Quality of Service

Quality of Service

Cisco UCS provides the following methods to implement quality of service:

- System classes that specify the global configuration for certain types of traffic across the entire system
- QoS policies that assign system classes for individual vNICs
- Flow control policies that determine how uplink Ethernet ports handle pause frames

Global QoS changes made to the QoS system class may result in brief data-plane interruptions for all traffic. Some examples of such changes are:

- Changing the MTU size for an enabled class
- Changing packet drop for an enabled class
- Changing the CoS value for an enabled class

Guidelines and Limitations for Quality of Service on Cisco UCS 6400 Series Fabric Interconnects and Cisco UCS 6536 Fabric Interconnects

- Multicast optimization is not supported.
- MTU is not configurable for drop type QoS system classes and is always set to 9216. MTU is only configurable for no-drop type QoS system classes, except for the fibre channel class.
- The default MTU size for the no-drop class is 1500 and the maximum supported size for this class is 9216.
- The MTU size for fibre channel is always 2240.

- Multicast is not supported on any no-drop QoS class.

Guidelines and Limitations for Quality of Service on Cisco UCS 6300 Series Fabric Interconnect

- Cisco UCS 6300 Series Fabric Interconnect uses a shared buffer for all system classes.
- Multicast optimization is not supported.
- Multicast is not supported on any no-drop QoS class.
- When you change the QoS parameters for any class causes traffic disruption to all classes. The following table lists the changes in the QoS system class and the conditions that trigger a system reboot.

QoS System class status	Condition	FI Reboot Status
Enabled	Change between drop and no drop	Yes
No-drop	Change between enable and disable	Yes
Enable and no-drop	Change in MTU size	Yes

- The subordinate FI reboots first as a result of the change in the QoS system class. The primary FI reboots only after you acknowledge it in **Pending Activities**.

Guidelines and Limitations for Quality of Service on Cisco UCS Mini

- Cisco UCS Mini uses a shared buffer for all system classes.
- The bronze class shares the buffer with SPAN. We recommend using either SPAN or the bronze class.
- Multicast optimization is not supported.
- Multicast is not supported on any no-drop QoS class.
- Changing the QoS parameters for any class causes traffic disruption to all classes.
- When mixing Ethernet and FC or FCoE traffic, the bandwidth distribution is not equal.
- Multiple streams of traffic from the same class may not be distributed equally.
- Use the same CoS values for all no-drop policies to avoid any FC or FCoE performance issues.
- Only the platinum and gold classes support no-drop policies.

Configuring System Classes

System Classes

Cisco UCS uses Data Center Ethernet (DCE) to handle all traffic inside a Cisco UCS domain. This industry standard enhancement to Ethernet divides the bandwidth of the Ethernet pipe into eight virtual lanes. Two virtual lanes are reserved for internal system and management traffic. You can configure quality of service

(QoS) for the other six virtual lanes. System classes determine how the DCE bandwidth in these six virtual lanes is allocated across the entire Cisco UCS domain.

Each system class reserves a specific segment of the bandwidth for a specific type of traffic, which provides a level of traffic management, even in an oversubscribed system. For example, you can configure the **Fibre Channel Priority** system class to determine the percentage of DCE bandwidth allocated to FCoE traffic.

The following table describes the system classes that you can configure.

Table 1: System Classes

System Class	Description
Platinum Gold Silver Bronze	A configurable set of system classes that you can include in the QoS policy for a service profile. Each system class manages one lane of traffic. All properties of these system classes are available for you to assign custom settings and policies. For Cisco UCS Mini, packet drop can only be disabled on the platinum and gold classes. Only one platinum and one gold class can be configured as a no drop class at a time.
Best Effort	A system class that sets the quality of service for the lane reserved for basic Ethernet traffic. Some properties of this system class are preset and cannot be modified. For example, this class has a drop policy that allows it to drop data packets if required. You cannot disable this system class.
Fibre Channel	A system class that sets the quality of service for the lane reserved for Fibre Channel over Ethernet traffic. Some properties of this system class are preset and cannot be modified. For example, this class has a no-drop policy that ensures it never drops data packets. You cannot disable this system class. **Note** FCoE traffic has a reserved QoS system class that should not be used by any other type of traffic. If any other type of traffic has a CoS value that is used by FCoE, the value is remarked to 0.

Configuring QoS System Classes

Note

- Under the network QoS policy, the MTU is used only for buffer carving when no-drop classes are configured. No additional MTU adjustments are required under the network QoS policy to support jumbo MTU.
- For Cisco VIC 1400 series, 14000 series and later version adapters, you can change the MTU size of the vNIC from the host interface settings. When the Overlay network is configured, make sure that the overall MTU size does not exceed the MTU value in the QoS system class. If this MTU value exceeds the MTU value in the QoS system class, packets could be dropped during data transmission.

Important Use the same CoS (Class of Service) values on UCS and N5K/N9K for all the no-drop policies. To insure that end-to-end PFC works correctly, have the same QoS policy configured on all intermediate switches.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **LAN** > **LAN Cloud**.

Step 3 Select the **QoS System Class** node. Packet drop should be unchecked to configure MTU.

MTU is not configurable for drop type QoS system classes and is always set to 9216. MTU is only configurable for no-drop type QoS system classes

Step 4 In the **Work** pane, click the **General** tab.

Step 5 Update the properties for the system class that you want to configure to meet the traffic management needs of the system.

Note Some properties may not be configurable for all system classes. The maximum value for MTU is 9216.

Step 6 Click **Save Changes**.

Enabling a QoS System Class

The Best Effort or Fibre Channel system classes are enabled by default.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **LAN** > **LAN Cloud**.

Step 3 Select the **QoS System Class** node.

Step 4 In the **Work** pane, click the **General** tab.

Step 5 Check the **Enabled** check box for the QoS system that you want to enable.

Step 6 Click **Save Changes**.

Disabling a QoS System Class

You cannot disable the Best Effort or Fibre Channel system classes.

All QoS policies that are associated with a disabled system class default to Best Effort or, if the disabled system class is configured with a Cos of 0, to the Cos 0 system class.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.
Step 2 Expand **LAN** > **LAN Cloud**.
Step 3 Select the **QoS System Class** node.
Step 4 In the **Work** pane, click the **General** tab.
Step 5 Uncheck the **Enabled** check box for the QoS system that you want to disable.
Step 6 Click **Save Changes**.

Configuring Quality of Service Policies

Quality of Service Policy

A quality of service (QoS) policy assigns a system class to the outgoing traffic for a vNIC or vHBA. This system class determines the quality of service for that traffic. For certain adapters, you can also specify additional controls on the outgoing traffic, such as burst and rate.

You must include a QoS policy in a vNIC policy or vHBA policy and then include that policy in a service profile to configure the vNIC or vHBA.

Creating a QoS Policy

Procedure

Step 1 In the **Navigation** pane, click **LAN**.
Step 2 Expand **LAN** > **Policies**.
Step 3 Expand the node for the organization where you want to create the pool.

If the system does not include multi tenancy, expand the **root** node.

Step 4 Right-click **QoS Policy** and select **Create QoS Policy**.
Step 5 In the **Create QoS Policy** dialog box, complete the required fields.
Step 6 Click **OK**.

What to do next

Include the QoS policy in a vNIC or vHBA template.

Deleting a QoS Policy

If you delete a QoS policy that is in use or you disable a system class that is used in a QoS policy, any vNIC or vHBA that uses that QoS policy is assigned to the Best Effort system class or to the system class with a CoS of 0. In a system that implements multitenancy, Cisco UCS Manager first attempts to find a matching QoS policy in the organization hierarchy.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **Servers** > **Policies** > ***Organization_Name***.

Step 3 Expand the **QoS Policies** node.

Step 4 Right-click the QoS policy you want to delete and select **Delete**.

Step 5 If a confirmation dialog box displays, click **Yes**.

Configuring Flow Control Policies

Flow Control Policy

Flow control policies determine whether the uplink Ethernet ports in a Cisco UCS domain send and receive IEEE 802.3x pause frames when the receive buffer for a port fills. These pause frames request that the transmitting port stop sending data for a few milliseconds until the buffer clears.

For flow control to work between a LAN port and an uplink Ethernet port, you must enable the corresponding receive and send flow control parameters for both ports. For Cisco UCS, the flow control policies configure these parameters.

When you enable the send function, the uplink Ethernet port sends a pause request to the network port if the incoming packet rate becomes too high. The pause remains in effect for a few milliseconds before traffic is reset to normal levels. If you enable the receive function, the uplink Ethernet port honors all pause requests from the network port. All traffic is halted on that uplink port until the network port cancels the pause request.

Because you assign the flow control policy to the port, changes to the policy have an immediate effect on how the port reacts to a pause frame or a full receive buffer.

Creating a Flow Control Policy

Before you begin

Configure the network port with the corresponding setting for the flow control that you need. For example, if you enable the send setting for flow-control pause frames in the policy, ensure that the receive parameter in the network port is set to on or to desired. If you want the Cisco UCS port to receive flow-control frames, ensure that the send parameter is set to on or to desire on the network port. If you do not want to use flow control, you can set the send and receive parameters on the network port to off.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **LAN** > **Policies**.

Step 3 Expand the **root** node.

You can only create a flow control policy in the root organization. You cannot create a flow control policy in a sub-organization.

Step 4 Right-click the **Flow Control Policies** node and select **Create Flow Control Policy**.

Step 5 In the **Create Flow Control Policy** wizard, complete the required fields.

Step 6 Click **OK**.

What to do next

Associate the flow control policy with an uplink Ethernet port or port channel.

Deleting a Flow Control Policy

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **LAN** > **Policies** > ***Organization_Name***.

Step 3 Expand the **Flow Control Policies** node.

Step 4 Right-click the policy you want to delete and select **Delete**.

Step 5 If a confirmation dialog box displays, click **Yes**.

Configuring Slow Drain

QoS Slow Drain Device Detection and Mitigation

All data traffic between end devices in the fabric is carried by Fibre Channel services that use link-level, per-hop-based, and buffer-to-buffer flow control. These classes of service do not support end-to-end flow control. When slow devices are attached to the fabric, the end devices do not accept the frames at the configured or negotiated rate. The slow devices lead to an Inter-Switch Link (ISL) credit shortage in the traffic that is destined for these devices, and they congest the links. The credit shortage affects the unrelated flows in the fabric that use the same ISL link even though destination devices do not experience a slow drain.

Similarly, in End-Host Mode, if a server that is directly attached to the Fabric Interconnect receives traffic slowly, it may congest the uplink port shared by other servers. If a slow server is attached to a HIF port on FEX/IOM, it may congest the fabric port and/or uplink port.

Cisco UCS Manager Release 4.0(2) introduces the QoS Slow Drain Detection and Mitigation feature on Cisco UCS 6454 Fabric Interconnects. This feature provides various enhancements that enable you to detect slow drain devices that cause congestion in the network, and also mitigate it. The enhancements are mainly on the edge ports and core ports that connect to the slow drain devices. This is done to minimize the frames stuck condition in the edge and core ports due to slow drain devices that are causing an ISL blockage. To avoid or minimize the stuck condition, you can configure smaller frame timeout for the ports. A smaller frame timeout value helps to alleviate the slow drain condition that affects the fabric by dropping the packets on the edge ports sooner than the time they actually get timed out. This function frees the buffer space in ISL, which can be used by other unrelated flows that do not experience the slow drain condition. Cisco UCS Manager Release 4.1 extends support of this feature to Cisco UCS 64108 Fabric Interconnects.

Note Another way of mitigating network congestion is to use the watchdog timer function, supported on Cisco UCS 6400 Series Fabric Interconnects starting with Cisco UCS Manager 4.2. However, the slow drain and watchdog timer functions are mutually exclusive.

In this release, slow drain detection and mitigation is supported on the following ports:

- FCoE
- Back-plane

Configuring Slow Drain

While configuring slow drain timeout timers, you can select the timeout value from the list of allowed values. You cannot configure custom timeout values.

Note

- In Cisco UCS Manager Release 4.1(3a), the slow drain timeout timer is enabled on the FCoE port by default with a 500ms timeout value. Starting with Cisco UCS Manager Release 4.2(1), the slow drain timeout timer is disabled and the watchdog timer is enabled by default.
- We recommend to change the default timeout values of the Core FCoE port and Edge FCoE Port only when the current default timeout values result in the dropping of packets during high traffic load.
- Slow drain and the watchdog timer cannot be used simultaneously. Attempting to do so will result in an error.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **LAN** > **LAN Cloud**.

Step 3 Starting in UCS Manager 4.2.1 click on the **QoS System Class** entry. in earlier releases, go to the **QoS** entry and click to go to the **Work** pane.

Step 4 In the **Work** pane, click the **QoS** tab.

Step 5 Click **Configure Slow Drain**.

Step 6 In the **Configure Slow Drain Timers** dialog box that appears, configure the following fields:

Name	Description
FCoE Port radio button	Whether slow drain timers are enabled on FCoE ports. • **Disabled**—Configuration of slow drain timers is disabled. This is the default option till the previous version of Cisco UCS Manager Release 4.1(3a). • **Enabled**—Configuration of slow drain timers is enabled. Starting with Cisco UCS Manager Release 4.1(3a), this is the default option.
Core FCoE port (ms) drop-down list	The time in milliseconds (ms) after which frames timeout on core FCoE ports. You can select from the following values: • 100 • 200 • 300 • 400 • 500—This is the default value • 600 • 700 • 800 • 900 • 1000

Name	Description
Edge FCoE Port (ms) drop-down list	The time in milliseconds (ms) after which frames timeout on edge FCoE ports. You can select from the following values: • 100 • 200 • 300 • 400 • 500—This is the default value • 600 • 700 • 800 • 900 • 1000

Step 7 Click **OK**.

Step 8 Click **Save Changes**.

Correcting a Slow Drain Condition

Correcting a slow drain condition will work only on those ports that are designated to be in the 'error-disabled' state because of 'slow-drain'.

Procedure

Step 1 In the **Navigation** pane, click **Equipment**.

Step 2 Expand **Equipment** > **Chassis** > ***Chassis Number*** > **IO Modules**.

Step 3 Choose the I/O module on which you want to recover backplane ports that are in the **error-disabled** state.

Step 4 In the **Work** pane, click the **General** tab.

Step 5 In the **Actions** area, click **Correct Slow Drain Condition**.

Step 6 If a confirmation dialog box displays, click **Yes**.

Configuring the Watchdog Timer

The Watchdog Timer

Cisco UCS Manager 4.2 supports the watchdog timer function on Cisco 6400 Series Fabric Interconnects with switch ports that are PFC mode enabled.

Priority flow control, (PFC) 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. A PFC storm could occur in the network due to a malfunctioning NIC or switch, causing the PFC frames to be propagated to all senders. This could cause a complete stall in network traffic. To mitigate a PFC storm, you can use a PFC watchdog timer. Configure a PFC watchdog interval to detect whether packets in a no-drop queue are being drained within a specified time period. If packets are present in buffer longer than the configured time period, once the time period is exceeded, all outgoing packets are dropped on the interfaces that match the PFC queue that is not being drained.

Use of watchdog timer and slow drain on are mutually exclusive options. You can enable either the slow drain or watchdog timer, but not both. Starting with Cisco UCS Manager Release 4.2(1), the slow drain timeout timer is disabled and the watchdog timer is enabled by default. The default watchdog interval is 500 with a shutdown multiplier of 1.

Configuring the Watchdog Timer

The watchdog timer cannot be used in connection with slow drain.

Before you begin

If slow drain was used previously, disable slow drain.

Procedure

Step 1 In the **Navigation** pane, click **LAN**.

Step 2 Expand **LAN** > **LAN Cloud**.

Step 3 Select a fabric interconnect and click on **QoS System Class**.

Step 4 In the **General** tab, click on **Configure Watchdog Timers**.

Step 5 In the **Configure WD Timers** window, click **On** for the **WD admin state** to globally enable the PFC watchdog interval for all interfaces, then select a watchdog interval betweem 100 and 1000 milliseconds and a shutdown multiplier between 1 and 10.

The default watchdog interval is 500 and the default shutdown multiplier is 1.

Step 6 Click **OK**.

Step 7 Click **Save Changes**.