps**] | **percent** *percent*}	Assigns a maximum rate on an output queue. You can configure a data rate by the bit rate or as a percentage of the underlying interface link rate. The default bit rate is in bits per second (bps). The data rate is from 8000 bps to 10 gbps. The percentage range is from 1 to 100. **Note** You can use only the **percent** keyword for interfaces set to autonegotiate. Repeat Steps 3 to 4 to configure shaping for other queuing classes.
Step 5	switch(config-pmap-c-que)# **exit**	Exits policy-map queue mode, and enters global configuration mode.
Step 6	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 7	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Configuring Queue Limits

You can configure the queue limit on both ingress and egress queues. The device drops any packets that exceed the queue limit. You use the system-defined queue classes for the type of module to which you want to apply the policy map. See Table 3.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.

	Command or Action	Purpose
Step 2	switch(config)# **policy-map type queuing** [**match-first**] {*policy-map-name* | **que-dynamic**}	Configures the policy map of type queuing, and then enters policy-map mode for the policy-map name that you specify. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
Step 3	switch(config-pmap-que)# **class type queuing** *class-queuing-name*	Configures the class map of type queuing, and then enters policy-map class queuing mode. You must select one of the system-defined priority queues.
Step 4	switch(config-pmap-c-que)# **queue-limit** {*threshold* [**packets** | **bytes** | **kbytes** | **mbytes** | **ms** | **us**] | **percent** *percent_of_queuelimit*}	Assigns a queue limit based on the queue size or percentage of the buffer memory used by the queue. The device will drop packets that exceed the specified threshold. You can configure the threshold by the number of packets, number of bytes, or the duration of time at the underlying interface minimum guaranteed link rate. The default threshold is in packets. The size is from 1 to 83886080. The duration is from 1 to 83886080. The percentage range is from 1 to 100.
Step 5	switch(config-pmap-c-que)# **exit**	Exits class-map queue mode and enters policy-map queue mode.
Step 6	switch(config-pmap-que)# **exit**	Exits policy-map queue mode and enters global configuration mode.
Step 7	(Optional) switch(config)# **show policy-map type queuing** [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
Step 8	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Enabling DSCP to Queue Mapping

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **hardware qos dscp-to-queue ingress module type** {**all** | **f-series** | **m-series** }	Enables the dscp-to-queue mapping on the specified module(s).

	Command or Action	Purpose
		Note Starting with Cisco NX-OS Release 8.0(1), use the **f-series** keyword to enable dscp-to-queue mapping on M3-Series I/O modules.
Step 3	(Optional) switch(config)# **show hardware qos dscp-to-queue ingress**	Displays information about the status of dscp-to-queue mapping in ingress direction.
Step 4	(Optional) switch(config)# **copy running-config startup-config**	Saves the running configuration to the startup configuration.

Verifying the Queuing and Scheduling Configuration

To configure queuing and scheduling, perform one of the following tasks:

Command	Purpose
show class-map type queuing[*class-queuing-name*]	Displays information about all configured class maps or a selected class map of type queuing. Class queuing names are listed in Table 3.
show policy-map type queuing [*policy-map-name* | **que-dynamic**]	Displays information about all configured policy maps or a selected policy map of type queuing.
show policy-map system	Displays information about all configured policy maps on the system.
show queuing interface ethernet *slot*/*port*[**.***subinterface* [*module* | **summary**]	Displays queuing information about the specified interface.

For more information about the fields in the output from these commands, see the *Cisco Nexus 7000 Series NX-OS Quality of Service Command Reference*.

Configuration Examples for Queuing and Scheduling

Example: Setting Ingress Port CoS Configuration

Note Setting the ingress port CoS value makes the specified interfaces untrusted.

Note Ensure that you are using the default queue for the port type that you are configuring. For information on the default queue for the port types, see "Using Modular QoS CLI."

The following example shows how to configure ingress port CoS for 1-Gigabit Ethernet ports:

```
configure terminal
 policy-map type queuing untrusted_port_cos
  class type queuing 2q4t-in-q-default
   set cos 5
 interface ethernet 2/1
  service-policy type queuing input untrusted_port_cos
```

The following example shows how to configure ingress port CoS for 10-Gigabit Ethernet ports:

```
configure terminal
 policy-map type queuing untrusted_port_cos
  class type queuing 8q2t-in-q-default
   set cos 5
 interface ethernet 2/1
  service-policy type queuing input untrusted_port_cos
```

Example: Priority and Queue Limit Configuration

The following example shows how to configure the priority and queue limit features:

```
configure terminal
 class-map type queuing match-any 1p3q4t-out-pq1
  match cos 5-7
 class-map type queuing match-any 1p3q4t-out-q2
  match cos 3-4
 class-map type queuing match-any 1p3q4t-out-q3
  match cos 0-2
policy-map type queuing priority_queue1
 class type queue 1p3q4t-out-pq1
  priority
 class type queue 1p3q4t-out-q2
  bandwidth remaining percent 60
  queue-limit 1 mbytes
 class type queue 1p3q4t-out-q3
  bandwidth remaining percent 40
  queue-limit 2 mbytes
```

Example: Shaping and Tail Drop Configuration

The following example shows how to configure the shaping and tail drop features:

```
configure terminal
 class-map type queuing match-any 1p3q4t-out-pq1
  match cos 5-7
 class-map type queuing match-any 1p3q4t-out-q2
  match cos 3-4
policy-map type queuing shape_dt
 class type queue 1p3q4t-out-pq1
  shape percent 50
  queue-limit cos 5 percent 10
  queue-limit cos 6 percent 10
 class type queue 1p3q4t-out-q2
  shape percent 25
  queue-limit cos 4 percent 15
```

Note If the **priority** keyword is not specified for a **pq1** queue, the queue is considered as a normal queue, not a priority queue.

Example: Bandwidth and WRED Configuration

The following example shows how to configure the bandwidth and WRED features for COS queues:

```
configure terminal
 class-map type queuing match-any 1p3q4t-out-pq1
  match cos 5-7
 class-map type queuing match-any 1p3q4t-out-q2
  match cos 3-4
policy-map type queuing bandwidth_wred
 class type queuing 1p3q4t-out-pq1
  bandwidth percent 50
  random-detect cos-based
  random-detect cos 5 minimum-threshold percent 10 maximum-threshold percent 30
  random-detect cos 6 minimum-threshold percent 40 maximum-threshold percent 60
 class type queuing 1p3q4t-out-q2
  bandwidth percent 25
  random-detect cos-based
  random-detect cos 4 minimum-threshold percent 20 maximum-threshold percent 40
```

The following example shows how to configure the bandwidth and WRED features for DSCP queues:

```
configure terminal
 class-map type queuing match-any 8q2t-in-q1
  match dscp 5-6
 class-map type queuing match-any 8q2t-in-q2
  match dscp 0-4
policy-map type queuing dscp_wred
 class type queuing 8q2t-in-q1
  bandwidth percent 50
  random-detect dscp-based
  random-detect dscp 5 minimum-threshold percent 10 maximum-threshold percent 30
  random-detect dscp 6 minimum-threshold percent 40 maximum-threshold percent 60
 class type queuing 8q2t-in-q2
  bandwidth percent 25
  random-detect dscp-based
  random-detect dscp 4 minimum-threshold percent 20 maximum-threshold percent 40
```

Example: Verifying the Status of DSCP-to-queue Mapping

The following sample output from the **show hardware qos dscp-to-queue ingress** command displays the status of DSCP-to-queue mapping enabled in ingress direction on M-series modules:

```
Switch# show hardware qos dscp-to-queue ingress

status: Enabled
module_type : m-series
```

Feature History for Queuing and Scheduling

The table below summarizes the new and changed features for this document and shows the releases in which each feature is supported. Your software release might not support all the features in this document. For the latest caveats and feature information, see the Bug Search Tool at https://tools.cisco.com/bugsearch/ and the release notes for your software release.

Table 3: Feature History for Queuing and Scheduling

Feature Name	Release	Feature Information
DSCP to Queue Mapping	8.0(1)	Use the **hardware qos dscp-to-queue ingress module type f-series** command to enable dscp-to-queue mapping on M3-Series I/O modules
DSCP to Queue Mapping	6.2(2)	This feature was introduced.
System-defined queue types	6.2(2)	Updated the System-Defined Queue Types table with new system-defined queue types for 4q8q policy templates on the Cisco Nexus 7710 switch and the Cisco Nexus 7718 switch.
No change from Release 4.1(2)	5.1(1)	—