# Cisco Nexus 3000 Series NX-OS QoS Configuration Guide, Release 7.x

**First Published:** 2015-08-24

**Last Modified:** 2020-08-31

**Americas Headquarters**
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED “AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: http://www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

© 2017–2020 Cisco Systems, Inc. All rights reserved.

# CONTENTS


PREFACE

**Preface** vii

Audience vii

Document Conventions vii

Related Documentation for Cisco Nexus 3000 Series Switches viii

Documentation Feedback viii

Communications, Services, and Additional Information viii

CHAPTER 1

**New and Changed Information** 1

New and Changed Information 1

CHAPTER 2

**Overview** 7

Quality of Service Overview 7

CHAPTER 3

**Configuring QoS** 9

Information About Quality of Service 9

Modular QoS CLI 9

System Classes 10

Default System Classes 10

Information About Policy Types 11

MTU 16

Trust Boundaries 16

Ingress Classification Policies 17

Priority Groups for No-Drop Classes 17

Egress Queuing Policies 17

QoS for Traffic Directed to the CPU 18

QoS Configuration Guidelines and Limitations 18

Configuring System Classes 21
Configuring Class Maps 21
Configuring ACL Classification 22
Configuring CoS Classification 23
Configuring DSCP Classification 24
Configuring IP RTP Classification 25
Configuring Precedence Classification 26
Creating Policy Maps 27
Configuring Type QoS Policies 29
Configuring Type Network QoS Policies 30
Configuring Type Queuing Policies 31
Configuring an ECN Threshold 34
Configuring Pause Buffer Thresholds and Priority Groups 37
Information About Marking 39
Configuring CoS Marking 39
Configuring a DSCP Wildcard Mask 40
Configuring DSCP Marking 42
Configuring IP Precedence Marking 43
QoS Configurations for Layer 3 Routing 45
Configuring Layer 3 Multicast Queuing 45
Configuring a Service Policy for a Layer 3 Interface 45
Changing the Bandwidth Allocated to Unicast and Multicast Traffic 46
Attaching the System Service Policy 47
Restoring the Default System Service Policies 47
Enabling the Jumbo MTU 48
Verifying the Jumbo MTU 49
Configuring QoS on Interfaces 56
Configuring Untagged CoS 56
Configuring an Interface Service Policy 56
Verifying the QoS Configuration 57
Monitoring the QoS Packet Buffer 67

CHAPTER 4 **Configuring Priority Flow Control** 69
About Priority Flow Control 69

Guidelines and Limitations for Priority Flow Control 70
Default Settings for Priority Flow Control 72
Enabling Priority Flow Control on a Traffic Class 72
Configuring Priority Flow Control 74
Reserving mmu-buffer for PFC 75
Configuring a Priority Flow Control Watchdog Interval 75
Verifying the Priority Flow Control Configuration 78
Monitoring PFC Frame Counter Statistics 78
Configuration Examples for Priority Flow Control 79

CHAPTER 5 **Configuring Policing** 81
About Policing 81
Licensing Requirements for Policing 81
Prerequisites for Policing 82
Guidelines and Limitations 82
Configuring Policing 82
Configuring 1-Rate and 2-Rate, 2-Color and 3-Color Policing 83
Configuring Ingress and Egress Policing 87
Configuring Markdown Policing 87
Verifying the Policing Configuration 89
Configuration Examples for Policing 89

CHAPTER 6 **Configuring Traffic Shaping** 91
About Traffic Shaping 91
Guidelines and Limitations for Traffic Shaping 92
Configuring Traffic Shaping 92
Verifying Traffic Shaping 93
Configuration Example for Traffic Shaping 93

# Preface

This preface includes the following sections:


- Audience, on page vii
- Document Conventions, on page vii
- Related Documentation for Cisco Nexus 3000 Series Switches, on page viii
- Documentation Feedback, on page viii
- Communications, Services, and Additional Information, on page viii


## Audience

This publication is for network administrators who install, configure, and maintain Cisco Nexus switches.

## Document Conventions

Command descriptions use the following conventions:

| Convention | Description |
|---|---|
| **bold** | Bold text indicates the commands and keywords that you enter literally as shown. |
| *Italic* | Italic text indicates arguments for which the user supplies the values. |
| [x] | Square brackets enclose an optional element (keyword or argument). |
| [x \| y] | Square brackets enclosing keywords or arguments separated by a vertical bar indicate an optional choice. |
| {x \| y} | Braces enclosing keywords or arguments separated by a vertical bar indicate a required choice. |
| [x {y \| z}] | Nested set of square brackets or braces indicate optional or required choices within optional or required elements. Braces and a vertical bar within square brackets indicate a required choice within an optional element. |

| Convention | Description |
|---|---|
| `variable` | Indicates a variable for which you supply values, in context where italics cannot be used. |
| string | A nonquoted set of characters. Do not use quotation marks around the string or the string will include the quotation marks. |

Examples use the following conventions:

| Convention | Description |
|---|---|
| `screen font` | Terminal sessions and information the switch displays are in screen font. |
| **`boldface screen font`** | Information you must enter is in boldface screen font. |
| *italic screen font* | Arguments for which you supply values are in italic screen font. |
| < > | Nonprinting characters, such as passwords, are in angle brackets. |
| [ ] | Default responses to system prompts are in square brackets. |
| !, # | An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line. |

# Related Documentation for Cisco Nexus 3000 Series Switches

The entire Cisco Nexus 3000 Series switch documentation set is available at the following URL:

https://www.cisco.com/c/en/us/support/switches/nexus-3000-series-switches/tsd-products-support-series-home.html

# Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to nexus3k-docfeedback@cisco.com. We appreciate your feedback.

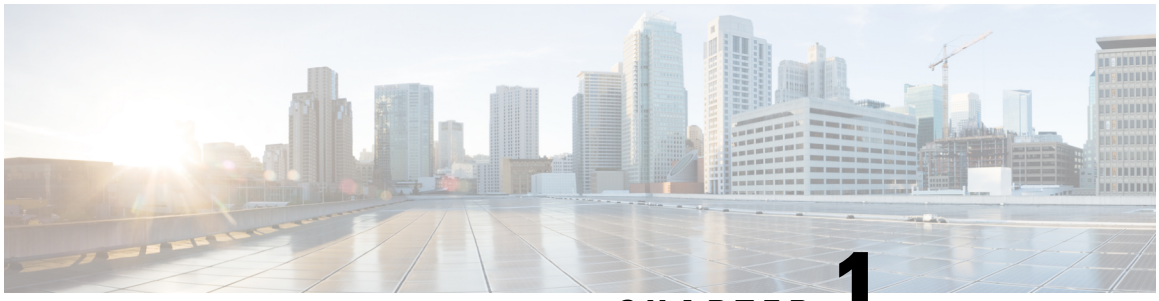
# Communications, Services, and Additional Information

- To receive timely, relevant information from Cisco, sign up at Cisco Profile Manager.
- To get the business impact you're looking for with the technologies that matter, visit Cisco Services.
- To submit a service request, visit Cisco Support.
- To discover and browse secure, validated enterprise-class apps, products, solutions and services, visit Cisco DevNet.
- To obtain general networking, training, and certification titles, visit Cisco Press.
- To find warranty information for a specific product or product family, access Cisco Warranty Finder.

### Cisco Bug Search Tool

Cisco Bug Search Tool (BST) is a web-based tool that acts as a gateway to the Cisco bug tracking system that maintains a comprehensive list of defects and vulnerabilities in Cisco products and software. BST provides you with detailed defect information about your products and software.

Preface



CHAPTER 1

# New and Changed Information

This chapter contains the following sections:


- New and Changed Information, on page 1


## New and Changed Information

See the following sections:

| Feature | Description | Added or Changed in Release | Where Documented |
|---|---|---|---|
| DSCP Wildcard Mask | Support for an additional filter when classifying traffic that is based on the DSCP value. | 7.0(3)I7(9) | Configuring a DSCP Wildcard Mask, on page 40 |
| Enabling syslog messages to account packet drops | Support to enable syslog messages to account packet drops on multicast queues for no-drop class. | 7.0(3)I7(8) | Guidelines and Limitations for Priority Flow Control, on page 70 |
| Write message to syslog when queue is stuck | Support to write message to syslog when queue is stuck. | 7.0(3)I7(4) | Guidelines and Limitations for Priority Flow Control, on page 70 |
| Drop PFC traffic and all PAUSE frames | Support to drop PFC traffic and drop all PAUSE frames. | 7.0(3)I7(4) | Guidelines and Limitations for Priority Flow Control, on page 70 |
| Configure PFC watchdog timers and multiplier per interface | Support to configure PFC watchdog timers and multiplier per interface. | 7.0(3)I7(4) | Configuring a Priority Flow Control Watchdog Interval, on page 75 |
| Drop multicast/broadcast traffic on no-drop configured class | Support to drop multicast/broadcast traffic on no-drop configured class. | 7.0(3)I7(4) | Guidelines and Limitations for Priority Flow Control, on page 70 |

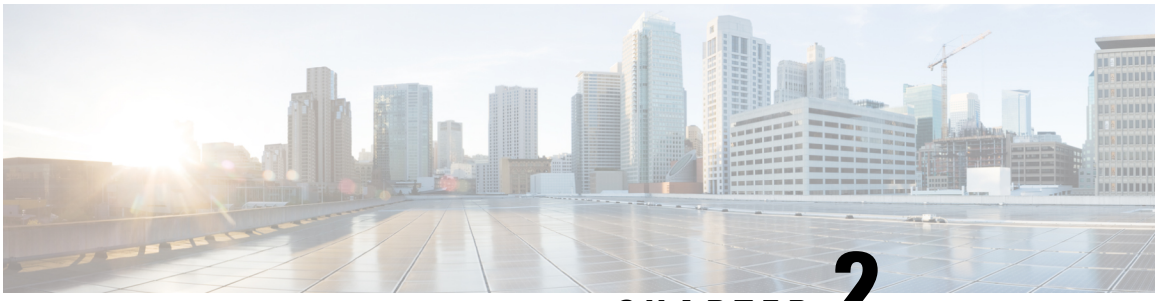
| Feature | Description | Added or Changed in Release | Where Documented |
|---|---|---|---|
| Create log entry when PFC is received on lossy group. | Support to create log entry when PFC is received on lossy (non-configured) group. | 7.0(3)I7(4) | Guidelines and Limitations for Priority Flow Control, on page 70 |
| Enable log-only option for PFC | Support for Enable log-only option for PFC. | 7.0(3)I7(4) | Configuring a Priority Flow Control Watchdog Interval, on page 75 |
| PFC watchdog Ingress drops: | Added support to count both egress and ingress drops using the show command on Cisco Nexus 3000 Series switches. | 7.0(3)I6(1) | Configuring a Priority Flow Control Watchdog Interval, on page 75 |
| Ingress and egress policing | Added ingress and egress policing support | 7.0(3)I5(1) | Guidelines and Limitations, on page 82 |
| DCBX Support | Changed the behaviour of the **no lldp tlv-select dcbxp** command is enhanced so that the PFC is disabled for interfaces on both sides of back-to-back switches | 7.0(3)I4(2) | Guidelines and Limitations for Priority Flow Control |
| Two additional options, module and summary displayed in the output of the **show queuing interface ethernet slot/chassis_number** command. | The **show queuing interface ethernet slot/chassis_number** command displays two more options: module and summary in the output. These additional options do not have any functionality impact. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |
| New switch prompt for **policy-map type network-qos** is switch(config-pmap-nqos)#. | The switch prompt for **policy-map type network-qos** was switch(config-pmap-nq)# in earlier releases. The new switch prompt for **policy-map type network-qos** is switch(config-pmap-nqos)#. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18<br>Verifying the QoS Configuration, on page 57<br>Enabling the Jumbo MTU, on page 48<br>Enabling Priority Flow Control on a Traffic Class, on page 72 |

| Feature | Description | Added or Changed in Release | Where Documented |
|---|---|---|---|
| The network-qos class-map is automatically created when it is used in the policy-map. | If a non-existing class is configured under a policy-map for network-qos, a new class-map is created and the prompt changes from (config-pmap-nqos) to (config-cmap-nqos). | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |
| Updated the output of the command **show queuing interface**. | The command **show queuing interface** displays the queues even without applying the network-qos policy. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |
| The output of the **show queuing interface** command displays an additional option. | The output of the **show queuing interface** command displays an option for the internal HiGig2 interface. These interfaces are not relevant and the additional option does not have any functional impact. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |
| Updated the output format and the fields of **show policy-map interface <> type queuing** CLI command. | The output format and the fields of **show policy-map interface <> type queuing** CLI command have been updated. For Class-map (queuing), the following fields are displayed: policy, bandwidth percent, queue dropped packets, and queue depth in bytes. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |

| Feature | Description | Added or Changed in Release | Where Documented |
|---|---|---|---|
| Removing the default bandwidth configuration sets the bandwidth to the default value 100%. | Removing the default bandwidth configuration from the default queuing class used to set the bandwidth to 50% in prior releases. Removing the default bandwidth configuration sets the bandwidth to the default value, 100%. You can set the bandwidth to 50% by configuring the CLI command **bandwidth percent 50**. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |
| Configuring priority levels 2 and 3 in the pmap configuration. | Prior to Release 7.0(3)I2(1), only priority level 1 was supported in the pmap configuration. Starting with Release 7.0(3)I2(1), you can configure priority levels 2 and 3 in the pmap configuration. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18 |
| Updated output format and the fields of the **show queuing interface eth** <> CLI command. | The output format and the fields of the **show queuing interface eth** <> CLI command have been updated. The output displays all qos-groups, control qos group, SPAN qos-group, and pfc statistics. The xon drops, xoff drops, and the HW MTU fields are not displayed in the new format. | 7.0(3)I2(1) | QoS Configuration Guidelines and Limitations, on page 18<br>Verifying the Jumbo MTU, on page 49 |
| Reserving mmu-buffer for Priority Flow Control. | Reserve the mmu-buffer for Priority Flow Control. | 7.0(3)I2(1) | Reserving mmu-buffer for PFC, on page 75 |
| Displaying the configured HW MTU for the qos-groups. | Use the **show policy-map system type network-qos** command to display the configured HW MTU for the qos-groups. | 7.0(3)I2(1) | Verifying the Jumbo MTU, on page 49 |

| Feature | Description | Added or Changed in Release | Where Documented |
|---|---|---|---|
| Updated MTU values | Updated MTU values on 10-Gigabit ports and on 40-Gigabit ports. | 7.0(3)I2(1) | Configuring Pause Buffer Thresholds and Priority Groups, on page 37 |

New and Changed Information



CHAPTER 2

# Overview

This chapter contains the following sections:

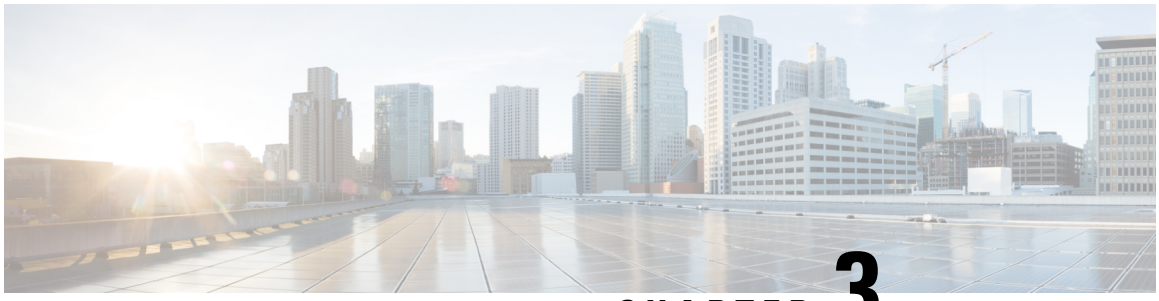

- Quality of Service Overview, on page 7


# Quality of Service Overview

This document describes the configurable Cisco NX-OS Quality of Service (QoS) features. You use the QoS features to provide the most desirable flow of traffic through a network. QoS allows you to classify the network traffic, prioritize the traffic flow, and provide congestion avoidance. The control of traffic is based on the fields in the packets that flow through the system. You use the Modular QoS CLI (MQC) to create the traffic classes and policies of the QoS features.

QoS features are applied using QoS policies and queuing policies, as follows:

- QoS policies include classification and marking features.
- Queuing policies use the queuing and scheduling features at egress. At ingress, they are used to configure buffer threshold and priority group mapping.
- Network QoS policies include configuring maximum transmission unit (MTU), pause no-drop, and queue-limit. It is also used to configure explicit congestion notification (ECN) and weighted random early detection (WRED).
- Priority flow control.



CHAPTER 3

# Configuring QoS

This chapter contains the following sections:


- Information About Quality of Service, on page 9
- QoS Configuration Guidelines and Limitations, on page 18
- Configuring System Classes, on page 21
- Configuring QoS on Interfaces, on page 56
- Verifying the QoS Configuration, on page 57
- Monitoring the QoS Packet Buffer, on page 67


# Information About Quality of Service

The configurable Cisco NX-OS quality of service (QoS) features allow you to classify the network traffic, prioritize the traffic flow, and provide congestion avoidance.

The default QoS configuration on the device provides best-effort service for Ethernet traffic. QoS can be configured to provide additional classes of service for Ethernet traffic. Cisco NX-OS QoS features are configured using Cisco Modular QoS CLI (MQC).

In the event of congestion or collisions, Ethernet will drop packets. The higher level protocols detect the missing data and retransmit the dropped packets.

## Modular QoS CLI

The Cisco Modular QoS CLI (MQC) provides a standard set of commands for configuring QoS.

You can use MQC to define additional traffic classes and to configure QoS policies for the whole system and for individual interfaces. Configuring a QoS policy with MQC consists of the following steps:

1. Define traffic classes.
2. Associate policies and actions with each traffic class.
3. Attach policies to logical or physical interfaces as well as at the global system level.

MQC provides two command types to define traffic classes and policies:

**class-map**

Defines a class map that represents a class of traffic based on packet-matching criteria. Class maps are referenced in policy maps.

The class map classifies incoming packets based on matching criteria, such as the IEEE 802.1p class of service (CoS) value. Unicast and multicast packets are classified.

**policy-map**

Defines a policy map that represents a set of policies to be applied on a class-by-class basis to class maps.

The policy map defines a set of actions to take on the associated traffic class, such as limiting the bandwidth or dropping packets.

You define the following class-map and policy-map object types when you create them:

**network-qos**

Defines MQC objects that you can use for system level related actions.

**qos**

Defines MQC objects that you can use for classification.

**queuing**

Defines MQC objects that you can use for queuing and scheduling at egress and for configuring buffer threshold and priority group mapping at ingress.

**Note** The qos type is the default for the **class-map** and **policy-map** commands, but not for the **service-policy** which requires that you specify an explicit type.

You can attach policies to interfaces or EtherChannels as well as at the global system level by using the **service-policy** command.

You can view all or individual values for MQC objects by using the **show class-map** and **show policy-map** commands.

An MQC target is an entity (such as an Ethernet interface) that represents a flow of packets. A service policy associates a policy map with an MQC target and specifies whether to apply the policy on incoming or outgoing packets. This mapping enables the configuration of QoS policies such as marking, bandwidth allocation, buffer allocation, and so on.

# 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.

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

# Default System Classes

The device provides the 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** By default, the Cisco Nexus 3000 Series devices only have a class-map, and policy map for the default queue. If more queues are required, you need to create your own configurations for class-map, qos-group, and policy-map.

# Information About Policy Types

The device supports a number of policy types. You create class maps in the policy types.

There are three policy types:

- Network-qos
- Queuing
- QoS

The following QoS parameters can be specified for each type of class:

- Type network-qos—A network-qos policy is used to instantiate system classes and associate parameters with those classes that are of system-wide scope.
  - Classification—The traffic that matches this class is as follows:
    - QoS Group—A class map of type network-qos identifies a system class and is matched by its associated qos-group.
  - Policy—The actions that are performed on the matching traffic are as follows:

**Note** A network-qos policy can only be attached to the system QoS target.

    - MTU—The MTU that needs to be enforced for the traffic that is mapped to a system class.

**Note** The Cisco Nexus device supports one MTU for traffic for all classes for all ports. However, you can have different MTUs for different classes. The MTUs are used for PFC buffer calculation.

    - Set CoS value—This configuration is used to mark 802.1p values for all traffic mapped to this system class.
    - Congestion Control WRED—Weighted random early detection (WRED) anticipates and avoids congestion before congestion occurs. WRED drops packets, based on the average queue length that exceeds a specific threshold value, to indicate congestion. You can configure congestion avoidance with WRED in egress policy maps. By default, tail-drop is the congestion control mechanism. To enable WRED, use the **congestion-control random-detect** command in network-qos policy map mode.

- ECN—ECN is an extension to WRED that marks packets instead of dropping them when the average queue length exceeds a specific threshold value. When configured with the WRED explicit congestion notification (ECN) feature, routers and end hosts use this marking as a signal that the network is congested to slow down sending packets. To enable an ECN, use the **congestion-control random-detect ecn** command in the network-qos policy map mode.

  ECN is supported on all types of Cisco Nexus 3000 series switches.

**Note** Enabling WRED and ECN on a class on a network-qos policy implies that WRED and ECN is enabled for all ports in the system.

- No drop—No drop specifies lossless service for the system class.

- Type queuing—The Cisco Nexus device supports type queuing in the ingress and egress directions. Egress type queuing policies are used to define the scheduling characteristics of the queues. Ingress type queuing policies are used to define the pause buffer thresholds, priority group, and queue limit.

**Note** Some configuration parameters when applied to an EtherChannel are not reflected on the configuration of the member ports.

- Classification—The traffic that matches this class is as follows:
  - QoS Group—A class map of type queuing identifies a system class and is matched by its associated QoS group.

- Policy—The actions that are performed on the matching traffic are as follows:

**Note** These policies can be attached to the system qos target or to any interface.

  - Egress queuing policy—The egress queuing policy is used to configure egress queues on the device.
    - Bandwidth—Sets the guaranteed scheduling deficit weighted round robin (DWRR) percentage for the system class.
    - Priority—Sets a system class for strict-priority scheduling. Only one system class can be configured for priority in a given queuing policy. For Cisco Nexus 3132 and 3172 switches, there are three strict priority levels.
    - Shape and minimum guarantee—Specifies the burst size and minimum guaranteed bandwidth for this queue.
    - Queue limit—Specifies either the static or dynamic queue limit for Cisco Nexus 3100 Series switches. The static queue limit defines the fixed size to which the queue can grow.

  - Ingress queuing policy—The ingress queuing policy is used to define the pause buffer thresholds, priority group, and queue limit.

- Pause buffer threshold—Sets the pause and resume buffer threshold settings for ingress traffic.
- Priority group—Classifies the traffic and monitors statistics on no-drop classes.
- Queue limit—Sets the shared buffer usage per priority group.

You can configure the threshold for using shared buffers both at ingress and egress based on the alpha value, which is derived from the index. The index ranges from 0 to 9 for Cisco Nexus 3000 series switches and from 0 to 10 for Cisco Nexus 3100 platform switches. At ingress, the alpha value is used to calculate the per port, per priority group share of the buffers available from the current free pool. At egress, the alpha value is used to calculate the per port, per queue share of the buffers available from the current free pool.

For the Cisco Nexus 3000 series switches, the alpha values are as follows:

| Index | Alpha Value |
|---|---|
| 0 | 1/64 |
| 1 | 1/32 |
| 2 | 1/16 |
| 3 | 1/8 |
| 4 | 1/4 |
| 5 | 1/2 |
| 6 | 1 |
| 7 | 2 |
| 8 | 4 |
| 9<br>**Note**<br>Index 9 is not applicable for Cisco Nexus 34180YC.<br>Cisco Nexus 34180YC has an index range from 0 to 8. | 8 |

For the Cisco Nexus 3100 platform switches, the alpha values are as follows:

| Index | Alpha Value |
|---|---|
| 0 | 1/128 |
| 1 | 1/64 |
| 2 | 1/32 |
| 3 | 1/16 |
| 4 | 1/8 |
| 5 | 1/4 |

| Index | Alpha Value |
|---|---|
| 6 | 1/2 |
| 7 | 1 |
| 8 | 2 |
| 9<br>**Note**<br>Index 9 is not applicable for Cisco Nexus 34180YC.<br>Cisco Nexus 34180YC has an index range from 0 to 8. | 4 |
| 10<br>**Note**<br>Index 10 is not applicable for Cisco Nexus 34180YC.<br>Cisco Nexus 34180YC has an index range from 0 to 8. | 8 |

- To determine the default shared alpha value, use one of the following commands:
  - For Cisco NX-OS 6.x, use the **test hardware internal bcm-usd bcm-diag-shell** command to determine the default shared alpha value.

    For example:

```
switch# test hardware internal bcm-usd bcm-diag-shell
Available Unit Numbers: 0
bcm-shell.0> d chg MMU_THDU_XPIPE_CONFIG_QUEUE
MMU_THDU_XPIPE_CONFIG_QUEUE.mmu0[0]:
<Q_SHARED_LIMIT_CELL=7,Q_SHARED_ALPHA_CELL=7,Q_MIN_LIMIT_CELL=0xb,
Q_LIMIT_DYNAMIC_CELL=1,Q_COLOR_LIMIT_DYNAMIC_CELL=1,DATA=0x000000000c00160007>
```

  - For Cisco NX-OS 7.x, use the **bcm-shell module 1** command to determine the default shared alpha value.

    For example:

```
switch# bcm-shell module 1
Available Unit Numbers: 0
bcm-shell.0> d chg MMU_THDU_XPIPE_CONFIG_QUEUE
MMU_THDU_XPIPE_CONFIG_QUEUE.mmu0[0]:
<Q_SHARED_LIMIT_CELL=7,Q_SHARED_ALPHA_CELL=7,Q_MIN_LIMIT_CELL=0xb,
Q_LIMIT_DYNAMIC_CELL=1,Q_COLOR_LIMIT_DYNAMIC_CELL=1,DATA=0x000000000c00160007>
```

- In dynamic mode, the amount of shared buffer a port can consume based on the alpha parameter is the queue's threshold: alpha value * (number of unused cells in the buffer). In general, as the number of unused cells decreases (ie. the buffer becomes fuller), the queue's threshold reduces.
- Type qos—A type QoS policy is used to classify traffic that is based on various Layer 2, Layer 3, and Layer 4 fields in the frame and to map it to system classes.

**Note** Some configuration parameters when applied to an EtherChannel are not reflected on the configuration of the member ports.

- Classification—The traffic that matches this class are as follows:
  - Access Control Lists—Classifies traffic based on the criteria in existing ACLs.
  - Class of Service—Matches traffic based on the CoS field in the frame header.
  - DSCP—Classifies traffic based on the Differentiated Services Code Point (DSCP) value in the DiffServ field of the IP header.
  - IP Real Time Protocol—Classifies traffic on the port numbers used by real-time applications.
  - Precedence—Classifies traffic based on the precedence value in the type of service (ToS) field of the IP header.
- Policy—The actions that are performed on the matching traffic are as follows:

**Note** This policy can be attached to the system or to any interface. It applies to input traffic only.

  - QoS Group—Sets the QoS group that corresponds to the system class this traffic flow is mapped to.

| | |
|---|---|
| The Cisco Nexus 3000 Series switches support: | • Eight QoS groups<br>• Eight queues for unicast<br>• Four queues for multicast<br>**Note**<br>Cisco Nexus 34180YC has eight queues for both unicast and multicast. |
| The Cisco Nexus 3100 platform switches support: | • Eight QoS groups<br>• Eight queues for unicast<br>• Eight queues for multicast<br>For Cisco Nexus 3100 platform switches, each QoS group is mapped to one multicast queue. The mapping is QoS group 0 mapped to multicast queue 1, QoS group 1 mapped to multicast queue 2, and so forth. |

# MTU

The Cisco Nexus device supports one MTU for all classes for all ports.

When configuring MTU, follow these guidelines:

- For the Cisco Nexus device, the MTU is controlled by the value configured on the class default.
- 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 1500 bytes and the maximum MTU is 9216 bytes.
- 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 default system class has a default MTU of 1500 bytes. You can configure this value.
- You can specify the MTU value for either a single Layer 3 interface or a range of Layer 3 interfaces. When you change the Layer 3 interface MTU value to the jumbo MTU value (1500 bytes or greater), you must also change the network QoS MTU value to 1500 bytes or greater.
- You can set the MTU per class of the network-qos policy. The MTU that is set is used to decide the buffer allocations for PFC. On a need basis, you can configure some classes to have an MTU of 9216 and some to have an MTU of 1500, depending on the type of traffic expected on that class. This will help the system configure the PFC buffers when a class is configured as a no-drop-class.
- On Cisco Nexus 3500 Switches, MTU for all classes must be same as the one configured for the default-class.
- Defining the MTU under the policy-map of type network-qos is mandatory for QoS to work:

```
class-map type network-qos nq-qos-3
  match qos-group 3
class-map type network-qos nq-qos-4
  match qos-group 4
class-map type network-qos nq-qos-5
  match qos-group 5

policy-map type network-qos nq-qos3-4-5
  class type network-qos nq-qos-3
    mtu 1500
  class type network-qos nq-qos-4
    mtu 1500
  class type network-qos nq-qos-5
    mtu 1500
  class type network-qos class-default
    mtu 1500
```

# Trust Boundaries

The trust boundary is enforced by the incoming interface as follows:

- By default, all Ethernet interfaces are trusted interfaces.The 802.1p CoS and DSCP are preserved unless the marking is configured. There is no default CoS to queue and DSCP to queue mapping. You can define and apply a policy to create these mappings. By default, without a user defined policy, all traffic is assigned to the default queue.

**Note** For Cisco Nexus 34180YC, all ports are trusted interfaces.

- Any packet that is not tagged with an 802.1p CoS value is classified into the default drop system class. If the untagged packet is sent over a trunk, it is tagged with the default untagged CoS value, which is zero.
- You can override the default untagged CoS value for an Ethernet interface or port channel.

After the system applies the untagged CoS value, QoS functions the same as for a packet that entered the system tagged with the CoS value.

# Ingress Classification Policies

You use classification to partition traffic into classes. You classify the traffic based on the packet property (CoS field) or the packet header fields that include IP precedence, Differentiated Services Code Point (DSCP), and Layer 2 to Layer 4 parameters. The values used to classify traffic are called match criteria.

Traffic that fails to match any class is assigned to a default class of traffic called class-default.

# Priority Groups for No-Drop Classes

In Cisco Nexus 3000 series switches and Cisco Nexus 3100 platform switches, packets are handled as cells. Each cell holds 208 bytes of data. One packet can be split into many cells, but each cell can contain a maximum of one packet. Priority groups are groups of cells on which the PFC thresholds are applied. They apply only to no-drop classes and are used for classifying traffic and monitoring statistics.

**Note** For Cisco Nexus 34180YC, the maximum amount of data in a cell is 80 bytes.

You can associate a no-drop class with a priority group number in the input queuing policy map to guarantee MTU buffers for the specified traffic class. The pause thresholds for the no-drop class are applied on the priority group.

By default, the priority group number is assigned by the system. You can override it by using the **priority-group** command.

**Note** You cannot have multiple no-drop classes mapped to the same priority group.

# Egress Queuing Policies

You can associate an egress policy map with an Ethernet interface to guarantee the bandwidth for the specified traffic class or to configure the egress queues.

Each Ethernet interface supports up to eight queues, one for each system class. The queues have the following default configuration:

- In addition to these queues, control traffic that is destined for the CPU uses strict priority queues. These queues are not accessible for user configuration.

**Note** For Cisco Nexus 34180YC, queue 7 is used for CPU traffic and is user configurable. However, as a best practice do not change queue 7 from strict priority to any other type of scheduling. Congesting queue 7 with data traffic leads to control packet loss.

- Standard Ethernet traffic in the default drop system class is assigned a queue. This queue uses WRR scheduling with 100 percent of the bandwidth.

If you add a system class, a queue is assigned to the class. You must reconfigure the bandwidth allocation on all affected interfaces. Bandwidth is not dedicated automatically to user-defined system classes.

You can configure one strict priority queue for Cisco Nexus 3000 series switches and Cisco Nexus 3500 platform switches. This queue is serviced before all other queues except the control traffic queue (which carries control rather than data traffic).

**Note** Cisco Nexus 34180YC supports a maximum of seven priority queues.

You can configure up to three strict priority queues with multiple priority levels on Cisco Nexus 3100 platform switches.

**Note** Cisco Nexus 34180YC supports a maximum of seven priority queues.

## QoS for Traffic Directed to the CPU

The device automatically applies QoS policies to traffic that is directed to the CPU to ensure that the CPU is not flooded with packets. Control traffic, such as bridge protocol data units (BPDU) frames, is given higher priority to ensure delivery.

# QoS Configuration Guidelines and Limitations

To maintain optimal switch performance, follow these guidelines when configuring system classes and policies:

- Beginning with Cisco Nexus NX-OS Release 7.0(3)I7(9), the following switches support DSCP wildcard mask:
  - N3K-C3048TP-1GE
  - N3K-C3064PQ-10GE
  - N3K-C3064PQ-10GX
  - N3K-C3064TQ-10GT

- N3K-C3132Q-40GX
- N3K-C3172PQ-10GE
- N3K-C3172PQ-XL
- N3K-C3172TQ-10GT
- N3K-C3172TQ-XL
- N3K-C3164Q-40GE

- Starting with Release 7.0(3)I2(1), the **show queuing interface ethernet slot/chassis_number** command displays two more options: module and summary in the output. These additional options do not have any functionality impact. See the following example of the output:

```
# show queuing interface eth1/1 ?
  <CR>
  ,         Multi range separator
  -         Range separator
  .         Sub interface separator
  >         Redirect it to a file
  >>        Redirect it to a file in append mode
  module    Slot/module
  summary   Summary
  |         Pipe command output to filter
```

- Prior to Release 7.0(3)I2(1), the switch prompt for **policy-map type network-qos** was switch(config-pmap-nq)#. Starting with Release 7.0(3)I2(1), the new switch prompt for **policy-map type network-qos** is switch(config-pmap-nqos)#. The change in the switch prompt is required to be taken care of in the scripts.

- Prior to Release 7.0(3)I2(1), if a non-existing class was configured under a policy-map, an error message would be displayed. Starting with Release 7.0(3)I2(1), if a non-existing class is configured under a policy-map, a new class-map is created and the prompt changes from config-pmap-nqos to config-cmap-nqos as displayed in the following example:

```
switch(config)# show class-map type network-qos

  Type network-qos class-maps
  ===========================
  class-map type network-qos pfcCos2
    match qos-group 2
  class-map type network-qos pfcCos3
    match qos-group 3
  class-map type network-qos pfcCos5
    match qos-group 5
  class-map type network-qos class-default
    match qos-group 0

switch(config)#
switch(config)# policy-map type network-qos pfcCos
switch(config-pmap-nqos)# class type network-qos pfcCos

switch(config-cmap-nqos)# show class-map type network-qos

  Type network-qos class-maps
  ===========================
  class-map type network-qos pfcCos

  class-map type network-qos pfcCos2
```

```
    match qos-group 2
  class-map type network-qos pfcCos3
    match qos-group 3
  class-map type network-qos pfcCos5
    match qos-group 5
  class-map type network-qos class-default
    match qos-group 0

switch(config-cmap-nqos)#
```

- Starting with Release 7.0(3)I2(1), the command **show queuing interface** displays the queues even without applying the network-qos policy.
- The output of the **show queuing interface** command displays an option for the internal HiGig2 interface. These interfaces are not relevant and the additional option does not have any functional impact.
- Starting with Release 7.0(3)I2(1), the output format and the fields of the **show policy-map interface <> type queuing** CLI command have been updated. For Class-map (queuing), the following fields are displayed: policy, bandwidth percent, queue dropped pkts, and queue depth in bytes.
- Prior to Release 7.0(3)I2(1), removing the default bandwidth configuration from the default queuing class used to set the bandwidth to 50%. Starting with Release 7.0(3)I2(1), removing the default bandwidth configuration sets the bandwidth to the default value of 100%. You can set the bandwidth to 50% by configuring **bandwidth percent 50**. You can use the **no bandwidth <bw-input>** CLI command to remove the bandwidth configuration.
- Prior to Release 7.0(3)I2(1), only priority level 1 was supported in the pmap configuration. Starting with Release 7.0(3)I2(1), you can configure priority levels 2 and 3 in the pmap configuration. Release 7.0(3)I2(1) supports only the functionality for priority level 1 for Cisco Nexus 3000 Series platforms even though both priority level 2 and 3 are allowed in the pmap configuration in the Cisco Nexus 3000 Series platforms. Cisco Nexus 3100 Series platforms support priority level 1, 2, 3 with functionality perspective.
- Starting with Release 7.0(3)I2(1), the output format and the fields of the **show queuing interface eth <>** CLI command have been updated. The output displays all qos-groups, control qos group, SPAN qos-group, and pfc statistics.The xon drops, xoff drops, and the HW MTU fields are not displayed in the new format.
- ECN is supported on the Cisco Nexus 3000 series switches and the Cisco Nexus 3132 switch. It is not supported on Cisco Nexus 3172 switches.
- WRED and ECN configuration is not supported on a class mapped to qos-group 1 for Cisco Nexus 3000 Series switches. However, WRED and ECN configuration is supported on a class mapped to qos-group 1 for the Cisco Nexus 3132 switch.
- Starting with Release 6.0(2)U5(1), the queue-limit CLI under the queuing policy is enhanced to support zero egress queue size to drop all packets on the queue. In scenarios where a particular type of traffic needs to be dropped without using the drop ACLs, the traffic can be mapped to a dedicated egress queue and then the queue-size 0 bytes can be applied using this enhancement. As a result, all UC and MC traffic mapped to the queue gets dropped completely.
- Starting with Release 6.0(2)U5(1), the switch allows different modes of min-buffer allocation that can reduce the overall min-buffers allocated per port, in turn increasing the shared pool size. If the shared pool size is higher, the burst absorption capability of the switch is better.

When configuring EtherChannels, note the following guidelines:

- The service policy configured on an EtherChannel applies to all member interfaces.

# Configuring System Classes

## Configuring Class Maps

You can create or modify a class map with the **class-map** command. The class map is a named object that represents a class of traffic. In the class map, you specify a set of match criteria for classifying the packets. You can then reference class maps in policy maps.

**Note** The class map type default is type qos and its match criteria default is match-all.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map** [**type** {**network-qos** \| **qos** \| **queuing**}] *class-map name* | Creates or accesses a named object that represents the specified class of traffic.<br>Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.<br>The three class-map configuration modes are as follows:<br>• **network-qos**—Network-wide (global) mode. CLI prompt: switch(config-cmap-nq)#<br>• **qos**—Classification mode; this is the default mode. CLI prompt: switch(config-cmap-qos)#<br>• **queuing**—Queuing mode. CLI prompt: switch(config-cmap-que)# |
| **Step 3** | (Optional) switch(config)# **class-map** [**type qos**] [**match-all** \| **match-any**] *class-map name* | Specifies that packets must match any or all criteria that is defined for a class map.<br>• **match-all**—Classifies traffic if packets match all criteria that is defined for a specified class map (for example, if both the defined CoS and the ACL criteria match). |

| | Command or Action | Purpose |
|---|---|---|
| | | • **match-any**—Classifies traffic if packets match any criteria that is defined for a specified class map (for example, if either the CoS or the ACL criteria matches).<br><br>Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 4** | (Optional) switch(config)# **no class-map** [**type** {**network-qos** \| **qos** \| **queuing**}] *class-name* | Deletes the specified class map.<br>**Note**<br>You cannot delete the system-defined class map: class-default.<br><br>Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |

## Configuring ACL Classification

You can classify traffic by matching packets based on an existing access control list (ACL). Traffic is classified by the criteria defined in the ACL. The **permit** and **deny** ACL keywords are ignored in the matching; even if a match criteria in the access-list has a **deny** action, it is still used for matching for this class.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type qos** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-qos)# **match access-group name** *acl-name* | Configures a traffic class by matching packets based on the *acl-name*. The **permit** and **deny** ACL keywords are ignored in the matching.<br>**Note**<br>You can only define a single ACL in a class map.<br>You cannot add any other match criteria to a class with a **match access-group** defined. |
| **Step 4** | (Optional) switch(config-cmap-qos)# **no match access-group name** *acl-name* | Removes the match from the traffic class. |

### Example

This example shows how to classify traffic by matching packets based on existing ACLs:

```
switch# configure terminal
switch(config)# class-map type qos class_acl
switch(config-cmap-qos)# match access-group name acl-01
```

Use the **show class-map** command to display the ACL class-map configuration:

```
switch# show class-map class_acl
```

## Configuring CoS Classification

You can classify traffic based on the class of service (CoS) in the IEEE 802.1Q header. This 3-bit field is defined in IEEE 802.1p to support QoS traffic classes. CoS is encoded in the high order 3 bits of the VLAN ID Tag field and is referred to as *user_priority*.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type qos** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-qos)# **match cos** *cos-value* | Specifies the CoS value to match for classifying packets into this class. You can configure a CoS value in the range of 0 to 7. |
| **Step 4** | (Optional) switch(config-cmap-qos)# **no match cos** *cos-value* | Removes the match from the traffic class. |

### Example

This example shows how to classify traffic by matching packets based on a defined CoS value:

```
switch# configure terminal
switch(config)# class-map type qos match-any class_cos
switch(config-cmap-qos)# match cos 4, 5-6
```

Use the **show class-map** command to display the CoS value class-map configuration:

```
switch# show class-map class_cos
```

# Configuring DSCP Classification

You can classify traffic based on the Differentiated Services Code Point (DSCP) value in the DiffServ field of the IP header (either IPv4 or IPv6).

***Table 1: Standard DSCP Values***

| Value | List of DSCP Values |
|---|---|
| af11 | AF11 dscp (001010)—decimal value 10 |
| af12 | AF12 dscp (001100)—decimal value 12 |
| af13 | AF13 dscp (001110)—decimal value 14 |
| af21 | AF21 dscp (010010)—decimal value 18 |
| af22 | AF22 dscp (010100)—decimal value 20 |
| af23 | AF23 dscp (010110)—decimal value 22 |
| af31 | AF31 dscp (011010)—decimal value 26 |
| af32 | AF32 dscp (011100)—decimal value 28 |
| af33 | AF33 dscp (011110)—decimal value 30 |
| af41 | AF41 dscp (100010)—decimal value 34 |
| af42 | AF42 dscp (100100)—decimal value 36 |
| af43 | AF43 dscp (100110)—decimal value 38 |
| cs1 | CS1 (precedence 1) dscp (001000)—decimal value 8 |
| cs2 | CS2 (precedence 2) dscp (010000)—decimal value 16 |
| cs3 | CS3 (precedence 3) dscp (011000)—decimal value 24 |
| cs4 | CS4 (precedence 4) dscp (100000)—decimal value 32 |
| cs5 | CS5 (precedence 5) dscp (101000)—decimal value 40 |
| cs6 | CS6 (precedence 6) dscp (110000)—decimal value 48 |
| cs7 | CS7 (precedence 7) dscp (111000)—decimal value 56 |
| default | Default dscp (000000)—decimal value 0 |
| ef | EF dscp (101110)—decimal value 46 |

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type qos** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-qos)# **match dscp** *dscp-list* | Configures the traffic class by matching packets based on the values in the *dscp-list* variable. For a list of DSCP values, see the Standard DSCP Values table. |
| **Step 4** | (Optional) switch(config-cmap-qos)# **no match dscp** *dscp-list* | Removes the match from the traffic class. For a list of DSCP values, see the Standard DSCP Values table. |

#### Example

This example shows how to classify traffic by matching packets based on the DSCP value in the DiffServ field of the IP header:

```
switch# configure terminal
switch(config)# class-map type qos match-any class_dscp
switch(config-cmap-qos)# match  dscp af21 af32
```

Use the **show class-map** command to display the DSCP class-map configuration:

```
switch# show class-map class_dscp
```

## Configuring IP RTP Classification

The IP Real-time Transport Protocol (RTP) is a transport protocol for real-time applications that transmits data such as audio or video and is defined by RFC 3550. Although RTP does not use a common TCP or UDP port, you typically configure RTP to use ports 16384 to 32767. UDP communications use an even port and the next higher odd port is used for RTP Control Protocol (RTCP) communications.

You can classify based on UDP port ranges, which are likely to target applications using RTP.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type qos** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, |

| | Command or Action | Purpose |
|---|---|---|
| | | are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-qos)# **match ip rtp** *port-number* | Configures the traffic class by matching packets based on a range of lower and upper UDP port numbers, which is likely to target applications using RTP. Values can range from 2000 to 65535. |
| **Step 4** | (Optional) switch(config-cmap-qos)# **no match ip rtp** *port-number* | Removes the match from the traffic class. |

### Example

The following example shows how to classify traffic by matching packets based on UDP port ranges that are typically used by RTP applications:

```
switch# configure terminal
switch(config)# class-map type qos match-any class_rtp
switch(config-cmap-qos)# match  ip rtp 2000-2100, 4000-4100
```

Use the **show class-map** command to display the RTP class-map configuration:

```
switch# show class-map class_rtp
```

## Configuring Precedence Classification

You can classify traffic based on the precedence value in the type of service (ToS) byte field of the IP header (either IPv4 or IPv6). The following table shows the precedence values:

***Table 2: Precedence Values***

| Value | List of Precedence Values |
|---|---|
| <0-7> | IP precedence value |
| critical | Critical precedence (5) |
| flash | Flash precedence (3) |
| flash-override | Flash override precedence (4) |
| immediate | Immediate precedence (2) |
| internet | Internetwork control precedence (6) |
| network | Network control precedence (7) |
| priority | Priority precedence (1) |
| routine | Routine precedence (0) |

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type qos match-any** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-qos)#**match precedence** *precedence-values* | Configures the traffic class by matching packets based on precedence values. For a list of precedence values, see the Precedence Values table. |
| **Step 4** | (Optional) switch((config-cmap-qos)# **no match precedence** *precedence-values* | Removes the match from the traffic class. For a list of precedence values, see the Precedence Values table. |

#### Example

This example shows how to classify traffic by matching packets based on the precedence value in the ToS byte field of the IP header:

```
switch# configure terminal
switch(config)# class-map type qos match-any class_precedence
switch(config)# class-map type qos c1
switch(config-cmap-qos)# match precedence 1-2, 5
```

Use the **show class-map** command to display the IP precedence value class-map configuration:

```
switch# show class-map class_precedence
```

# Creating Policy Maps

The **policy-map** command is used to create a named object that represents a set of policies that are to be applied to a set of traffic classes.

The device provides one default system class: a drop class for best-effort service (class-default). You can define up to four additional system classes for Ethernet traffic.

The following predefined policy maps are used as default service policies:

- network-qos: default-nq-policy
- Input qos: default-in-policy
- Output queuing: default-out-policy
- Input queuing: default-in-policy

You need to create a policy map to specify the policies for any user-defined class. In the policy map, you can configure the QoS parameters for each class. You can use the same policy map to modify the configuration of the default classes.

The device distributes all the policy-map configuration values to the attached network adapters.

### Before you begin

Before creating the policy map, define a class map for each new system class.

## Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **policy-map** [**type** {**network-qos** \| **qos** \| **queuing**}] *policy-name* | Creates a named object representing a set of policies that are to be applied to a set of traffic classes. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.<br>The three policy-map configuration modes are as follows:<br>• network-qos—Network-wide (global) mode. CLI prompt: switch(config-pmap-nq)#<br>• qos—Classification mode; this is the default mode. CLI prompt: switch(config-pmap-qos)#<br>• queuing—Queuing mode. CLI prompt: switch(config-pmap-que)# |
| **Step 3** | (Optional) switch(config)# **no policy-map** [**type** {**network-qos** \| **qos** \| **queuing**}] *policy-name* | Deletes the specified policy map. |
| **Step 4** | switch(config-pmap)# **class** [**type** {**network-qos** \| **qos** \| **queuing**}] *class-name* | Associates a class map with the policy map, and enters configuration mode for the specified system class. The three class-map configuration modes are as follows:<br>• network-qos—Network-wide (global) mode. CLI prompt: switch(config-pmap-c-nq)#<br>• qos—Classification mode; this is the default mode. CLI prompt: switch(config-pmap-c-qos)# |

| | Command or Action | Purpose |
|---|---|---|
| | | • queuing—Queuing mode. CLI prompt: switch(config-pmap-c-que)#<br><br>**Note**<br>The associated class map must be the same type as the policy-map type. |
| **Step 5** | (Optional) switch(config-pmap)# **no class** [**type** {**network-qos** \| **qos** \| **queuing**}] *class-name* | Deletes the class map association. |

## Configuring Type QoS Policies

Type qos policies are used for classifying the traffic of a specific system class identified by a unique qos-group value. A type qos policy can be attached to the system or to individual interfaces for ingress traffic only.

You can set a maximum of eight QoS groups for ingress traffic.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **policy-map type qos** *policy-name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. 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-qos)# [**class** \| **class-default**] **type qos** *class-name* | Associates a class map with the policy map, and enters configuration mode for the specified system class.<br><br>**Note**<br>The associated class map must be the same type as the policy map type. |
| **Step 4** | switch(config-pmap-c-qos)# **set qos-group** *qos-group-value* | Configures one or more **qos-group** values to match on for classification of traffic into this class map. The list below identifies the ranges of the *qos-group-value* . There is no default value.<br><br>**Note**<br>The Cisco 3000 Series Switches supports 8 QoS groups(0-7).<br><br>The Cisco 3500 Series Switches supports 4 QoS groups(1-4). |

| | Command or Action | Purpose |
|---|---|---|
| **Step 5** | (Optional) switch(config-pmap-c-qos)# **no set qos-group** *qos-group-value* | Removes the **qos-group** values from this class. |

### Example

This example shows how to define a type qos policy map:

```
switch# configure terminal
switch(config)# policy-map type qos policy-s1
switch(config-pmap-qos)# class type qos class-s1
switch(config-pmap-c-qos)# set qos-group 2
```

## Configuring Type Network QoS Policies

Type network qos policies can only be configured on the system qos attachment point. They are applied to the entire switch for a particular class.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **policy-map type network-qos** *policy-name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. 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-nq)# **class type network-qos** *class-name* | Associates a class map with the policy map, and enters configuration mode for the specified system class.<br>**Note**<br>The associated class map must be the same type as the policy map type. |
| **Step 4** | switch(config-pmap-c-nq)# **mtu** *mtu-value* | Specifies the MTU value in bytes.<br>**Note**<br>The *mtu-value* that you configure must be less than the value set by the **system jumbomtu** command. |
| **Step 5** | (Optional) switch(config-pmap-c-nq)# **no mtu** | Resets the MTU value in this class. |
| **Step 6** | switch(config-pmap-c-nq)# **pause no-drop** | Configures a no-drop class.<br>**Note** |

| | Command or Action | Purpose |
|---|---|---|
| | | This option works only on class-map matching qos-group 1 in Cisco Nexus 3500 series switches. |
| **Step 7** | switch(config-pmap-c-nq)# **set cos** *cos-value* | Specifies a 802.1Q CoS value which is used to mark packets on this interface. The value range is from 0 to 7. |
| **Step 8** | (Optional) switch(config-pmap-c-nq)# **no set cos** *cos-value* | Disables the marking operation in this class. |

### Example

This example shows how to define a type network-qos policy map:

```
switch# configure terminal
switch(config)# policy-map type network-qos policy-que2
switch(config-pmap-nq)# class type network-qos class-que2
switch(config-pmap-c-nq)# mtu 5000
switch(config-pmap-c-nq)# pause no-drop
switch(config-pmap-c-nq)# congestion-control random-detect
switch(config-pmap-c-nq)# set cos 4
```

## Configuring Type Queuing Policies

Type queuing policies for egress are used for scheduling and buffering the traffic of a specific system class. A type queuing policy is identified by its QoS group and can be attached to the system or to individual interfaces for input or output traffic.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **policy-map type queuing** *policy-name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. 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-name* | Associates a class map with the policy map, and enters configuration mode for the specified system class. |
| **Step 4** | switch(config-pmap-c-que)# **priority** | Specifies that traffic in this class is mapped to a strict priority queue.<br>**Note** |

| | Command or Action | Purpose |
|---|---|---|
| | | In Cisco Nexus 3000 series switches and Cisco Nexus 3500 platform switches, only one class in each policy map can have strict priority set on it. Cisco Nexus 3100 platform switches allow multiple priority levels. A maximum of three classes can have priority levels configured. However, only one class can be configured with a particular level. (The Cisco Nexus 34180YC switch supports a maximum of seven priority queues.) |
| **Step 5** | (Optional) switch(config-pmap-c-que)# **no priority** | Removes the strict priority queuing from the traffic in this class. |
| **Step 6** | switch(config-pmap-c-que)# **shape** {**kbps** \| **mbps** \| **gbps**} *burst size* **min** *minimum bandwidth* | Specifies the burst size and minimum guaranteed bandwidth for this queue.<br>**Note**<br>This command is not suported on Cisco Nexus 3500 Series switches.<br>**Note**<br>For the Cisco Nexus 34180YC switch, the command syntax is switch(config-pmap-c-que)# **shape** {**kbps** \| **mbps** \| **gbps**} **min** *minimum bandwidth* **max** *maximum bandwidth*. |
| **Step 7** | switch(config-pmap-c-que)# **bandwidth percent** *percentage* | Assigns a weight to the class. The class will receive the assigned percentage of interface bandwidth if there are no strict-priority queues. If there are strict-priority queues, however, the strict-priority queues receive their share of the bandwidth first. The remaining bandwidth is shared in a weighted manner among the class configured with a bandwidth percent. For example, if strict-priority queues take 90 percent of the bandwidth, and you configure 75 percent for a class, the class will receive 75 percent of the remaining 10 percent of the bandwidth.<br>**Note**<br>Before you can successfully allocate bandwidth to the class, you must first reduce the default bandwidth configuration on class-default and class-fcoe. |
| **Step 8** | (Optional) switch(config-pmap-c-que)# **no bandwidth percent** *percentage* | Removes the bandwidth specification from this class. |

| | Command or Action | Purpose |
|---|---|---|
| **Step 9** | (Optional) switch(config-pmap-c-que)# **priority level** *level* | Specifies the strict priority levels for the Cisco Nexus 3100 platform switches. These levels can be 1, 2, or 3.<br>**Note**<br>The Cisco Nexus 34180YC switch supports a maximum of seven priority queues. |
| **Step 10** | (Optional) switch(config-pmap-c-que)# **queue-limit** *queue size* [**dynamic** *dynamic threshold*] | Specifies either the static or dynamic shared limit available to the queue for Cisco Nexus 3100 Series switches. The static queue limit defines the fixed size to which the queue can grow.<br>The dynamic queue limit allows the queue's threshold size to be decided depending on the number of free cells available, in terms of the alpha value.<br>**Note**<br>The queue-limit CLI under queuing policy is enhanced to support 0 queue size. Configuring queue-limit as 0 bytes drops all the packets on the queue.<br>**Note**<br>The queue-limit range for the Cisco Nexus 34180YC switch is from 64 to 12582912 bytes. |

### Example

This example shows how to define a type queuing policy map for Cisco Nexus 3000 series switches:

```
switch# configure terminal
switch(config)# policy-map type queuing policy-queue1
switch(config-pmap-que)# class type queuing class-queue1
switch(config-pmap-c-que)# priority
switch(config-pmap-c-que)# shape kbps 30000000 min 18000000
switch(config-pmap-c-que)# bandwidth percent 20
```

This example shows how to define a type queuing policy map for Cisco Nexus 3100 platform switches:

```
switch# configure terminal
switch(config)# policy-map type queuing p1
switch(config-pmap-que)# class type queuing q3
switch(config-pmap-c-que)# priority level 2
switch(config-pmap-c-que)# shape kbps 30000000 min 18000000
switch(config-pmap-c-que)# class type queuing q2
switch(config-pmap-c-que)# priority level 3
switch(config-pmap-c-que)# class type queuing q1
switch(config-pmap-c-que)# bandwidth percent 30
switch(config-pmap-c-que)# class type queuing q4
```

```
switch(config-pmap-c-que)# bandwidth percent 10
switch(config-pmap-c-que)# class type queuing q5
switch(config-pmap-c-que)# bandwidth percent 10
switch(config-pmap-c-que)# class type queuing q6
switch(config-pmap-c-que)# priority level 1
switch(config-pmap-c-que)# class type queuing q7
switch(config-pmap-c-que)# bandwidth percent 10
switch(config-pmap-c-que)# class type queuing class-default
switch(config-pmap-c-que)# queue-limit dynamic 4
switch(config-pmap-c-que)# bandwidth percent 0
```

This example shows how to define a type queuing policy map for Cisco Nexus 3500 platform switches:

```
switch# configure terminal
switch(config)# policy-map type queuing p1
switch(config-pmap-que)# class type queuing q1
switch(config-pmap-c-que)# bandwidth percent 10
switch(config-pmap-c-que)# priority
switch(config-pmap-c-que)# class type queuing q2
switch(config-pmap-c-que)# bandwidth percent 20
switch(config-pmap-c-que)# class type queuing q3
switch(config-pmap-c-que)# bandwidth percent 30
switch(config-pmap-c-que)# class type queuing q4
switch(config-pmap-c-que)# bandwidth percent 40
switch(config-pmap-c-que)# class type queuing class-default
switch(config-pmap-c-que)# bandwidth percent 10
```

## Configuring an ECN Threshold

You can configure an explicit congestion notification (ECN) threshold per class in a queuing policy and apply it to an interface.

Prior to Cisco NX-OS Release 5.0(3)U4(1), WRED and an ECN can only be enabled or disabled on QoS class in the networkqos policy (with static thresholds). Starting with Cisco NX-OS Release 5.0(3)U4(1), an enhanced ECN marking is supported as follows:

**Note** This feature is not supported on Cisco Nexus 3500 platform switches.

- WRED and ECN thresholds can be configured corresponding to a class from the queueing policy by using the following Steps 1 through 8.

**Note** A WRED and ECN still need to be enabled by the network-qos policy class configuration mode. Not applicable for Cisco Nexus 34180YC switches.

- WRED and ECN can be enabled on a global basis outside the MQC command line. You can configure WRED and an ECN at a global buffer level where you enable WRED and an ECN and specify a threshold at the system level by using the following Steps 1 through 9. If this threshold is exceeded, WRED and ECN are applied on all WRED/ECN enabled classes in the system.
- By default, when WRED and an ECN are enabled, the marking or drop happens based on the class or queue threshold. However, when the global based WRED and ECN is also enabled, by using the **congestion-control random-detect global-buffer** and **wred-queue qos-group-map queue-only**

commands, the WRED and ECN marking behavior initiates when either of the class thresholds or global threshold is exceeded.

### Before you begin

Ensure that you have enabled ECN or WRED on the QoS group that you want in the network-qos policy.

## Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type queuing** *class-map name* | Creates or accesses a named object that represents the specified class of traffic in queuing mode. Class map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-que)# **match qos-group** *qos-group-number* | Associates a QoS group to the queuing class map. |
| **Step 4** | switch(config-cmap-que)# **exit** | Exits class mode. |
| **Step 5** | switch(config)# **policy-map type queuing** *policy-map name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes in queuing mode. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 6** | switch(config-pmap-que)# **class type queuing** *class-map name* | Associates a queuing class map with the policy map, and enters configuration mode for the specified system class. |
| **Step 7** | switch(config-pmap-c-que)# **random-detect minimum-threshold** {*min-threshold* [**bytes** \| **kilobytes** \| **megabytes** \| **packets** } **maximum-threshold** {*max-threshold* [**bytes** \| **kilobytes** \| **megabytes** \| **packets**} **drop-probability** *drop probability* **weight** *weight* **cap-average** | Configures WRED or ECN based on whether the network-qos policy has ECN or WRED enabled. You can specify minimum and maximum thresholds used to drop (WRED) or mark (ECN) packets from the queue. You can configure thresholds by the number of packets, number of bytes, and packets where 1 packet maps to 208 bytes. 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.<br>You can also specify the following optional parameters: |

| | Command or Action | Purpose |
|---|---|---|
| | | • drop probability—This option specifies the drop probability when the average queue size is between the minimum queue length and maximum queue size. When the queue size is greater than the maximum queue size, 100 percent of frames are dropped. The default drop probability value is 100.<br>• weight—This option is used to derive the actual queue size from the current queue size. For example, a weight of 0 means that the actual queue size is equal to the current queue size. The default weight value is 0.<br>• cap-average—This option is used to replace the average queue size with the current queue size if the average queue size is greater than the current queue size.<br>**Note**<br>N3K-C34180YC does not support cap-average. |
| **Step 8** | switch(config-cmap-que)# **exit** | Exits policy mode. |
| **Step 9** | switch(config)# **congestion-control random-detect global-buffer minimum-threshold** {*min-threshold* [**bytes** \| **kilobytes** \| **megabytes** \| **packets**]} **maximum-threshold** {*max-threshold* [**bytes** \| **kilobytes** \| **megabytes packets**} | Configures the global ECN threshold. You can specify minimum and maximum global buffer thresholds. If the global buffer exceeds these thresholds, the packets in individual queues, which are ECN or WRED enabled, are marked or dropped even if the queue thresholds have not been exceeded. You can configure thresholds by the number of packets, number of bytes, and packets where 1 packet maps to 208 bytes. 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.<br>**Note**<br>The Cisco Nexus 34180YC switch does not support **congestion-control random-detect global-buffer minimum-threshold**. |
| **Step 10** | (Optional) switch(config-pmap-nq)# **wred-queue qos-group-map queue-only** *queue-group* | Enables ECN marking for the specified QoS group that is based only on a class threshold |

| | Command or Action | Purpose |
|---|---|---|
| | | and independent of the global buffer threshold configuration. |
| **Step 11** | (Optional) switch(config-pmap-nq)# **show wred-queue qos-group-map** | Displays the configuration for the queue-only QoS group maps. |
| **Step 12** | (Optional) switch# **copy running-config startup-config** | Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration. |

### Example

The following example shows how to configure an ECN threshold per class:

```
switch# configure terminal
switch(config)# class-map type queuing cque_ecn
switch(config-cmap-que)# match qos-group 1
switch(config-cmap-que)# exit
switch(config)# policy-map type queuing pque_ecn
switch(config-pmap-que)# class type queuing cque_ecn
switch(config-pmap-c-que)# random-detect minimum-threshold 20 kilobytes
maximum-threshold 60 kilobytes drop-probability 70 weight 11 cap-average
```

The following example shows how to configure a global ECN threshold in packets, bytes, kilobytes, and megabytes:

```
switch(config)# congestion-control random-detect global-buffer
minimum-threshold 1000 bytes maximum-threshold 1000 bytes
switch(config)#
```

# Configuring Pause Buffer Thresholds and Priority Groups

The ingress queuing policy is used to configure pause buffer thresholds and priority groups. This feature is not supported on Cisco Nexus 3500 Series switches.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **policy-map type queuing** *policy-map name* | Enters policy-map queuing class mode and identifies the policy map assigned to the type queuing policy map. |
| **Step 3** | switch(config-pmap-que)# **class type queuing** *class-map name* | References an existing queuing class map in a policy map and enters class mode.<br>**Important**<br>This must be defined as a no-drop class in the network-qos policy applied in the system qos. |

| | Command or Action | Purpose |
|---|---|---|
| **Step 4** | switch(config-pmap-c-que)# **pause buffer-size** *buffer-size* **pause-threshold** *xoff-size* **resume-threshold** *xon-size* | Specifies the buffer threshold settings for pause and resume.<br>**Note**<br>The values mentioned next to each parameter are the recommended values. These values are assigned by the system and will be the default values if no value is defined in the input queuing policy. You can use different values based on your requirement.<br>On 10-Gigabit ports, if the no-drop class has an MTU <= 2240:<br>• buffer-size—27,456 bytes<br>• pause threshold—12,480 bytes<br>• resume-threshold—0 bytes<br>On 10-Gigabit ports, if the no-drop class has an MTU > 2240:<br>• buffer-size—89,440 bytes<br>• pause threshold—34,320 bytes<br>• resume-threshold—21,840 bytes<br>On 40-Gigabit ports, if the no-drop class has an MTU of MTU <=2240:<br>• buffer-size—83,616 bytes<br>• pause threshold—40,352 bytes<br>• resume-threshold—19,552 bytes<br>On 40-Gigabit ports, if the no-drop class has an MTU > 2240:<br>• buffer-size—1,58,080 bytes<br>• pause threshold— 77,376 bytes<br>• resume-threshold—56,576 bytes |
| **Step 5** | switch(config-pmap-c-que)# **no pause buffer-size** *buffer-size* **pause-threshold** *xoff-size* **resume-threshold** *xon-size* | Removes the buffer threshold settings for pause and resume. |
| **Step 6** | switch(config-pmap-c-que)# **pause priority-group** *priority group number* | Maps the no-drop class traffic to the priority group specified. |
| **Step 7** | (Optional) switch(config-pmap-c-que)# **queue-limit** *queue size* [**dynamic** *dynamic threshold*] | Specifies either the static or dynamic shared limit available to the queue for Cisco Nexus 3100 Series switches. The static queue limit defines the fixed size to which the queue can grow. The dynamic queue limit allows the queue's threshold size to be decided depending on the number of free cells available, in terms of the alpha value. |

### Example

This example shows how to configure priority groups for no-drop classes:

```
switch# configure terminal
switch(config-pmap-que)# policy-map type queuing p1
switch(config-pmap-que)# class type queuing c1
switch(config-pmap-c-que)# pause buffer-size 39936 pause-threshold 24960 resume-threshold
12480
switch(config-pmap-c-que)# pause priority-group 1
```

# Information About Marking

Marking is a method that you use to modify the QoS fields of the incoming and outgoing packets.

You can use marking commands in traffic classes that are referenced in a policy map. The marking features that you can configure are listed below:

- DSCP
- IP precedence
- CoS

# Configuring CoS Marking

The value of the CoS field is recorded in the high-order three bits of the VLAN ID Tag field in the IEEE 802.1Q header.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config) # **policy-map** [**type network-qos**] *policy-map name* | Creates or accesses the policy map named *policy-map-name* and enters policy-map mode.<br>The policy-map name can contain alphabetic, hyphen, or underscore characters, is case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-pmap-nq) # **class** [**type network-qos**] {*class-map name* \|**class-default**} | Creates a reference to the *class-map-name* and enters policy-map class configuration mode.<br>Use the **class-default** keyword to select all traffic that is not currently matched by classes in the policy map. |
| **Step 4** | switch(config-pmap-c-nq) # **set cos** *cos-value* | Specifies the CoS value to cos-value.<br>The *cos-value* can range from 0 to 7.<br>**Note** |

| | Command or Action | Purpose |
|---|---|---|
| | | This command is supported only for egress policies. |

# Configuring a DSCP Wildcard Mask

Use the DSCP wildcard mask feature to classify multiple DSCP values from a set of IP flows recognized by an ACL and the DSCP value. Classification of IP information and DSCP values occurs in a more granular way by using multiple parameters. With this granularity, you can treat these flows by policing them to protest the rest of the traffic, or assign them to a qos-group for further QoS operations.

**Note** Only the following Cisco Nexus 3000 series switches support the DSCP wildcard mask feature:

- Cisco Nexus 3048TP-1GE
- Cisco Nexus 3064PQ-10GE
- Cisco Nexus 3064PQ-10GX
- Cisco Nexus 3064TQ-10GT
- Cisco Nexus 3132Q-40GX
- Cisco Nexus 3172PQ-10GE
- Cisco Nexus 3172PQ-XL
- Cisco Nexus 3172TQ-10GT
- Cisco Nexus 3172TQ-XL
- Cisco Nexus 3164Q-40GE

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal**<br>**Example:**<br>`switch# configure terminal`<br>`switch(config)#` | Enters global configuration mode. |
| **Step 2** | **ip access-list** *acl-name*<br>**Example:**<br>`switch(config)# ip access-list acl-01`<br>`switch(config-acl)` | Enters the ACL configuration mode and creates an ACL with the entered name. |

| | Command or Action | Purpose |
|---|---|---|
| **Step 3** | [ *sequence-number* ] { **permit** \| **deny** } *protocol* { *source-ip-prefix* \| *source-ip-mask* } { *destination-ip-prefix* \| *destination-ip-mask* } [ **dscp** *dscp-value* **dscp-mask** *dscp-mask-value* ]<br>**Example:**<br>`switch(config-acl)# 10 permit ip 10.1.1.1/24 20.1.1.2/24 dscp 33 dscp-mask 30` | Creates an ACL entry that matches or filters traffic that is based on a DSCP wildcard bit mask.<br>The *sequence-number* argument can be a whole number from 1 through 4294967295.<br>**dscp** *dscp-value*: Match packets with a specific DSCP value.<br>**dscp-mask** *dscp-mask-value*: Configures the DSCP wildcard mask which matches on any bit in the DSCP value to filter traffic. Range is from 0 to 0x3F. |
| **Step 4** | **exit**<br>**Example:**<br>`switch(config-acl)# exit`<br>`switch(config)#` | Exits ACL configuration mode and enters global configuration mode. |
| **Step 5** | **class-map** [**type qos**] [**match-any** \| **match-all**] *class-name*<br>**Example:**<br>`switch(config)# class-map type qos match-any class_dscp_mask`<br>`switch(config-cmap-qos)#` | Creates or accesses the class map that is named by the *class-name* variable and enters the class-map mode. The class-map name can contain alphabetic, hyphen, or underscore characters, and can be up to 40 characters. |
| **Step 6** | **match access-list** *acl-name*<br>**Example:**<br>`switch(config-cmap-qos)# match access-list acl-01`<br>`switch(config-cmap-qos)#` | Configures the traffic class by matching packets that are based on the IP access list. |

### Example

In the following example, an ACL looks at traffic that is sent from subnet 10.1.1.0 to subnet 20.1.1.0. The ACL also checks for traffic with DSCP 33, and any subsequent DSCP values from 33 through 63, with a mask value of 30. The ACL is set to a class map that is matching this ACL for further QoS operations.

```
switch# configure terminal
switch(config)# ip access-list acl-01
switch(config-acl)# 10 permit ip 10.1.1.1/24 20.1.1.2/24 dscp 33 dscp-mask 30
switch(config-acl)# exit
switch(config)# class-map type qos match-any class_dscp_mask
switch(config-cmap-qos)# match access-list acl-01
```

# Configuring DSCP Marking

You can set the DSCP value in the six most significant bits of the DiffServ field of the IP header to a specified value. You can enter numeric values from 0 to 60, in addition to the standard DSCP values shown in the table below:

**Note** For Cisco Nexus 34180YC switches, the numeric values range from 0 to 63

**Note** You can set DSCP or IP precedence but you cannot set both values because they modify the same field in the IP packet.

***Table 3: Standard DSCP Values***

| Value | List of DSCP Values |
|---|---|
| af11 | AF11 dscp (001010)—decimal value 10 |
| af12 | AF12 dscp (001100)—decimal value 12 |
| af13 | AF13 dscp (001110)—decimal value 14 |
| af21 | AF21 dscp (010010)—decimal value 18 |
| af22 | AF22 dscp (010100)—decimal value 20 |
| af23 | AF23 dscp (010110)—decimal value 22 |
| af31 | AF31 dscp (011010)—decimal value 26 |
| af32 | AF40 dscp (011100)—decimal value 28 |
| af33 | AF33 dscp (011110)—decimal value 30 |
| af41 | AF41 dscp (100010)—decimal value 34 |
| af42 | AF42 dscp (100100)—decimal value 36 |
| af43 | AF43 dscp (100110)—decimal value 38 |
| cs1 | CS1 (precedence 1) dscp (001000)—decimal value 8 |
| cs2 | CS2 (precedence 2) dscp (010000)—decimal value 16 |
| cs3 | CS3 (precedence 3) dscp (011000)—decimal value 24 |
| cs4 | CS4 (precedence 4) dscp (100000)—decimal value 32 |
| cs5 | CS5 (precedence 5) dscp (101000)—decimal value 40 |
| cs6 | CS6 (precedence 6) dscp (110000)—decimal value 48 |

| Value | List of DSCP Values |
|---|---|
| cs7 | CS7 (precedence 7) dscp (111000)—decimal value 56 |
| default | Default dscp (000000)—decimal value 0 |
| ef | EF dscp (101110)—decimal value 46 |

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal** | Enters global configuration mode. |
| **Step 2** | **policy-map type qos** *qos-policy-map-name* | Creates or accesses the policy map named qos-policy-map-name, and then enters policy-map mode. The policy-map name can contain alphabetic, hyphen, or underscore characters, is case sensitive, and can be up to 40 characters. |
| **Step 3** | **class** [**type qos**] {*class-map-name* \| **class-default**} | Creates a reference to class-map-name, and enters policy-map class configuration mode. Use the **class-default** keyword to select all traffic that is not currently matched by classes in the policy map. |
| **Step 4** | **set dscp** *dscp-value* | Sets the DSCP value to dscp-value. See the Standards DSCP Values table. |

### Example

This example shows how to display the policy-map configuration as shown below:

```
switch# show policy-map policy1
```

# Configuring IP Precedence Marking

You can set the value of the IP precedence field in bits 0 to 2 of the IPv4 type of service (ToS) field or the equivalent Traffic Class field for IPv6 of the IP header. The following table shows the precedence values:

**Note** You can set IP precedence or DSCP but you cannot set both values because they modify the same field in the IP packet.

*Table 4: Precedence Values*

| Value | List of Precedence Values |
|---|---|
| *0-7* | IP precedence value |
| critical | Critical precedence (5) |
| flash | Flash precedence (3) |
| flash-override | Flash override precedence (4) |
| immediate | Immediate precedence (2) |
| internet | Internetwork control precedence (6) |
| network | Network control precedence (7) |
| priority | Priority precedence (1) |
| routine | Routine precedence (0) |

## Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **config terminal** | Enters global configuration mode. |
| **Step 2** | switch(config) # **policy-map** [**type qos**] *qos-policy-map-name* | Creates or accesses the policy map named *policy-map-name*, and then enters policy-map mode. The policy-map name can contain alphabetic, hyphen, or underscore characters, is case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-pmap-nq) # **class** [**type qos**] {*class-map-name* \| **class-default**} | Creates a reference to class-map-name, and enters policy-map class configuration mode. Use the **class-default** keyword to select all traffic that is not currently matched by classes in the policy map. |
| **Step 4** | switch(config-pmap-c-nq) # **set precedence** *precedence-value* | Sets the IP precedence value to precedence-value. You can enter one of the values shown in the Precedence Values table. |

### Example

This example shows how to set the precedence marking to 5:

```
switch(config)# policy-map type qos my_policy
switch(config-pmap-qos)# class type qos my_class
switch(config-pmap-c-qos)# set precedence 5
switch(config-pmap-c-qos)#
```

# QoS Configurations for Layer 3 Routing

## Configuring Layer 3 Multicast Queuing

You can use this procedure to distribute traffic into different queues, where each queue is configured with different weighted round robin (WRR) parameters.

### Procedure

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

### Example

This example shows how to configure a Layer 3 interface:

## Configuring a Service Policy for a Layer 3 Interface

You can configure a service policy for a Layer 3 interface.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **interface ethernet** *slot*/*port* | Enters the configuration mode for the specified interface. |
| **Step 3** | switch(config-if)# **no switchport** | Selects the Layer 3 interface. |
| **Step 4** | switch(config-if)# **service-policy** [**type** {**qos** **input** \| **queuing** [**input** \| **output**]}*policy-name* | Specifies the policy map to use as the service policy for the Layer 3 interface. There are two policy-map configuration modes:<br>• qos—Classification mode; this is the default mode.<br>• queuing—Queuing mode.<br>**Note**<br>The **output** keyword specifies that this policy map should be applied to traffic transmitted from an interface and the **input** keyword specifies that this policy map should be applied to traffic transmitted to an interface. You can apply both **output** and **input** to a queuing policy. The policy applied at input should have buffer configurations and the policy applied at |

| | Command or Action | Purpose |
|---|---|---|
| | | output should have scheduling and queuing configurations. |

### Example

The following example shows how to attach a queuing policy map to a Layer 3 interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/5
switch(config-if)# no switchport
switch(config-if)# service-policy type queuing output my_output_q_policy
switch(config-if)#
```

The following example shows how to attach an input qos policy map to a Layer 3 interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/5
switch(config-if)# no switchport
switch(config-if)# service-policy type qos input my_input_qos_policy
switch(config-if)#
```

## Changing the Bandwidth Allocated to Unicast and Multicast Traffic

You can change the bandwidth allocated to unicast and multicast traffic by assigning weighted round-robin (WRR) weights as a percentage of the interface data rate to the egress queues. The **wrr unicast-bandwidth** command configures this bandwidth percentage globally even if you use it on a specific interface.

**Note** The Cisco Nexus 34180YC switch does not supports the **wrr unicast-bandwidth** command.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **interface ethernet** *slot*/*port* | Enters configuration mode for the specified interface. |
| **Step 3** | switch(config-if)# **wrr unicast-bandwidth** percentage-value | Changes the bandwidth allocated to unicast and multicast traffic on traffic congestion globally. The bandwidth-value percentage ranges from 0 to 100 percent. |

### Example

This example shows how to attach a queuing policy map to a Layer 3 interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/5
switch(config-if)# wrr unicast-bandwidth 75
switch(config-if)#
```

# 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 input** \|**queuing** [**input** \| **output**]} *policy-name* | Specifies the policy map to use as the service policy for the system. There are three policy-map configuration modes:<br>• network-qos—Network-wide (system qos) mode.<br>• qos—Classification mode (system qos input or interface input only).<br>• queuing—Queuing mode (output at system qos and interface).<br>**Note**<br>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 only apply **output** to a queuing policy. |

### Example

# 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** *policy-map name* | 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** *policy-map name* | Resets the network-wide policy map. |
| **Step 5** | switch(config-sys-qos)# **no service-policy type queuing [input \| output]***policy-map name* | Resets the output queuing mode policy map. |

### Example

# 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
```

**Note** The Cisco Nexus device supports 1 MTU for all classes for all ports.

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-nqos)# class type network-qos class-fcoe
switch(config-pmap-nqos-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-nqos)# class type network-qos class-default
switch(config-pmap-c-nqos)# 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. Use the **show queueing interface ethernet** *slot/chassis_number* command to display all qos-groups, control qos group, span qos-group, and pfc statistics.

Use the **show policy-map system type network-qos** command to display the configured HW MTU for the qos-groups.

```
switch(config)# show policy-map system type network-qos

  Type network-qos policy-maps
  ============================
  policy-map type network-qos pn_system
    class type network-qos cn_1
      match qos-group 1
      mtu 9216
    class type network-qos cn_2
      match qos-group 2
      mtu 9216
    class type network-qos cn_3
      match qos-group 3
      mtu 9216
    class type network-qos cn_4
      match qos-group 4
      mtu 9216
    class type network-qos cn_5
      match qos-group 5
      mtu 9216
    class type network-qos cn_6
      match qos-group 6
      mtu 9216
    class type network-qos cn_7
      match qos-group 7
      mtu 9216
    class type network-qos class-default
      match qos-group 0
      mtu 1500

switch(config)#
```

This example shows queuing interface for the Cisco Nexus 34180YC switch:

```
switch(config)# show queuing interface spq_shape_wred_8q
policy-map type queuing msft_spq_8q
    class type queuing c-out-8q-q5
      priority level 3
    class type queuing c-out-8q-q4
      priority level 4
    class type queuing c-out-8q-q7
      priority level 1
    class type queuing c-out-8q-q6
      priority level 2
    class type queuing c-out-8q-q1
      bandwidth remaining percent 60
```

```
     random-detect minimum-threshold 120 kbytes maximum-threshold 120 kbytes drop-probability
 1 weight 0 ecn
    class type queuing c-out-8q-q3
      priority level 5
      shape min 2500 mbps max 2500 mbps
    class type queuing c-out-8q-q2
      priority level 6
      shape min 5000 mbps max 5000 mbps
    class type queuing c-out-8q-q-default
      bandwidth remaining percent 40
     random-detect minimum-threshold 120 kbytes maximum-threshold 120 kbytes drop-probability
 1 weight 0
switch(config-if)#

switch(config-if)# sh queuing interf eth1/29

slot  1
=======


Egress Queuing for Ethernet1/29 [Interface]
------------------------------------------------------------------------------
QoS-Group# Bandwidth% PrioLevel                Shape                     QLimit
                                     Min          Max        Units
------------------------------------------------------------------------------
      7             -         1           -            -     -            2(D)
      6             -         2           -            -     -            2(D)
      5             -         3           -            -     -            2(D)
      4             -         4           -            -     -            2(D)
      3             -         5  2500000000   2500000000   bps            2(D)
      2             -         6  5000000000   5000000000   bps            2(D)
      1            60         -           -            -     -            2(D)
      0            40         -           -            -     -            2(D)
+-------------------------------------------------------------+
|                             QOS GROUP 0                     |
+-------------------------------------------------------------+
|                           |  Ingress       |Egress          |
+-------------------------------------------------------------+
|                   Tx Pkts |              0|      2887259853|
|                   Tx Byts |              0|      1562745132|
|             Tail Drop Pkts |              0|      1440300408|
|               Q Depth Byts |              0|          126160|
|              ECN Mark Pkts |              0|               0|
|              ECN Mark Byts |              0|               0|
|             WRED Drop Pkts |              0|         8925945|
|             WRED Drop Byts |              0|     13388917500|
+-------------------------------------------------------------+
|                             QOS GROUP 1                     |
+-------------------------------------------------------------+
|                           |  Ingress       |Egress          |
+-------------------------------------------------------------+
|                   Tx Pkts |              0|        15763298|
|                   Tx Byts |              0|      2170110520|
|             Tail Drop Pkts |              0|        20771625|
|               Q Depth Byts |              0|               0|
|              ECN Mark Pkts |              0|               0|
|              ECN Mark Byts |              0|               0|
|             WRED Drop Pkts |              0|               0|
|             WRED Drop Byts |              0|               0|
+-------------------------------------------------------------+
|                             QOS GROUP 2                     |
+-------------------------------------------------------------+
|                           |  Ingress       |Egress          |
+-------------------------------------------------------------+
```

```
|                    Tx Pkts |               0|        29990985|
|                    Tx Byts |               0|      2036804540|
|             Tail Drop Pkts |               0|        31638183|
|               Q Depth Byts |               0|               0|
|              ECN Mark Pkts |               0|               0|
|              ECN Mark Byts |               0|               0|
|             WRED Drop Pkts |               0|               0|
|             WRED Drop Byts |               0|               0|
+-----------------------------------------------------------------+
|                              QOS GROUP 3                        |
+-----------------------------------------------------------------+
|                            |  Ingress       |Egress            |
+-----------------------------------------------------------------+
|                    Tx Pkts |               0|        14911143|
|                    Tx Byts |               0|       891878020|
|             Tail Drop Pkts |               0|        46351570|
|               Q Depth Byts |               0|               0|
|              ECN Mark Pkts |               0|               0|
|              ECN Mark Byts |               0|               0|
|             WRED Drop Pkts |               0|               0|
|             WRED Drop Byts |               0|               0|
+-----------------------------------------------------------------+
|                              QOS GROUP 4                        |
+-----------------------------------------------------------------+
|                            |  Ingress       |Egress            |
+-----------------------------------------------------------------+
|                    Tx Pkts |               0|               0|
|                    Tx Byts |               0|               0|
|             Tail Drop Pkts |               0|               0|
|               Q Depth Byts |               0|               0|
|              ECN Mark Pkts |               0|               0|
|              ECN Mark Byts |               0|               0|
|             WRED Drop Pkts |               0|               0|
|             WRED Drop Byts |               0|               0|
+-----------------------------------------------------------------+
|                              QOS GROUP 5                        |
+-----------------------------------------------------------------+
|                            |  Ingress       |Egress            |
+-----------------------------------------------------------------+
|                    Tx Pkts |               0|               0|
|                    Tx Byts |               0|               0|
|             Tail Drop Pkts |               0|               0|
|               Q Depth Byts |               0|               0|
|              ECN Mark Pkts |               0|               0|
|              ECN Mark Byts |               0|               0|
|             WRED Drop Pkts |               0|               0|
|             WRED Drop Byts |               0|               0|
+-----------------------------------------------------------------+
|                              QOS GROUP 6                        |
+-----------------------------------------------------------------+
|                            |  Ingress       |Egress            |
+-----------------------------------------------------------------+
|                    Tx Pkts |               0|               0|
|                    Tx Byts |               0|               0|
|             Tail Drop Pkts |               0|               0|
|               Q Depth Byts |               0|               0|
|              ECN Mark Pkts |               0|               0|
|              ECN Mark Byts |               0|               0|
|             WRED Drop Pkts |               0|               0|
|             WRED Drop Byts |               0|               0|
+-----------------------------------------------------------------+
|                              QOS GROUP 7                        |
+-----------------------------------------------------------------+
|                            |  Ingress       |Egress            |
```

```
+-------------------------------------------------------------+
|                    Tx Pkts |               0|            248|
|                    Tx Byts |               0|          57692|
|             Tail Drop Pkts |               0|              0|
|               Q Depth Byts |               0|              0|
|              ECN Mark Pkts |               0|              0|
|              ECN Mark Byts |               0|              0|
|             WRED Drop Pkts |               0|              0|
|             WRED Drop Byts |               0|              0|

switch(config-if)#
```

This example shows how to display all qos-groups, control qos group, span qos-group, and pfc statistics:

```
switch# show queuing interface ethernet1/11
Egress Queuing for Ethernet1/11 [System]
------------------------------------------------------------------------------
QoS-Group# Bandwidth% PrioLevel                Shape                     QLimit
                                     Min          Max        Units
------------------------------------------------------------------------------
      0          100         -           -            -     -            7(D)
+-------------------------------------------------------------------+
|                              QOS GROUP 0                          |
+-------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-------------------------------------------------------------------+
|                              QOS GROUP 1                          |
+-------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-------------------------------------------------------------------+
|                              QOS GROUP 2                          |
+-------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-------------------------------------------------------------------+
|                              QOS GROUP 3                          |
+-------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-------------------------------------------------------------------+
|                              QOS GROUP 4                          |
+-------------------------------------------------------------------+
```

```
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-----------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-----------------------------------------------------------------------+
|                             QOS GROUP 5                               |
+-----------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-----------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-----------------------------------------------------------------------+
|                             QOS GROUP 6                               |
+-----------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-----------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-----------------------------------------------------------------------+
|                             QOS GROUP 7                               |
+-----------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-----------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-----------------------------------------------------------------------+
|                        CONTROL QOS GROUP                              |
+-----------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-----------------------------------------------------------------------+
|        Tx Pkts |            1055|               0|               0|
|        Tx Byts |           83011|               0|               0|
|   Dropped Pkts |               7|               0|               0|
|   Dropped Byts |             508|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-----------------------------------------------------------------------+
|                          SPAN QOS GROUP                               |
+-----------------------------------------------------------------------+
|                |  Unicast       | OOBFC Unicast  |  Multicast     |
+-----------------------------------------------------------------------+
|        Tx Pkts |               0|               0|               0|
|        Tx Byts |               0|               0|               0|
|   Dropped Pkts |               0|               0|               0|
|   Dropped Byts |               0|               0|               0|
|   Q Depth Byts |               0|               0|               0|
+-----------------------------------------------------------------------+

Port Egress Statistics
------------------------------------------------------
WRED Drop Pkts                              0
```

```
Ingress Queuing for Ethernet1/11
------------------------------------------------------------
QoS-Group#                 Pause                       QLimit
           Buff Size       Pause Th      Resume Th
------------------------------------------------------------
      7              -            -            -     11210992(S)
      6              -            -            -     11210992(S)
      5              -            -            -     11210992(S)
      4              -            -            -     11210992(S)
      3              -            -            -     11210992(S)
      2              -            -            -     11210992(S)
      1              -            -            -     11210992(S)
      0              -            -            -     11210992(S)


Port Ingress Statistics
------------------------------------------------------
Ingress MMU Drop Pkts                              0
Ingress MMU Drop Bytes                             0

PFC Statistics
------------------------------------------------------------------------------
TxPPP:                    0, RxPPP:                    0
------------------------------------------------------------------------------
 COS QOS Group        PG   TxPause   TxCount         RxPause         RxCount
   0         -         7  Inactive         0        Inactive               0
   1         -         7  Inactive         0        Inactive               0
   2         -         7  Inactive         0        Inactive               0
   3         -         7  Inactive         0        Inactive               0
   4         -         7  Inactive         0        Inactive               0
   5         -         7  Inactive         0        Inactive               0
   6         -         7  Inactive         0        Inactive               0
   7         -         7  Inactive         0        Inactive               0
------------------------------------------------------------------------------
```

This example shows how to display all qos-groups, control qos group, span qos-group, and pfc statistics on Cisco Nexus 3500 platform switches:

```
switch# show queuing interface ethernet1/1
Egress Queuing for Ethernet1/11 [System]
slot 1
=======


HW MTU of Ethernet1/1 : 1500 bytes

Egress Queuing for Ethernet1/1 [System]
------------------------------------------------------------------------------
QoS-Group# Bandwidth% PrioLevel Shape QLimit
Min Max Units
------------------------------------------------------------------------------
0 100 - - - - 7(D)

Mcast pkts dropped : 0

+---------------------------------------------------------------------+
| QOS GROUP 0 |
+---------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+---------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+---------------------------------------------------------------------+
| QOS GROUP 1 |
```

```
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+
| QOS GROUP 2 |
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+
| QOS GROUP 3 |
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+
| QOS GROUP 4 |
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+
| QOS GROUP 5 |
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+
| QOS GROUP 6 |
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+
| QOS GROUP 7 |
+-------------------------------------------------------------------+
| | Unicast | OOBFC Unicast | Multicast |
+-------------------------------------------------------------------+
| Dropped Pkts | 0| 0| 0|
+-------------------------------------------------------------------+


Ingress Queuing for Ethernet1/1
------------------------------------------------------------------------
QoS-Group# Pause QLimit
Buff Size Pause Th Resume Th
------------------------------------------------------------------------
7 - - - 1511506826(S)
6 - - - 1511506826(S)
5 - - - 1511506826(S)
4 - - - 1511506826(S)
3 - - - 1511506826(S)
2 - - - 1511506826(S)
1 - - - 1511506826(S)
0 - - - 1511506826(S)


Port Ingress Statistics
------------------------------------------------------------
Ingress MMU Drop Pkts 0
Ingress MMU Drop Bytes 0

switch#
```

# Configuring QoS on Interfaces

## Configuring Untagged CoS

Any incoming packet not tagged with an 802.1p CoS value is assigned the default untagged CoS value of zero (which maps to the default Ethernet drop system class). You can override the default untagged CoS value for an Ethernet or EtherChannel interface.

**Note** The Cisco Nexus 34180YC switch does not support **untagged cos** *cos-value* .

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **interface** {**ethernet** [*chassis*/]*slot*/*port* \| **port-channel** *channel-number*} | Enters the configuration mode for the specified interface or port channel. |
| **Step 3** | switch(config-if)# **untagged cos** *cos-value* | Configures the untagged CoS value. Values can be from 1 to 7. |

#### Example

The following example shows how to set the CoS value to 4 for untagged frames received on an interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/2
switch(config-if)# untagged cos 4
```

## Configuring an Interface Service Policy

An input qos policy is a service policy applied to incoming traffic on an Ethernet interface for classification. For type queuing, the output policy is applied to all outgoing traffic that matches the specified class.

### Procedure

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

| | Command or Action | Purpose |
|---|---|---|
| **Step 2** | switch(config)# **interface** {**ethernet** [*chassis*/]*slot*/*port* \| **port-channel** *channel-number*} | Enters the configuration mode for the specified interface.<br>**Note**<br>The service policy on a port channel applies to all member interfaces. |
| **Step 3** | switch(config-if)# **service-policy input** *policy-name* | Applies the policy map to the interface.<br>**Note**<br>There is a restriction that system type qos policy cannot be the same as any the type qos policy applied to an interface or EtherChannel. |

### Example

This example shows how to apply a policy to an Ethernet interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/1
switch(config-if)# service-policy type qos input policy1
```

# Verifying the QoS Configuration

To verify the QoS configurations, perform one of these tasks:

**Note** All the commands and/or options may not work on Cisco Nexus 3500 switches. For more information on the supported commands you may refer *Cisco Nexus 3548 Switch NX-OS Quality of Service Configuration Guide, Release 6.x*

| Command | Purpose |
|---|---|
| switch# **show class-map** | Displays the class maps defined on the device. |
| switch# **show policy-map** [*name*] | Displays the policy maps defined on the device. Optionally, you can display the named policy only. |
| switch# **show policy-map interface** [*interface number*] | Displays the policy map settings for an interface or all interfaces. |
| switch# **show policy-map system** | Displays the policy map settings attached to the system qos. |

| Command | Purpose |
|---|---|
| switch# **show policy-map type** {**network-qos** \| **qos** \| **queuing**} [*name*] | Displays the policy map settings for a specific policy type. Optionally, you can display the named policy only. |
| switch# **show interface untagged-cos** [**module** *number*] | Displays the untagged CoS values for all interfaces. |
| switch# **show wrr-queue cos-map** [*var*] | Displays the mapped CoS values to egress queues.<br>**Note**<br>This command may not work on Cisco Nexus 3500 Series switches. |
| switch# **show running-config ipqos** | Displays information about the running configuration for QoS. |
| switch# **show startup-config ipqos** | Displays information about the startup configuration for QoS. |
| switch# **show queuing interface ethernet** *slot-no/port-no* | Displays the queuing information on interfaces. |
| switch# **show queuing** | Displays the queuing information configured for all interfaces. It includes shaper configuration information for each class, the control queue Tx and drop statistics for each port, and WRED drop packet counts. |

This example shows how to configure a network QoS policy:

```
switch(config)# class-map type network-qos cnq1
switch(config-cmap-nq)# match qos-group 1
switch(config-cmap-nq)# exit
switch(config)# class-map type network-qos cnq6
switch(config-cmap-nq)# match qos-group 6
switch(config-cmap-nq)# exit
switch(config)# policy-map type network-qos pnqos
switch(config-pmap-nqos)# class type network-qos cnq1
switch(config-pmap-nqos-c)# mtu 2200
switch(config-pmap-nqos-c)# pause no-drop
switch(config-pmap-nqos-c)# set cos 4
switch(config-pmap-nqos-c)# exit
switch(config-pmap-nqos)# class type network-qos cnq6
switch(config-pmap-nqos-c)# mtu 2200
switch(config-pmap-nqos-c)# pause no-drop
switch(config-pmap-nqos-c)# set cos 5
switch(config-pmap-nqos-c)# congestion-control random-detect ecn
switch(config-pmap-nqos-c)# exit
switch(config-pmap-nqos)# class type network-qos class-default
switch(config-pmap-nqos-c)# mtu 9216
switch(config-pmap-nqos-c)# exit
switch(config-pmap-nqos)# system qos
switch(config-sys-qos)# service-policy type network-qos pnqos
switch(config-sys-qos)#
```

This example shows how to configure an output queuing policy:

```
switch(config)# class-map type queuing cqu1
switch(config-cmap-que)# match qos-group 1
switch(config-cmap-que)# exit
switch(config)# class-map type queuing cqu6
switch(config-cmap-que)# match qos-group 6
switch(config-cmap-que)# exit
switch(config)# policy-map type queuing pqu
switch(config-pmap-que)# class type queuing class-default
switch(config-pmap-c-que)# bandwidth percent 70
switch(config-pmap-c-que)# exit
switch(config-pmap-que)# class type queuing cqu1
switch(config-pmap-c-que)# bandwidth percent 10
switch(config-pmap-c-que)# exit
switch(config-pmap-que)# class type queuing cqu6
switch(config-pmap-c-que)# bandwidth percent 20
switch(config-pmap-c-que)# exit
switch(config-pmap-que)# exit
switch(config)# system qos
switch(config-sys-qos)# service-policy type queuing output pqu
switch(config-sys-qos)#
```

This example shows how to configure an input queuing policy:

```
switch(config)# class-map type queuing cqu1
switch(config-cmap-que)# match qos-group 1
switch(config-cmap-que)# exit
switch(config)# class-map type queuing cqu6
switch(config-cmap-que)# match qos-group 6
switch(config-cmap-que)# exit
switch(config)# policy-map type queuing piqu
switch(config-pmap-que)# class type queuing cqu1
switch(config-pmap-c-que)# pause buffer-size 39936 pause-threshold 24960 resume-threshold
12480
switch(config-pmap-c-que)# pause priority-group 1
switch(config-pmap-c-que)# exit
switch(config-pmap-que)# class type queuing cqu6
switch(config-pmap-c-que)# pause priority-group 3
switch(config-pmap-c-que)# exit
switch(config-pmap-que)# class type queuing class-default
switch(config-pmap-c-que)# queue-limit dynamic 2
switch(config-pmap-c-que)# exit
switch(config-pmap-que)# exit
switch(config)# system qos
switch(config-sys-qos)# service-policy type queuing output piqu
switch(config-sys-qos)#
```

This example shows how to configure a QoS policy:

```
switch(config)# class-map type qos cqos1
switch(config-cmap-qos)# match cos 1
switch(config-cmap-qos)# exit
switch(config)# class-map type qos cqos6
switch(config-cmap-qos)# match cos 6
switch(config-cmap-qos)# exit
switch(config)# policy-map type qos pqos
switch(config-pmap-qos)# class type qos cqos1
switch(config-pmap-c-qos)# set qos-group 1
switch(config-pmap-c-qos)# exit
switch(config-pmap-qos)# class type qos cqos6
switch(config-pmap-c-qos)# set qos-group 6
switch(config-pmap-c-qos)# exit
switch(config-pmap-qos)# exit
switch(config)# system qos
```

```
switch(config-sys-qos)# service-policy type qos input pqos
switch(config-sys-qos)#
```

This example shows how to verify the untagged-cos configuration on interfaces:

```
switch(config-if)# show interface untagged-cos
================================

Interface      Untagged-CoS
================================

Ethernet1/1  4
Ethernet1/2
Ethernet1/3  5
Ethernet1/4
Ethernet1/5
Ethernet1/6
Ethernet1/7
Ethernet1/8
Ethernet1/9
Ethernet1/10
Ethernet1/11
Ethernet1/12
Ethernet1/13
Ethernet1/14
Ethernet1/15
Ethernet1/16
Ethernet1/17
```

This example shows how to display the QoS running configuration:

```
switch(config)# show running-config ipqos
!Command: show running-config ipqos
!Time: Tue Dec 10 08:29:08 2013

version 6.0(2)U2(1)
class-map type qos match-all c1
  match cos 1
class-map type qos match-all c2
  match cos 2
.
.
.
class-map type qos match-any cq1
  match cos 4
class-map type qos match-any cq2
  match cos 5
class-map type qos match-all dscp
  match precedence 0
class-map type qos match-all cq1_1
  match cos 4
  match precedence 7
class-map type qos match-all cq2_1
  match cos 5
  match precedence 3
class-map type qos match-all cMap_Cos_1
  match cos 1
class-map type qos match-all cMap_Cos_2
  match cos 2
.
.
.

class-map type queuing cMap_Qing_match_qGrp_7
  match qos-group 7
```

```
policy-map type qos inpq
  class c5
    set qos-group 5
.
.
.
policy-map type queuing piqu
  class type queuing cqu1
    pause buffer-size 39936 pause-threshold 24960 resume-threshold 12480
    pause priority-group 1
  class type queuing cqu6
    pause priority-group 3
.
.
.

class-map type network-qos cMap_NQ_match_qGrp_7
  match qos-group 7
policy-map type network-qos pnqos
  class type network-qos cnq1
    mtu 2200
    pause no-drop
    set cos 4
  class type network-qos cnq6
    mtu 2200
    pause no-drop
    set cos 5
    congestion-control random-detect ecn
.
.
.
system qos
  service-policy type qos input pMap_Qos_system
  service-policy type network-qos pMap_NQ_system
  service-policy type queuing output pMap_Qing_system


interface Ethernet1/1/1
  priority-flow-control mode on

interface Ethernet1/32
  priority-flow-control mode on
```

This example shows how to display the QoS groups that are mapped to the egress queue:

```
switch(config)# wrr-queue qos-group-map 3 1
switch(config)# show wrr-queue qos-group-map
MCAST Queue ID           Qos-Group Map
0                       0
1                       2 3
2                       4 5
3                       1 6 7
switch(config)#
```

This example shows how to display the class map configuration:

```
switch(config)# show class-map


  Type qos class-maps
  ===================

    class-map type qos match-all cqos1
      match cos 1
```

```
    class-map type qos match-all cqos6
      match cos 6

    class-map type qos match-any class-default
      match any


  Type queuing class-maps
  =======================

    class-map type queuing cqu1
      match qos-group 1

    class-map type queuing cqu6
      match qos-group 6

    class-map type queuing class-default
      match qos-group 0



  Type network-qos class-maps
  ===========================

    class-map type network-qos cnq1
      match qos-group 1

    class-map type network-qos cnq6
      match qos-group 6

    class-map type network-qos class-default
      match qos-group 0

switch(config)#
```

This example shows how to display the policy map configuration:

```
switch(config)# show policy-map


  Type qos policy-maps
  ====================

  policy-map type qos pqos
    class type qos cqos1
      set qos-group 1
    class type qos cqos6
      set qos-group 6
    class type qos class-default
      set qos-group 0
  policy-map type qos default-in-policy
    class type qos class-default
      set qos-group 0

  Type queuing policy-maps
  ========================
  policy-map type queuing piqu
    class type queuing cqu1
      pause buffer-size 39936 pause-threshold 24960 resume-threshold 12480
      pause priority-group 1
    class type queuing cqu6
      pause priority-group 3
    class type queuing class-default
      bandwidth percent 100
```

```
      queue-limit dynamic 2


  Type network-qos policy-maps
  ============================

  policy-map type network-qos pnqos
    class type network-qos cnq1
      mtu 1500
      set cos 4
    class type network-qos cnq6
      mtu 1500
      set cos 5
      congestion-control random-detect ecn
    class type network-qos class-default
      mtu 9216
  policy-map type network-qos default-nq-policy
    class type network-qos class-default
      mtu 1500
switch(config)#
```

This example shows how to display all active policy maps in the system:

```
switch(config)# show policy-map system
 Type network-qos policy-maps
 ============================
 policy-map type network-qos pnqos
    class type network-qos cnq1
      match qos-group 1

      mtu 2200
      pause no-drop
      set cos 4
    class type network-qos cnq6
      match qos-group 6

      mtu 2200
      pause no-drop
      set cos 5
      congestion-control random-detect ecn
    class type network-qos class-default
      match qos-group 0

      mtu 9216

  Service-policy (qos) input:   pMap_Qos_system
    policy statistics status:   enabled

Class-map (qos):   cqos1 (match-all)
      Match: cos 1
      set qos-group 1

    Class-map (qos):   cqos6 (match-all)
      Match: cos 6
      set qos-group 6

    Class-map (qos):   class-default (match-any)
      Match: any
      set qos-group 0

  Service-policy (queuing) output:   pqu
    policy statistics status:   disabled

    Class-map (queuing):   cqu1 (match-any)
      Match: qos-group 1
```

```
      bandwidth percent 10

    Class-map (queuing):   cqu6 (match-any)
      Match: qos-group 6
      bandwidth percent 20

    Class-map (queuing):   class-default (match-any)
      Match: qos-group 0
      bandwidth percent 70

switch(config)#
```

This example shows how to display the service policy maps configured on the interfaces:

```
switch(config)# show policy-map interface ethernet 1/1


Global statistics status :   enabled

Ethernet1/1

  Service-policy (qos) input:   pqos
    policy statistics status:   enabled

    Class-map (qos):   cqos1 (match-all)
      Match: cos 1
      set qos-group 1

    Class-map (qos):   cqos6 (match-all)
      Match: cos 6
      set qos-group 6

    Class-map (qos):   class-default (match-any)
      Match: any
      set qos-group 0

  Service-policy (queuing) output:   pqu
    policy statistics status:   enabled

    Class-map (queuing):   cqu1 (match-any)
      Match: qos-group 1
      bandwidth percent 10

    Class-map (queuing):   cqu6 (match-any)
      Match: qos-group 6
      bandwidth percent 20

    Class-map (queuing):   class-default (match-any)
      Match: qos-group 0
      bandwidth percent 70

switch(config)#
```

This example shows how to display the queuing information configured for all interfaces:

```
switch# show queuing
Egress Queuing for Ethernet1/1 [Interface]
------------------------------------------------------------------
QoS-Group# Bandwidth% PrioLevel                Shape
                                     Min          Max        Units
------------------------------------------------------------------
      0            10         -          0          0              -
      1            10         -          0          0              -
      2            10         -          0          0              -
```

```
      3            10         1          0          0               -
      4            10         -          0          0               -
      5            10         2          0          0               -
      6            10         -          0          0               -
      7            10         -          0          0               -
      9             0         -          0          0               -
+-------------------------------------------------------------------+
|                            QOS GROUP 0                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                            QOS GROUP 1                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                            QOS GROUP 2                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                            QOS GROUP 3                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                            QOS GROUP 4                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                            QOS GROUP 5                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                            QOS GROUP 6                            |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
```

```
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                             QOS GROUP 7                           |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                           CONTROL QOS GROUP 9                     |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|            1901|               0
|         Tx Byts|          145235|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+---------------------------------------------------+
Port Egress Statistics
-----------------------------------------------------
WRED Drop Pkts                0


...

Egress Queuing for Ethernet1/4 [Interface]
-------------------------------------------------------------------
QoS-Group# Bandwidth% PrioLevel                Shape
                                    Min          Max        Units
-------------------------------------------------------------------
      0          100         -          0          0               -
      9            0         -          0          0               -
+-------------------------------------------------------------------+
|                             QOS GROUP 0                           |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|               0|               0
|         Tx Byts|               0|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+-------------------------------------------------------------------+
|                           CONTROL QOS GROUP 9                     |
+-------------------------------------------------------------------+
|                |  Unicast       |  Multicast     |
+-------------------------------------------------------------------+
|         Tx Pkts|            8634|               0
|         Tx Byts|         1218248|               0
|    Dropped Pkts|               0|               0
|    Dropped Byts|               0|               0
+---------------------------------------------------+
Port Egress Statistics
-----------------------------------------------------
WRED Drop Pkts                0
```

# Monitoring the QoS Packet Buffer

Cisco Nexus 3000 series switches have a 9-MB buffer memory and the Cisco Nexus 3100 platform switches have a 12-MB buffer memory that divides into a dedicated per port and dynamic shared memory. Each front-panel port has eight unicast and four multicast queues in egress. The Cisco Nexus 3100 platform switches have eight unicast and eight multicast queues in egress. In the scenario of burst or congestion, each egress port consumes buffers from the dynamic shared memory.

For the Cisco Nexus 3000 series switches, you can display the real-time status of the shared buffer per port. For the Cisco Nexus 3100 platform switches, you can also display the peak status of the shared buffer per port. All counters are displayed in terms of the number of cells. Each cell is 208 bytes in size. You can also display the global level buffer consumption in terms of consumption and available number of cells.

**Note** For the Cisco Nexus 34180YC switch:

- The maximum amount of data in a cell is 80 bytes.
- The total number of cells available is 194224.
- The total memory is 15 MB (194224 * 80 = 15537920 bytes).

```
switch(config-if)# sh hardware internal buffer info pkt-stats

slot  1
=======


INSTANCE: 0

============


|-------------------------------------------------------------------------------------|
                     Output Buffer Pool Utilization (in cells)
                                    BP-0           BP-1           BP-2           BP-3
|-------------------------------------------------------------------------------------|

Total Instant Usage                 1537           0              -              -
Remaining Instant Usage             192687         5200           -              -
Peak/Max Cells Used                 38965          228            -              -
Switch Cell Count                   194224         5200           -              -
switch(config-if)#
```

Starting with Release 6.0(2)U5(1), the buffer is carved into two pools: pool 0 and pool 1. Cisco Nexus 3000 Series switches that have Broadcom T+ have 9MB of total buffer and Broadcom T2 have 12MB of total buffer. There are three buffer modes – all, default, and none. The default mode of buffer in Cisco Nexus 3000 series switches is all.

**Note** The Cisco Nexus 34180YC switch does not support the all, default, and none buffer modes.

In default mode all for T+, there are 10 unicast queues and 5 multicast queues. Each queue in T+ consumes 8 cells. Therefore, 10x8+5x8=120 cells are consumed by each port. T+ has total 46080 cells. The all mode consumes 120 cells/port in T+.

In default mode all for T2, there are 10 unicast queues and 10 multicast queues. Each queue in T2 consumes 11 cells. Therefore, 10x11+10x11=220 cells are consumed by each port. T2 has total 61440 cells. The all mode consumes 220 cells/port in T2.

Cisco Nexus 3000 series switches that are based out of T+ and T2 have 48 CPU multicast queues.

The **default** mode consumes for one default unicast queue, multicast queue, unicast CPU queue (q9), and multicast cpu queue (q9) totaling 11x4=44 cells. The **none** mode consumes only for unicast CPU queue (q9) and multicast CPU queue (q9) – totaling 11x2=22 cells.

The maximum cells consumed in each platform are:

- Platform T2 based Product– Max Cells for All Mode is 23586 cells.
- Platform T2 based Product– Max Cells for Default Mode is 29903 cells.
- Platform T2 based Product – Max Cells for None Mode is 30677 cells.

The **show running | grep hardware** command displays the configured mode in the CLI output:

```
switch# show running | grep hardware
hardware profile portmode 48X10G+breakout6x40g
hardware qos min-buffer qos-group none
```

Usage information of shared buffer resources:

- Total Instant Usage—Current buffer usage in terms of the number of cells on a global basis.
- Remaining Instant Usage—Effective free number of cells available on a global basis.
- Max Cell Usage—Maximum buffer usage that is seen until the system buffer maximum cell usage counter is last cleared.
- Switch Cell Count—Total global buffer space available in the platform in terms of the number of cells on a global basis.

UC and MC represent the 8 unicast (Q1-Q8) and 8 multicast (Q1-Q8) instant cell usage. The example above shows that the multicast queue Q1 is consuming 3807 cells instantaneously on port 9.

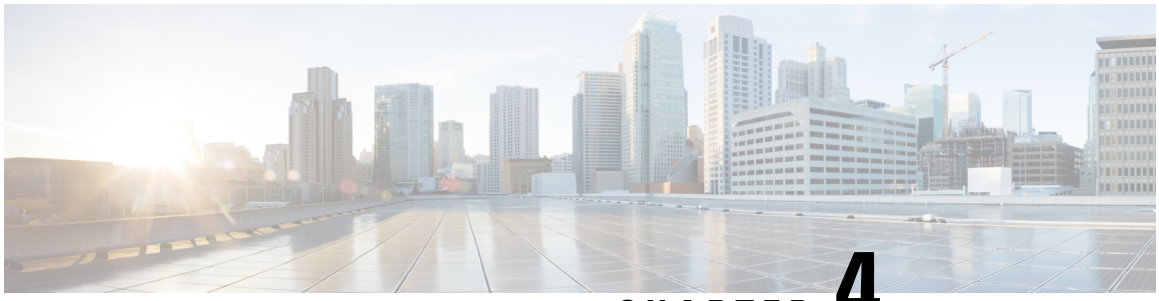
The following example shows how to clear the system buffer maximum cell usage counter:

```
switch# clear counters buffers
Max Cell Usage has been reset successfully
```

The following example shows how to set a buffer utilization threshold on a per port basis. If the buffer occupancy exceeds this number, you can generate a syslog.

**Note** The Cisco Nexus 34180YC switch does not support buffer utilization threshold on a per port basis.

```
switch# hardware profile buffer info port-threshold front-port  1 threshold  10
Port threshold changed successfully
```



CHAPTER 4

# Configuring Priority Flow Control

This chapter contains the following sections:


- About Priority Flow Control, on page 69
- Guidelines and Limitations for Priority Flow Control, on page 70
- Default Settings for Priority Flow Control, on page 72
- Enabling Priority Flow Control on a Traffic Class, on page 72
- Configuring Priority Flow Control, on page 74
- Reserving mmu-buffer for PFC, on page 75
- Configuring a Priority Flow Control Watchdog Interval, on page 75
- Verifying the Priority Flow Control Configuration, on page 78
- Monitoring PFC Frame Counter Statistics, on page 78
- Configuration Examples for Priority Flow Control , on page 79


## About Priority Flow Control

Priority flow control (PFC; IEEE 802.1Qbb), which is also referred to as Class-based Flow Control (CBFC) or Per Priority Pause (PPP), is a mechanism that prevents frame loss that is due to congestion. PFC functions on a per class-of-service (CoS) basis.

When a buffer threshold is exceeded due to congestion, PFC sends a pause frame that indicates which CoS value needs to be paused. A PFC pause frame contains a 2-octet timer value for each CoS that indicates the length of time that the traffic needs to be paused. The unit of time for the timer is specified in pause quanta. A quanta is the time that is required for transmitting 512 bits at the speed of the port. The range is from 0 to 65535. A pause frame with a pause quanta of 0 indicates a resume frame to restart the paused traffic.

By default, PFC is in the auto mode. However, no particular traffic class is enabled for pause.

**Note** Only certain classes of service of traffic can be flow controlled while other classes are allowed to operate normally.

PFC asks the peer to stop sending frames of a particular CoS value by sending a pause frame to a well-known multicast address. This pause frame is a one-hop frame that is not forwarded when received by the peer. When the congestion is mitigated, PFC can request the peer to restart transmitting frames.

**Note** RDMA over Converged Ethernet (RoCE) v1 and v2 protocols are supported on Cisco Nexus 3000 Series switches.

# Guidelines and Limitations for Priority Flow Control

PFC has the following guidelines and limitations:

- If PFC is enabled on a port or a port channel, it does not cause a port flap.
- Ensure that ports or port channels have enough resources before enabling PFC on them.
- PFC configuration enables PFC in both the send (Tx) and receive (Rx) direction.
- Only an exact match of the no-drop CoS is considered as a successful negotiation of PFC by the Data Center Bridging Exchange Protocol (DCBXP).
- Configuration time quanta of the pause frames is not supported.
- The configuration does not support pausing selected streams that are mapped to a particular traffic-class queue. All flows that are mapped to the class are treated as no-drop. It blocks out scheduling for the entire queue, which pauses traffic for all the streams in the queue. To achieve lossless service for a no-drop class, we highly recommend that you have only the no-drop class traffic on the queue.
- For VLAN-tagged packets, priority is always assigned based on the 802.1p field in the VLAN tag and takes precedence over the assigned internal priority(qos-group). DSCP or IP access-list classification cannot be performed on VLAN-tagged frames
- When a no-drop class is classified based on 802.1p CoS x and assigned an internal priority value (qos-group) of y, we recommend that you use the internal priority value x to classify traffic on 802.1p CoS only, and not on any other field. For x, the packet priority assigned is x if the classification is not based on CoS, which results in packets of the internal priority that is x and y to map to the same priority x.
- The PFC feature supports up to three no-drop classes of any MTU size. However, there is a limit on the number of PFC-enabled interfaces based on the following factors:
  - MTU size of the no-drop class
  - Number of 10G and 40G ports
  - Pause buffer size configuration in the input queuing policies
- Interface QoS policy takes precedence over the system policy. PFC priority derivation also occurs in the same order.
- Ensure that you apply the same interface-level QoS policy on all PFC-enabled interfaces for both ingress and egress.

**Caution** Irrespective of the PFC configuration, we recommend that you stop traffic before you apply or remove the queuing policy that has strict priority levels at the interface level or the system level.

- To achieve end-to-end lossless service over the network, we recommend that you enable PFC on each interface through which the no-drop class traffic flows #(Tx/Rx).
- To achieve lossless service for a no-drop class, it is recommended that you maintain only no-drop class traffic on the egress queue.
- We recommend that you change the PFC configuration when there is no traffic; otherwise, packets already in the memory management unit (MMU) of the system might not get the expected treatment.
- The buffers required for PFC are best allocated automatically. However, you can change the buffer thresholds by configuring input queuing policies.
- For no-drop classes classified based on DSCP/IP access-lists (non CoS based classifications), we highly recommend that you use the same qos-group value as the match CoS value.
- Do not enable WRED on a no-drop class because it results in egress queue drops.
- When you configure a port from the 40 Gigabit Ethernet mode to the 10 Gigabit Ethernet mode or from the 10 Gigabit Ethernet mode to the 40 Gigabit Ethernet mode, the affected ports will be administratively unavailable and PFC will be disabled on these ports. To make these ports available, use the **no shut** command. After the ports are available, PFC will become enabled on them.
- Beginning with Cisco NX-OS Release 7.0(3)I4(2), the **no lldp tlv-select dcbxp** command is enhanced so that the PFC is disabled for interfaces on both sides of back-to-back switches.
- When the 'hardware qos pfc mc-drop' configuration is enabled on Cisco Nexus 3064 switches, and if the no-drop and drop qos-groups are mapped to the same queue, then flapping the link or removing/adding PFC configurations will result in the drop-multicast feature not working correctly. This is due to the hardware limitation of having only four multicast queues in Cisco Nexus 3064 switches. To avoid this issue, do one of the following:
  - Specify both of the qos-groups that correspond to a single multicast queue as no-drop.
  - Change the mapping of the multicast queue to qos-groups using the **wrr-queue qos-group-map <queue-no> <qos-groups that are no-drop>** command.
- Beginning with Cisco NX-OS Release 7.0(3)I7(8), you can see additional syslog messages for multicast queue drops on no-drop class when you enable **hardware qos pfc mc-drop** global configuration on Cisco Nexus 3000 and 3100 switches. However this feature is not supported on Cisco Nexus C3132Q-V, C31108PC-V, C31108TC-V and C3132C-Z switches.
- Beginning with Cisco NX-OS Release 7.0(3)I7(4), when PFC is received on a lossy priority group (non-configured), the event is recorded in the syslog for subsequent analysis.

  Cisco Nexus N3000 and N3100 series switches report BCM_UNEXPECTED_PFC_FRAMES syslog whenever PFC frames are received with unexpected CoS. The syslog contains the approximate count of the unexpected PFC frames received for a particular CoS in the last two seconds. The packets per second (PPS) metric can be derived by dividing this number by two.

- Beginning with Cisco NX-OS Release 7.0(3)I7(4), switches can be configured to drop multicast/broadcast traffic on no-drop configured classes with the **hardware qos pfc mc-drop** command.
- Beginning with Cisco NX-OS Release 7.0(3)I7(4), traffic across all no-drop queues and incoming PFC frames for no-drop classes can be configured to be dropped with the **priority-flow-control watch-dog forced on** command. (Use the **no priority-flow-control watch-dog forced on** command to re-enable the traffic for no-drop classes.)
- When a queue (under an interface) becomes stuck, you can use the **priority-flow-control watch-dog-interval on disable-action** command to send a message to the syslog that describes the status of the queue instead of shutting the queue (Cisco NX-OS Release 7.0(3)I7(4) and later).

  Example:

```
switch(config)# interface ethernet 1/12

switch(config-if)# priority-flow-control watch-dog-interval on disable-action
```

# Default Settings for Priority Flow Control

The following table lists the default setting for PFC.

***Table 5: Default PFC Setting***

| Parameter | Default |
|---|---|
| PFC | Auto |

# Enabling Priority Flow Control on a Traffic Class

You can enable PFC on a particular traffic class:

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **class-map type qos** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 3** | switch(config-cmap-qos)# **match cos** *cos-value* | Specifies the CoS value to match for classifying packets into this class. You can configure a CoS value in the range of 0 to 7. |
| **Step 4** | switch(config-cmap-qos)# **exit** | Exits class-map mode and enters global configuration mode. |

| | Command or Action | Purpose |
|---|---|---|
| **Step 5** | switch(config)# **policy-map type qos** *policy-name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 6** | switch(config-pmap-qos)# **class type qos** *class-name* | Associates a class map with the policy map, and enters configuration mode for the specified system class.<br>**Note**<br>The associated class map must be the same type as the policy map type. |
| **Step 7** | switch(config-pmap-c-qos)# **set qos-group** *qos-group-value* | Configures one or more qos-group values to match on for classification of traffic into this class map. There is no default value. |
| **Step 8** | switch(config-pmap-c-qos)# **exit** | Exits policy-map mode and enters global configuration mode. |
| **Step 9** | switch(config)# **interface** *type slot/port* | Enters the configuration mode for the specified interface. |
| **Step 10** | switch(config-if)# **service-policy type qos input** *policy-name* | Applies the policy map of type qos to the specific interface. |
| **Step 11** | switch(config-if)# **exit** | Exits interface configuration mode and enters global configuration mode. |
| **Step 12** | switch(config)# **class-map type network-qos** *class-name* | Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters. |
| **Step 13** | switch(config-cmap-nq)# **match qos-group** *qos-group-value* | Configures the traffic class by matching packets based on a list of QoS group values. Values can range from 0 to 7. QoS group 0 is equivalent to class-default. |
| **Step 14** | switch(config-cmap-nq)# **exit** | Exits class-map mode and enters global configuration mode. |
| **Step 15** | switch(config)# **policy-map type network-qos** *policy-name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. 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 16** | switch(config-pmap-nqos)# **class type network-qos** *class-name* | Associates a class map with the policy map, and enters configuration mode for the specified system class.<br>**Note**<br>The associated class map must be the same type as the policy map type. |
| **Step 17** | switch(config-pmap-c-nq)# **pause no-drop** | Configures a no-drop class. |
| **Step 18** | switch(config-pmap-c-nq)# **exit** | Exits policy-map mode and enters global configuration mode. |
| **Step 19** | switch(config)# **system qos** | Enters system class configuration mode. |
| **Step 20** | switch(config-sys-qos)# **service-policy type network-qos** *policy-name* | Applies the policy map of type network-qos at the system level or to the specific interface. |

### Example

This example shows how to enable PFC on a traffic class .

```
switch# configure terminal
switch(config)# class-map type qos c1
switch(config-cmap-qos)# match cos 3
switch(config-cmap-qos)# exit
switch(config)# policy-map type qos p1
switch(config-pmap-qos)# class type qos c1
switch(config-pmap-c-qos)# set qos-group 3
switch(config-pmap-c-qos)# exit
switch(config)# interface ethernet 1/1
switch(config-if)# service-policy type qos input p1
switch(config-if)# exit
switch(config)# class-map type network-qos c1
switch(config-cmap-nq)# match qos-group 3
switch(config-cmap-nq)# exit
switch(config)# policy-map type network-qos p1
switch(config-pmap-nqos)# class type network-qos c1
switch(config-pmap-nqos-c)# pause no-drop
switch(config-pmap-nqos-c)# exit
switch(config)# system qos
switch(config-sys-qos)# service-policy type network-qos p1
```

# Configuring Priority Flow Control

You can configure PFC on a per-port basis to enable the no-drop behavior for the CoS as defined by the active network qos policy. PFC can be configured in one of these three modes:

- auto—Enables the no-drop CoS values to be advertised by the DCBXP and negotiated with the peer. A successful negotiation enables PFC on the no-drop CoS. Any failures because of a mismatch in the capability of peers causes the PFC not to be enabled.

- on—Enables PFC on the local port regardless of the capability of the peers.
- off—Disables PFC on the local port.

**Procedure**

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal** | Enters global configuration mode. |
| **Step 2** | **interface ethernet** [*slot/port-number*] | Enter the interface mode on the specified interface. |
| **Step 3** | **priority-flow-control mode** {**auto** \| **off** \| **on**}<br>**priority-flow-control mode {auto \|off \| on}** | Sets the PFC to the auto, off, or on mode. By default, PFC mode is set to auto on all ports. |
| **Step 4** | **exit** | Exits interface configuration mode. |
| **Step 5** | **show interface priority-flow-control** | Displays the status of PFC on all interfaces. |

# Reserving mmu-buffer for PFC

Complete the following steps to reserve mmu-buffer for PFC.

**Procedure**

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal**l | Enters global configuration mode. |
| **Step 2** | **hardware profile pfc mmu buffer-reservation ?**<br>**Example:**<br>switch(config)# hardware profile pfc mmu buffer-reservation ? | Reserves the mmu-buffer for PFC.<br><0-100> Percentage of shared pool buffers to be reserved |

# Configuring a Priority Flow Control Watchdog Interval

You can configure a PFC watchdog interval to detect whether packets in a no-drop queue are being drained within a specified time period. When the time period is exceeded, all incoming and outgoing packets are dropped on interfaces that match the PFC queue that is not being drained. This feature is supported beginning with Cisco NX-OS Release 6.0(3)U6(9) and only for Cisco Nexus 3000 Series switches.

Starting with Cisco NX-OS Release 7.0(3)I6(1), ingress packets will be dropped for matching to shutdown PFC queue or qos-group of a physical port, and the show command displays the status.

## Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal**<br>**Example:**<br>`switch# configure terminal`<br>`switch(config)#` | Enters global configuration mode. |
| **Step 2** | **priority-flow-control auto-restore multiplier** *value* | Configures a value for the PFC auto-restore multiplier. |
| **Step 3** | **priority-flow-control fixed-restore multiplier** *value* | Configures a value for the PFC fixed-restore multiplier. |
| **Step 4** | **priority-flow-control watch-dog-interval** {**on** \| **off**}<br>**Example:**<br>`switch(config)# priority-flow-control watch-dog-interval on` | Globally enables or disables the PFC watchdog interval for all interfaces.<br>**Note**<br>You can use this same command in interface configuration mode to enable or disable the PFC watchdog interval for a specific interface.<br>See the following example of the command configured at an interface with a specific shutdown multiplier value (NX-OS 7.0(3)I7(4) and later releases):<br>`switch(config)# int e1/36`<br>`switch(config-if)# priority-flow-control watch-dog-interval on interface-multiplier 10`<br>**Note**<br>Range of values for interface-multiplier is 1 - 10. |
| **Step 5** | **priority-flow-control watch-dog interval** *value*<br>**Example:**<br>`switch(config)# priority-flow-control watch-dog interval 200` | Specifies the watchdog interval value. The range is from 100 to 1000 milliseconds. |
| **Step 6** | **priority-flow-control watch-dog shutdown-multiplier** *multiplier*<br>**Example:**<br>`switch(config)# priority-flow-control watch-dog shutdown-multiplier 5` | Specifies when to declare the PFC queue as stuck. The range is from 1 to 10, and the default value is 1.<br>**Note**<br>When the PFC queue is declared as stuck, a syslog entry is created to record the conditions of the PFC queue. (NX-OS 7.0(3)I7(4) and later releases) |

| | Command or Action | Purpose |
|---|---|---|
| **Step 7** | (Optional) **sh queuing pfc-queue [interface] [ethernet] [detail]**<br>**Example:**<br>`switch(config)# sh queuing pfc-queue interface ethernet 1/1 detail` | Displays the PFCWD statistics. Starting with Cisco NX-OS Release 7.0(3)I6(1), you can use the detail option to account for Ingress drops.<br>`\| QOS GROUP 2 [Active] PFC [YES] PFC-COS [1]`<br>`+----------------------------------------+`<br>`\|                                        \|`<br>`Stats            \|`<br>`+----------------------------------------+`<br>`\|                            Shutdown\|`<br>`0\|`<br>`\|                            Restored\|`<br>`0\|`<br>`\|               Total bytes drained\|`<br>`0\|`<br>`\|                 Total pkts dropped\|`<br>`0\|`<br>`\|            Aggregate pkts dropped\|`<br>`0\|`<br>`\|              Ingress Pkts dropped\|`<br>`0\|`<br>`\|    Aggregate Ingress Pkts dropped\|`<br>`0\|`<br>`Global watch-dog interval [Enabled]`<br>`+----------------------------------------+`<br>`+----------------------------------------+`<br>`Global PFC watchdog configuration details`<br>`PFC watch-dog poll interval : 200 ms`<br>`PFC watch-dog shutdown multiplier : 5`<br>`PFC watch-dog auto-restore multiplier : 10`<br>`PFC watch-dog fixed-restore multiplier : 0`<br>`PFC watchdog internal-interface multiplier : 0` |
| **Step 8** | (Optional) **clear queuing pfc-queue [interface] [ethernet] [intf-name]**<br>**Example:**<br>`switch(config)# clear queuing pfc-queue interface ethernet 1/1` | Clears the environment variable PFCWD statistics. |
| **Step 9** | (Optional) **priority-flow-control recover interface [ethernet] [intf-name] [qos-group <0-7>]**<br>**Example:**<br>`switch# priority-flow-control recover interface ethernet 1/1 qos-group 3` | Recovers the interface manually. |

# Verifying the Priority Flow Control Configuration

To display the PFC configuration, perform the following task:

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal** | Enter global configuration mode. |
| **Step 2** | **show interface priority-flow-control** | Displays the status of PFC on all interfaces. |
| **Step 3** | (Optional) **show interface priority-flow-control detail** | Displays the status of PFC for each priority level for each interface. |

# Monitoring PFC Frame Counter Statistics

You can monitor the Tx and Rx counters for PFC-enabled devices either at an interface level or at a per priority (CoS) level for each interface.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **show int priority-flow-control** [**detail**] | |

### Example

This example shows how to display PFC frame counter statistics for each priority level for each interface:

```
switch# show int priority-flow-control detail


Ethernet1/1/1:
    Admin Mode: On
    Oper Mode: On
    VL bitmap: (14)
    Total Rx PFC Frames: 0
    Total Tx PFC Frames: 0
    ------------------------------------------------------------------------------
----------------------------------------
        |  Priority0  |  Priority1  |  Priority2  |  Priority3  |  Priority4  |
 Priority5  |  Priority6  |  Priority7  |
    ------------------------------------------------------------------------------
----------------------------------------
    Rx  |0            |0            |0            |0            |0            |0
            |0            |0
    ------------------------------------------------------------------------------
----------------------------------------
    Tx  |0            |0            |0            |0            |0            |0
            |0            |0
```

```
Ethernet1/1/2:
    Admin Mode: Auto
    Oper Mode: Off
    VL bitmap:
    Total Rx PFC Frames: 0
    Total Tx PFC Frames: 0
    ------------------------------------------------------------------------------
----------------------------------------
        |  Priority0  |  Priority1  |  Priority2  |  Priority3  |  Priority4  |
 Priority5  |  Priority6  |  Priority7  |
    ------------------------------------------------------------------------------
----------------------------------------
    Rx  |0            |0            |0            |0            |0            |0
            |0            |0
    ------------------------------------------------------------------------------
----------------------------------------
    Tx  |0            |0            |0            |0            |0            |0
            |0            |0
```

This example shows how to display PFC frame counter statistics for each interface:

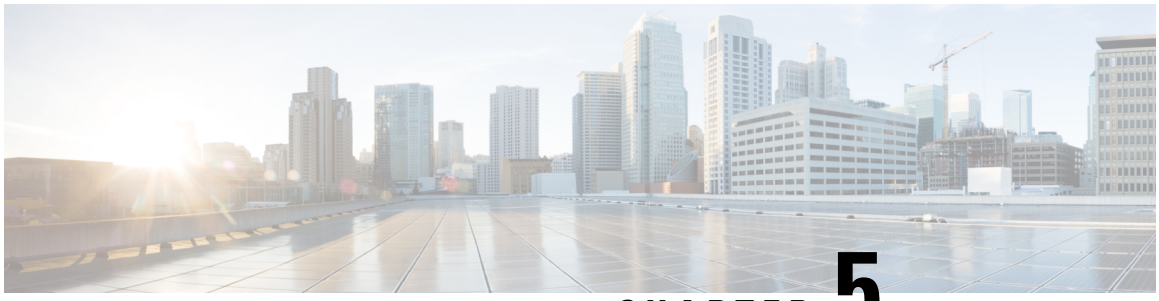
```
switch# show int priority-flow-control
===============================================================
Port               Mode Oper(VL bmap)  RxPPP       TxPPP
===============================================================

Ethernet1/1/1      On   On  (14)      0           0
Ethernet1/1/2      Auto Off           0           0
Ethernet1/1/3      Auto On  (14)      0           0
Ethernet1/15       Auto On  (14)      0           0
Ethernet1/15       Auto On  (14)      0           0
Ethernet1/15       Auto On  (14)      0           0
Ethernet1/15       Auto On  (14)      0           0
Ethernet1/25       Auto On  (14)      0           0
Ethernet1/32       On   On  (14)      0           0
switch#
```

# Configuration Examples for Priority Flow Control

The following example shows how to configure PFC.

```
switch# configure terminal
switch(config)# interface ethernet 5/5
switch(config-if)# priority-flow-control mode on
```



CHAPTER 5

# Configuring Policing


- About Policing, on page 81
- Licensing Requirements for Policing, on page 81
- Prerequisites for Policing, on page 82
- Guidelines and Limitations, on page 82
- Configuring Policing, on page 82
- Verifying the Policing Configuration, on page 89
- Configuration Examples for Policing, on page 89


## About Policing

Policing is the monitoring of the data rates for a particular class of traffic. When the data rate exceeds user-configured values, marking or dropping of packets occurs immediately. Policing does not buffer the traffic; therefore, the transmission delay is not affected. When traffic exceeds the data rate, you instruct the system to either drop the packets or mark QoS fields in them.

You can define single-rate and dual-rate policers.

Single-rate policers monitor the committed information rate (CIR) of traffic. Dual-rate policers monitor both CIR and peak information rate (PIR) of traffic. In addition, the system monitors associated burst sizes. Three colors, or conditions, are determined by the policer for each packet depending on the data rate parameters supplied: conform (green), exceed (yellow), or violate (red).

You can configure only one action for each condition. For example, you might police for traffic in a class to conform to the data rate of 256000 bits per second, with up to 200 millisecond bursts. The system would apply the conform action to traffic that falls within this rate, and it would apply the violate action to traffic that exceeds this rate.

For more information about policers, see RFC 2697 and RFC 2698.

## Licensing Requirements for Policing

The following table shows the licensing requirements for this feature:

| Product | License Requirement |
|---|---|
| Cisco NX-OS | The QoS feature does not a require license. Any feature not included in a license package is bundled with the NX-OS image and is provided at no extra charge to you. For a complete explanation of the Cisco NX-OS licensing scheme, see the *Cisco NX-OS Licensing Guide*. |

# Prerequisites for Policing

Policing has the following prerequisites:

- You must be familiar with using modular QoS CLI.
- You are logged on to the device.

# Guidelines and Limitations

Policing has the following configuration guidelines and limitations:

- Starting with Cisco NX-OS Release 7.0(3)I2(1), ingress and egress policing is supported on the Cisco Nexus 3100 platform switches. Starting with Cisco NX-OS Release 7.0(3)I5(1) ingress policing is supported on the Cisco Nexus 3000 series switches. Egress policing is not supported on Cisco Nexus 3000 series switches.
- Each module polices independently, which might affect QoS features that are being applied to traffic that is distributed across more than one module. Policers are applied to portchannel interfaces, however policers are not supported on sub-interfaces and VLANs
- All policers in the ingress direction must use the same mode.
- The **set qos-group** command can only be used in ingress policies.
- When egress RACL and egress QoS are applied together, statistics can only be enabled for one or the other, not both.
- Egress QoS policies on ALE uplink ports on top-of-rack (TOR) platforms is not supported.
- When using egress QoS, it is recommended to use appropriate match criteria to exclusively match data traffic. (Avoid match criteria like **permit ip any any**.)

# Configuring Policing

You can configure a single or dual-rate policer.

# Configuring 1-Rate and 2-Rate, 2-Color and 3-Color Policing

The type of policer created by the device is based on a combination of the **police** command arguments described in the following Arguments to the police Command table.

**Note** You must specify the identical value for **pir** and **cir** to configure 1-rate 3-color policing.

**Note** A 1-rate 2-color policer with the violate markdown action is not supported.

**Note** If the same policer enabled QoS policy is applied across multiple ingress interfaces on Cisco Nexus 3000 Series switches, the **qos qos-policies statistics** command should be enabled. Otherwise, the policer entry is shared between the interfaces that results in aggregate policing. The command, **qos qos-policies statistics** enables separate policer entries for each ingress interface and also enables policer statistics.

*Table 6: Arguments to the police Command*

| Argument | Description |
|---|---|
| **cir** | Committed information rate, or desired bandwidth, specified as a bit rate or a percentage of the link rate. Although a value for cir is required, the argument itself is optional. The range of values is from 1 to 80000000000. The range of policing values is from 8000 to 80 Gbps. |
| **percent** | Rate as a percentage of the interface rate. The range of values is from 1 to 100 percent. |
| **bc** | Indication of how much the cir can be exceeded, either as a bit rate or an amount of time at cir. The default is 200 milliseconds of traffic at the configured rate. The default data rate units are bytes. |
| **pir** | Peak information rate, specified as a PIR bit rate or a percentage of the link rate. There is no default. The range of values is from 1 to 80000000000; the range of policing values is from 8000 bps to 480 Gbps. The range of percentage values is from 1 to 100 percent. |
| **be** | Indication of how much the pir can be exceeded, either as a bit rate or an amount of time at pir. When the bc value is not specified, the default is 200 milliseconds of traffic at the configured rate. The default data rate units are bytes.<br>**Note**<br>You must specify a value for pir before the device displays this argument. |
| **conform** | Single action to take if the traffic data rate is within bounds. The basic actions are transmit or one of the set commands listed in the following Policer Actions for Conform table. The default is transmit. |
| **exceed** | Single action to take if the traffic data rate is exceeded. The basic actions are drop or markdown. The default is drop. |

| Argument | Description |
|---|---|
| **violate** | Single action to take if the traffic data rate violates the configured rate values. The basic actions are drop or markdown. The default is drop. |

Although all the arguments in the above Arguments to the police Command table are optional, you must specify a value for **cir**. In this section, **cir** indicates its value but not necessarily the keyword itself. The combination of these arguments and the resulting policer types and actions are shown in the following Policer Types and Actions from Police Arguments Present table.

***Table 7: Policer Types and Actions from Police Arguments Present***

| Police Arguments Present | Policer Type | Policer Action |
|---|---|---|
| **cir**, but not **pir**, **be**, or **violate** | 1-rate, 2-color | <= **cir**, **conform**; else **violate** |
| **cir** and **pir** | 2-rate, 3-color | <= **cir**, conform; <= **pir**, exceed; else **violate** |

The policer actions that you can specify are described in the following Policer Actions for Exceed or Violate table and the following Policer Actions for Conform table.

***Table 8: Policer Actions for Exceed or Violate***

| Action | Description |
|---|---|
| **drop** | Drops the packet. This action is available only when the packet exceeds or violates the parameters. |
| **set-cos-transmit** | Sets CoS and transmits the packet. |
| **set-dscp-transmit** | Sets DSCP and transmits the packet. |
| **set-prec-transmit** | Sets precedence and transmits the packet. |
| **set-qos-transmit** | Sets qos-group and transmits the packet. |

***Table 9: Policer Actions for Conform***

| Action | Description |
|---|---|
| **transmit** | Transmits the packet. This action is available only when the packet conforms to the parameters. |
| **set-prec-transmit** | Sets the IP precedence field to a specified value and transmits the packet. This action is available only when the packet conforms to the parameters. |
| **set-dscp-transmit** | Sets the differentiated service code point (DSCP) field to a specified value and transmits the packet. This action is available only when the packet conforms to the parameters. |

| Action | Description |
|---|---|
| **set-cos-transmit** | Sets the class of service (CoS) field to a specified value and transmits the packet. This action is available only when the packet conforms to the parameters. |
| **set-qos-transmit** | Sets the QoS group internal label to a specified value and transmits the packet. This action can be used only in input policies and is available only when the packet conforms to the parameters. |

**Note** The policer can only drop or mark down packets that exceed or violate the specified parameters. For information on marking down packets, see the Configuring Marking section.

The data rates used in the **police** command are described in the following Data Rates for the police Command table.

***Table 10: Data Rates for the police Command***

| Rate | Description |
|---|---|
| bps | Bits per second (default) |
| kbps | 1,000 bits per seconds |
| mbps | 1,000,000 bits per second |
| gbps | 1,000,000,000 bits per second |

Burst sizes used in the **police** command are described in the following Burst Sizes for the police Command table.

***Table 11: Burst Sizes for the police Command***

| Speed | Description |
|---|---|
| bytes | bytes |
| kbytes | 1,000 bytes |
| mbytes | 1,000,000 bytes |
| ms | milliseconds |
| us | microseconds |

.

## Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal**<br>**Example:**<br>`switch# configure terminal`<br>`switch(config)#` | Enters global configuration mode. |
| **Step 2** | **policy-map** [**type qos**] [**match-first**] [*policy-map-name*]<br>**Example:**<br>`switch(config)# policy-map policy1`<br>`switch(config-pmap-qos)#` | Creates or accesses the policy map named *policy-map-name* and then enters policy-map mode. The policy-map name can contain alphabetic, hyphen, or underscore characters, is case sensitive, and can be up to 40 characters. |
| **Step 3** | **class** [**type qos**] {*class-map-name* \| **class-default**} [**insert-before** *before-class-name*]<br>**Example:**<br>`switch(config-pmap-qos)# class class-default`<br>`switch(config-pmap-c-qos)#` | Creates a reference to *class-map-name* and enters policy-map class configuration mode. The class is added to the end of the policy map unless **insert-before** is used to specify the class to insert before. Use the **class-default** keyword to select all traffic that is not currently matched by classes in the policy map. |
| **Step 4** | **police** [**cir**] {*committed-rate* [*data-rate*] \| **percent** *cir-link-percent*} [**bc** *committed-burst-rate* [*link-speed*]][**pir**] {*peak-rate* [*data-rate*] \| **percent** *cir-link-percent*} [**be** *peak-burst-rate* [*link-speed*]] [**conform** {**transmit** \| **set-prec-transmit** \| **set-dscp-transmit** \| **set-cos-transmit** \| **set-qos-transmit**} [**exceed** {**drop**} [**violate** {**drop**}]]} | Polices **cir** in bits or as a percentage of the link rate. The **conform** action is taken if the data rate is <= cir. If **be** and **pir** are not specified, all other traffic takes the **violate** action. If **be** or **violate** are specified, the **exceed** action is taken if the data rate <= **pir**, and the **violate** action is taken otherwise. The actions are described in the Policer Actions for Exceed or Violate table and the Policer Actions for Conform table. The data rates and link speeds are described in the Data Rates for the police Command table and the Burst Sizes for the police Command table. |
| **Step 5** | **exit**<br>**Example:**<br>`switch(config-pmap-c-qos)# exit`<br>`switch(config-pmap-qos)#` | Exits policy-map class configuration mode and enters policy-map mode. |
| **Step 6** | **exit**<br>**Example:**<br>`switch(config-pmap-qos)# exit`<br>`switch(config)#` | Exits policy-map mode and enters global configuration mode. |

| | Command or Action | Purpose |
|---|---|---|
| **Step 7** | **show policy-map** [**type qos**] [*policy-map-name* \| **qos-dynamic**]<br>**Example:**<br>`switch(config)# show policy-map` | (Optional) Displays information about all configured policy maps or a selected policy map of type qos. |
| **Step 8** | **copy running-config startup-config**<br>**Example:**<br>`switch(config)# copy running-config startup-config` | (Optional) Saves the running configuration to the startup configuration. |

### Example

This example shows how to display the policy1 policy-map configuration:

```
switch# show policy-map policy1
```

# Configuring Ingress and Egress Policing

You can apply the policing instructions in a QoS policy map to ingress or egress packets by attaching that QoS policy map to an interface. To select ingress or egress, you specify the **input** keyword or the **output** keyword in the **service-policy** command. For more information on attaching and detaching a QoS policy action from an interface, see the Using Modular QoS CLI. section.

For egress QoS purposes, TCAM regions can be specified with the **hardware access-list tcam region** [**e-qos** | **e-qos-lite** | **e-ipv6-qos** | **e-mac-qos**] *tcam-size* command.

**Note** All TCAM regions for egress QoS purposes are double wide, however the e-qos-lite region is single wide

**Notes for Egress QoS and TCAM Regions**

- Only violated and non-violated statistics are supported for policing action when the double width TCAM is used.
- Only non-violated statistics are supported for policing action when the single width TCAM (e-qos-lite) is used.
- Statistics are disabled when the optional **no-stats** keyword is used and policies are shared (where applicable).
- The **set qos-group** command is not supported for egress QoS policies.

# Configuring Markdown Policing

Markdown policing is the setting of a QoS field in a packet when traffic exceeds or violates the policed data rates. You can configure markdown policing by using the set commands for policing action described in the Policer Actions for Exceed or Violate table and the Policer Actions for Conform table.

**Note** You must specify the identical value for **pir** and **cir** to configure 1-rate 3-color policing.

## Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | **configure terminal**<br>**Example:**<br>`switch# configure terminal`<br>`switch(config)#` | Enters global configuration mode. |
| **Step 2** | **policy-map** [**type qos**] [**match-first**] [*policy-map-name*]<br>**Example:**<br>`switch(config)# policy-map policy1`<br>`switch(config-pmap-qos)#` | Creates or accesses the policy map named *policy-map-name* and then enters policy-map mode. The policy-map name can contain alphabetic, hyphen, or underscore characters, is case sensitive, and can be up to 40 characters. |
| **Step 3** | **class** [**type qos**] {*class-name* \| **class-default**} [**insert-before** *before-class-name*]<br>**Example:**<br>`switch(config-pmap-qos)# class class-default`<br>`switch(config-pmap-c-qos)#` | Creates a reference to *class-name* and enters policy-map class configuration mode. The class is added to the end of the policy map unless **insert-before** is used to specify the class to insert before. Use the **class-default** keyword to select all traffic that is not currently matched by classes in the policy map. |
| **Step 4** | **police** [**cir**] {*committed-rate* [*data-rate*] \| **percent** *cir-link-percent*} [[**bc** \| **burst**] *burst-rate* [*link-speed*]] [[**be** \| **peak-burst**] *peak-burst-rate* [*link-speed*]] [**conform** *conform-action* [**exceed** [**violate drop set dscp dscp table** *pir-markdown-map*]]} | Polices **cir** in bits or as a percentage of the link rate. The **conform** action is taken if the data rate is <= cir. If **be** and **pir** are not specified, all other traffic takes the **violate** action. If **be** or **violate** are specified, the **exceed** action is taken if the data rate <= **pir**, and the **violate** action is taken otherwise. The actions are described in the Policer Actions for Exceed or Violate table and the Policer Actions for Conform table. The data rates and link speeds are described in the Data Rates for the police Command table and the Burst Sizes for the police Command table. |
| **Step 5** | **exit**<br>**Example:**<br>`switch(config-pmap-c-qos)# exit`<br>`switch(config-pmap-qos)#` | Exits policy-map class configuration mode and enters policy-map mode. |
| **Step 6** | **exit**<br>**Example:** | Exits policy-map mode and enters global configuration mode. |

| | Command or Action | Purpose |
|---|---|---|
| | `switch(config-pmap-qos)# exit`<br>`switch(config)#` | |
| **Step 7** | **show policy-map** [**type qos**] [*policy-map-name*]<br>**Example:**<br>`switch(config)# show policy-map` | (Optional) Displays information about all configured policy maps or a selected policy map of type qos. |
| **Step 8** | **copy running-config startup-config**<br>**Example:**<br>`switch(config)# copy running-config startup-config` | (Optional) Saves the running configuration to the startup configuration. |

# Verifying the Policing Configuration

To display the policing configuration information, perform one of the following tasks:

| Command | Purpose |
|---|---|
| **show policy-map** | Displays information about policy maps and policing. |

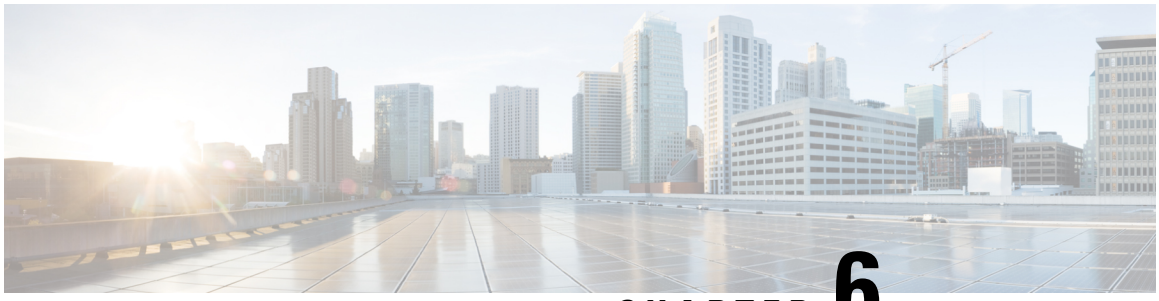
# Configuration Examples for Policing

The following example shows how to configure policing for a 1-rate, 2-color policer:

```
configure terminal
  policy-map policy1
    class one_rate_2_color_policer
      police cir 256000 conform transmit violate drop
```

The following example shows how to configure policing for a 1-rate, 2-color policer with DSCP markdown:

```
configure terminal
  policy-map policy2
    class one_rate_2_color_policer_with_dscp_markdown
      police cir 256000 conform transmit violate set-dscp-transmit 10
```



CHAPTER 6

# Configuring Traffic Shaping

This chapter contains the following sections:


- About Traffic Shaping, on page 91
- Guidelines and Limitations for Traffic Shaping, on page 92
- Configuring Traffic Shaping, on page 92
- Verifying Traffic Shaping, on page 93
- Configuration Example for Traffic Shaping, on page 93


## About Traffic Shaping

Traffic shaping allows you to control the traffic going out an interface in order to match its flow to the speed of the remote target interface and to ensure that the traffic conforms to policies contracted for it. Traffic that is adhering to a particular profile can be shaped to meet downstream requirements, which eliminates bottlenecks in topologies with data-rate mismatches.

Traffic shaping regulates and smooths out the packet flow by imposing a maximum traffic rate for each port's egress queue. Packets that exceed the threshold are placed in the queue and are transmitted later. This process is similar to traffic policing; however, the packets are not dropped. Because packets are buffered, traffic shaping minimizes packet loss (based on the queue length), which provides a better traffic behavior for TCP traffic.

Using traffic shaping, you can control access to available bandwidth, ensure that traffic conforms to the policies established for it, and regulate the flow of traffic in order to avoid congestion that can occur when the egress traffic exceeds the access speed of its remote, target interface. For example, you can control access to the bandwidth when policy dictates that the rate of a given interface should not, on average, exceed a certain rate even though the access rate exceeds the speed.

The traffic shaping rate can be configured in kilobits per second (kbps) or packets per second (PPS) and is applied to unicast queues. Queue length thresholds are configured using weighted randam early detection (WRED) configuration.

Traffic shaping can be configured at the system level or the interface level. System level queuing policies can be overridden by interface queuing policies.

# Guidelines and Limitations for Traffic Shaping

- Traffic shaping might increase the latency of packets due to queuing, because it falls back to store-and-forward mode when packets get queued.
- For traffic shaping to function properly, the store and forwarding switching mode needs to be enabled.

# Configuring Traffic Shaping

You can configure a maximum traffic rate to regulate traffic flow.

### Before you begin

Configure random-detect minimum and maximum thresholds for packets.

Configure congestion control random detection on the network QoS class map by using the **congestion-control random detect** command under the network-qos class-map.

Both QoS and network QoS policies must be applied for queuing to work. This prerequisite exists for configuring any queuing policy.

### Procedure

| | Command or Action | Purpose |
|---|---|---|
| **Step 1** | switch# **configure terminal** | Enters global configuration mode. |
| **Step 2** | switch(config)# **policy-map type queuing** *policy-name* | Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. 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-name* | Associates a class map with the policy map and enters configuration mode for the specified system class. |
| **Step 4** | switch(config-pmap-que)# **shape** {**kbps** \| **mbps** \| **gbps**} *burst size* **min** *minimum bandwidth* | Specifies the burst size and minimum guaranteed bandwidth for this queue. |
| **Step 5** | switch(config-pmap-que)# **exit** | Exits the current configuration mode. |
| **Step 6** | (Optional) switch(config)# **copy running-config startup-config** | Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration. |

### Example

This example shows how to configuring packet shaping using 200000 packets per second (pps):

```
switch# configuration terminal
class-map type qos match-all cq
  match access-group name test
class-map type queuing cqu
  match qos-group 2
policy-map type qos pq
  class cq
    set qos-group 2
policy-map type queuing pqu
  class type queuing cqu
    random-detect minimum-threshold 100 packets maximum-threshold 350 packets
switch(config)# policy-map type queuing pqu
switch(config-pmap-que)# class type queuing cqu
switch(config-pmap-que)# shape pps 200000
switch(config-pmap-que)# exit
switch(config)# copy running-config startup-config
```

# Verifying Traffic Shaping

To display traffic shaping configuration information, perform one of the following tasks:

| Command | Purpose |
|---|---|
| **show queuing** | Displays the queuing information configured for all interfaces. It includes shaper configuration information for each class, the control queue Tx and drop statistics for each port, and WRED drop packet counts. |
| **show queuing interface** *slot/port* | Displays the queuing information configured on the specified interface. |
| **show interface** *slot/port* | Shows the aggregated output traffic rate on all egress queues of the specified interface. |

# Configuration Example for Traffic Shaping

The following example shows sample configuration for the following tasks:

- Defining an ACL that matches a range of IP flows named test-qos-acl
- Marking traffic matching the test-qos-acl that ingresses on eth1/97 and eth1/99 to q2 (qos-group 2)
- Traffic shaping all traffic on q2 (qos-group 2) to a maximum of 200 mbps on egress interface eth1/100
- Applying the service-policy for qos and queuing to the respective ingress and egress interface(s)

```
ip access-list test-qos-acl
  10 permit ip 10.1.1.0/24 any
  20 permit ip 10.2.2.0/24 any
```

```
  30 permit ip 10.3.3.0/24 any

class-map type qos match-all cq
  match access-group name test-qos-acl

policy-map type qos set-queue-2
  class cq
    set qos-group 2

policy-map type queuing custom-shaper-policy
  class type queuing c-out-8q-q3
    bandwidth percent 20
  class type queuing c-out-8q-q2
    bandwidth percent 30
    shape min 200 mbps max 200 mbps
  class type queuing c-out-8q-q-default
    bandwidth percent 50
  class type queuing c-out-8q-q1
    bandwidth percent 0
  class type queuing c-out-8q-q4
    bandwidth percent 0
  class type queuing c-out-8q-q5
    bandwidth percent 0
  class type queuing c-out-8q-q6
    bandwidth percent 0
  class type queuing c-out-8q-q7
    bandwidth percent 0

interface Ethernet1/97
  service-policy type qos input set-queue-2

interface Ethernet1/99
  service-policy type qos input set-queue-2

interface Ethernet1/100
  service-policy type queuing output custom-shaper-policy
```