QTECH
МИР ДОСТУПНЕЕ

РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

Data Center Configuration

www.qtech.ru

Оглавление

1. ONFIGURING DCB

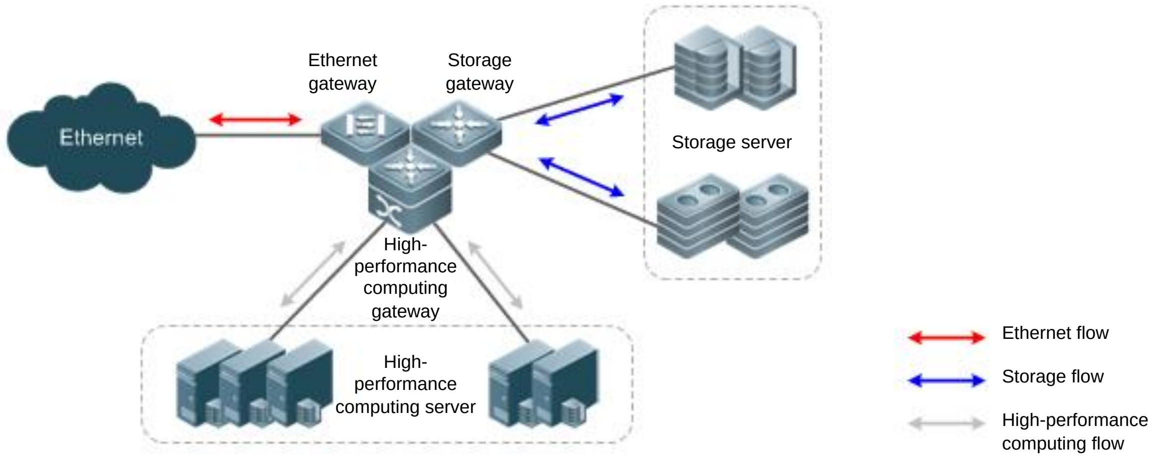
1.1. Overview

A data center (DC) mainly processes the following three traffic types:

- Storage traffic: The requirement of packet integrity must be met.
- High-performance computing traffic: The requirement of low delay must be met.
- Ethernet traffic: Certain degree of packet loss and delay are allowed.

Due to the different requirements for the three traffic types, traditional DC bears the traffic on three types of networks, as shown in Figure 1-1.

Figure 1-1 Three types of DC networks



This topology configuration is feasible for a small-sized DC. If the DC grows, running the three types of networks becomes costly. An alternative method is to bear the three traffic types on Ethernet and define additional mechanisms, so that the Ethernet can meet the requirements of the three types of networks. The following definitions are formulated in IEEE to implement the method:

- Priority-based Flow Control (PFC): a mechanism that avoids packet loss in storage traffic and impact on other traffic types if the three traffic types coexist on Ethernet.
- Data Center Bridging Capability Exchange Protocol (DCBX): a protocol that is defined by IEEE 802.1Qaz for negotiating the data center bridge (DCB) capability. With the Link Layer Discovery Protocol (LLDP), DCBX discovers devices and exchanges DCB capability.

Protocols and Standards

- IEEE Std 802.1Qaz
- IEEE 802.1Qbb
- IEEE 802.1Qau
- IEEE 802.3bd
- http://www.ieee802.org/1/files/public/docs2008/az-wadekar-dcbx-capability-exchange-discovery-protocol-1108-v1.01.pdf

1.2. Applications

Application	Description

Configuring DC Network	The Ethernet of a DC bears three types of data traffic.

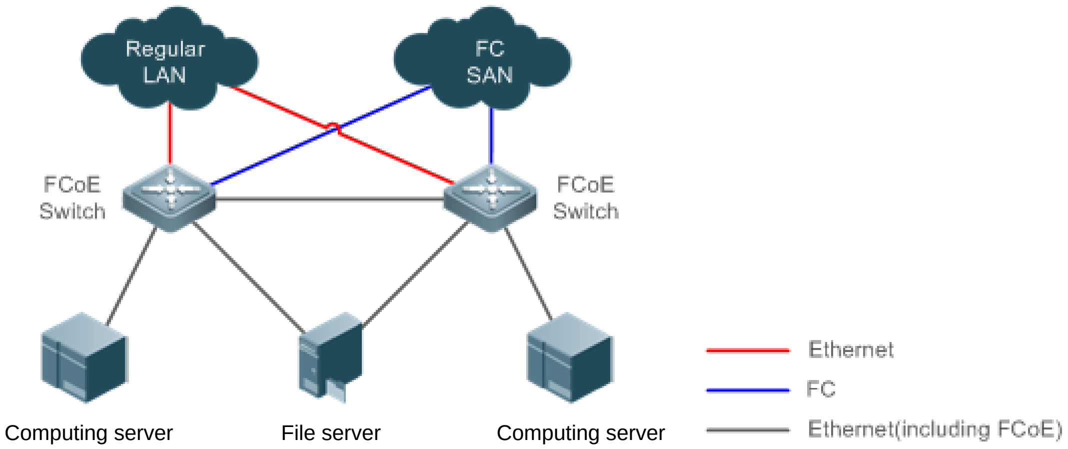
1.1.2. Configuring DC Network

Scenario

The DC Ethernet bearing three types of data traffic must meet the following requirements:

- Storage traffic: The requirement of packet integrity must be met.
- High-performance computing traffic: The requirement of low delay must be met.
- Ethernet traffic: Certain degree of packet loss and delay are allowed.

Figure 1-2



Deployment

- Enable PFC for storage traffic to meet the requirement of packet integrity.
- Map the packet priority of high-performance computing traffic to the traffic class (TC) which boasts strict priority (SP) algorithm.
- Allocate bandwidth to the TC of Ethernet traffic.

1.3. Features

Basic Concepts

802.1p (Cos)

In 802.1p, the L2 header Virtual Local Area Network (VLAN) tag field falls into 8 priorities, ranging from 0 to 7.

16bit	3bit	1bit	12bit
Tag Protocol ID (0x8100)	priority	CFI	VLAN ID

Traffic with priorities in the range 0~2 indicates common Ethernet traffic that is transmitted with maximum capability; traffic with priority 3 indicates Fiber Channel Over Ethernet (FCoE) Storage Area Network (SAN) traffic; traffic with priorities 4 and 5 indicates delay-sensitive voice traffic; and traffic with priorities 6 and 7 indicates delay-sensitive main traffic. Traffic with priority 7 indicates network control packets.

Trust edge and application priority identification

By default, Ethernet ports are untrusted ports and need to be set to trust mode, so that packets carrying 802.1p value tags can be mapped to corresponding TC. The priority of untagged packets is 0 by default and can be modified.

The identified priority of FCoE packets is 3 by default and can be modified. The priority identification function must be enabled on a network interface card (NIC) which supports DCBX. If the NIC does not support DCBX, the priority identification function will be ignored.

If NICs do not support DCBX or users require more detailed priority identification function, configure Quality of Service (QoS) policies on ports to map an access control list (ACL) and identify application priority; or modify the FCoE packet priority on the servers directly connected to switches.

Queue and TC

A switch port has 3~8 queues for caching packets. A traffic class (TC) can be a single queue caching packets with different priorities or composed of multiple queues, with each queue caching packets of specified priority. The SP algorithm can be performed for a TC.

For example, TC 1 contains queue 1 only, and queue 1 caches packets with priorities 0~2; or TC 1 contains queues 0~2, where, queue 0 caches packets with priority 0, queue 1 caches packets with priority 1, and queue 2 caches packets with priority 2. Meanwhile, the SP algorithm may be enabled for queues 0~2 to ensure that packets with priority 2 are transmitted first.

Mapping list

Packets are mapped to a TC in the mapping list according to identified priority and then allocated to a specific queue of the TC.

- A TC is composed of multiple queues.

Each packet with a priority is mapped to a specific queue of the TC. The following table lists the mapping relationship between priorities and queues.

Cos	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Queue	0	1	2	3	4	5	6	7

Queues in a TC are associated with the priorities mapped. According to the mapping list, users can modify queues contained in a TC. For example, users can map priorities 0~2 to TC 1, then TC 1 contains queues 0~2 that are associative with priorities 0~2.

- A TC contains a single queue.

A TC is associative with a specific queue. The following table lists the mapping relationship between queues and TCs.

Queue	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
TC	0	1	2	3	4	5	6	7

According to the mapping list, the mapping relationship between priorities and TCs is the mapping relationship between priorities and queues. For example, if priority 3 is mapped to TC 1, packets with priority 3 will be mapping to queue 1 of TC 1.

Overview

Feature	Description
PFC	Implements PFC on packets based on 802.1p value. If traffic of specified priority becomes congested, stops forwarding packets of this priority and instructs the transmitting device to stop sending such packets. With this function, packet integrity of Ethernet can be ensured.
DCBX	The DCB device on the edge access layer instructs the servers connected to automatically synchronize with the DCB device.

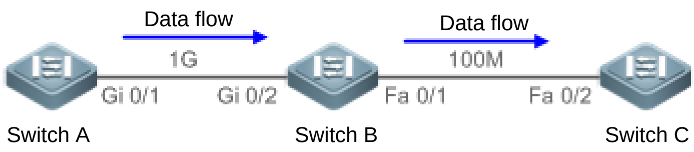
1.1.3. Configuring PFC

With PFC, packet integrity of the storage traffic can be ensured and impact on other traffic types can be avoided if the three traffic types, storage traffic, high-performance computing traffic, and Ethernet traffic, coexists on Ethernet. An 802.1p value is allocated to the storage traffic, and the PFC function is enabled for the priority. Other 802.1p values are allocated to the high-performance computing traffic and Ethernet traffic and the PFC function is disabled.

Working Principle

PFC is an enhanced version of the flow control function. Generally, if a receiver is unable to process received packets, the receiver instructs the packet sender to stop sending packets, which avoids packet loss. The flow control function is implemented by receiving or sending PAUSE frames via ports.

Figure 1-3 Flow control

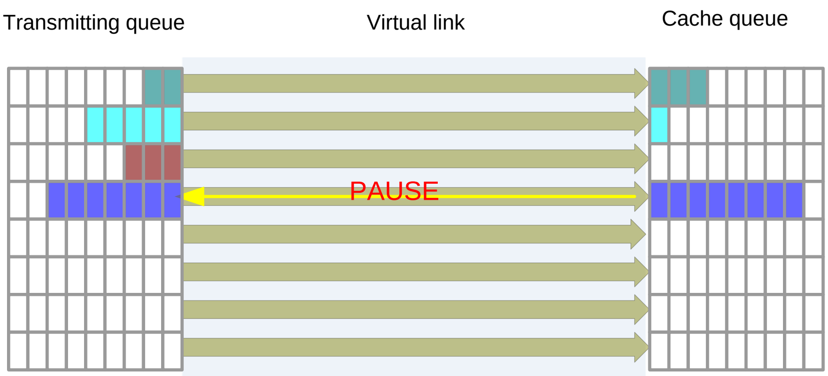


As shown in Figure 1-3, when ports Gi 0/1 and Gi 0/2 forward packets at 1 Gbps, port Fa 0/1 will be congested. To avoid packet loss, enable the flow control function on ports Gi 0/1 and Gi 0/2.

- If port Fa 0/1 is congested in the course of forwarding packets, switch B will cache the packets. When the traffic exceeds the forwarding capability of switch B, packets will be lost. In this case, port Gi 0/2 will send a PAUSE frame to port Gi 0/1, instructing port Gi 0/1 to stop forwarding packets.
- Upon receiving the PAUSE frame, port Gi 0/1 stops forwarding packets to port Gi 0/2. The pause time is subject to the information carried in the PAUSE frame. If the congestion problem persists after the pause time is up, the preceding procedure will repeat until the congestion problem is cleared.

As described above, the flow control function can avoid packet loss, but it will result in link stop as well. That is, the entire link is stopped. The PFC function, however, allows eight virtual links to be created over an Ethernet link and specifies an 802.1p value for each virtual link. Any virtual link can be paused or restarted independently, without causing traffic interruption to other virtual links.

Figure 1-4 PFC



As shown in the preceding figure, enable the PFC function on ports Gi 0/1 and Gi 0/2 for priority 3.

- If port Fa 0/1 is congested in the course of forwarding packets with priority 3, switch B will cache the packets. When the traffic exceeds the forwarding capability of switch B, port Gi 0/2 will send a PFC PAUSE frame to port Gi 0/1, instructing port Gi 0/1 to stop forwarding packets with 802.1p value 3.
- Upon receiving the PFC PAUSE frame, port Gi 0/1 stops forwarding packets to port Gi 0/2. The pause time is subject to the information carried in the PFC PAUSE frame. If the congestion problem persists after the pause time is up, the preceding procedure will repeat until the congestion problem is cleared. As shown in the preceding figure, stopping forwarding packets with priority 3 will not have impact on packet transmission of other links between Gi 0/1 and Gi 0/2.

Related Configuration

Enabling the PFC Function

The PFC function is disabled for any priority by default.

Run the **priority-flow-control nodrop** *cos-value-list* {**off** | **on**} command to enable the PFC function for a specified priority. *cos-value-list* indicates a priority list (range: 0~7); **on** indicates that the PFC function is enabled; and **off** indicates that the PFC function is disabled.

Configuring Flow Rate Warning for Pause Frames for a Priority of an Interface

By default, no warning is generated when a device interface sends or receives Pause frames.

Run the **priority-flow-control early-warning** *cos-value-list* **{ output pps** *rate-num*] **| input pps** *rate-num***}** command to configure the flow rate warning thresholds for Pause frames in the input and output directions for a specified PFC priority of an interface. *cos-value-list* indicates the configured priority list, and a priority ranges from **0** to **7**. **output pps** *rate-num* indicates the threshold of Pause frames in the output direction while **input pps** *rate-num* indicates the threshold of Pause frames in the input direction.

If the rate for an interface at a specified priority to send/receive Pause frames exceeds the configured threshold (packets per second), a warning is generated and logged.

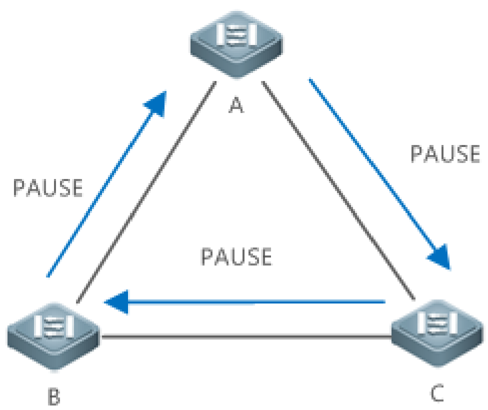
1.3.1.1. PFC Deadlock Detection

After the PFC priority is configured on the switch, the Ethernet is apt to encounter deadlock if a link is congested. This is called PFC deadlock.

Working Principle

The figure below shows a typical scenario, in which devices in a data center have cyclic buffer areas that are mutually dependent. In the ring topology, any device may affect the buffer areas of the uplink device, and wait for its downlink device to release some buffer areas to recover the data packet transmission (dependency relationship). When Device A, Device B, and Device C reach the XOFF waterline and transmit PFC frames to the peer device at the same time, all switches in the topology are in the flow pause state, that is, deadlock occurs. The throughput of the entire network or part of the network becomes zero due to the back pressure effect of PFC Pause frames. A lot of facts show that deadlock surely occurs in such networks. Deadlock may also occur when a transient loop forms in a closed-loop network. In a data center, when a link fault occurs, a transient loop arises during the rerouting of the Border Gateway Protocol (BGP, an autonomous system routing protocol running over TCP); a transient loop may also appear during the update of the SDN-based data center. Though the transient loop disappears rapidly, the deadlock incurred persists. The system cannot automatically clear deadlock even after the error cause (for example, configuration error or link fault/update) is eliminated.

Figure 1-5 Deadlock Principle



Remarks	Device A, Device B, and Device C are switches.

A buffer time can be set for PFC priority queues of a device interface to eliminate the deadlock. When a queue is congested and PFC Pause frames are received, a few time slices are still allocated for packet processing, to recover from the deadlock state. When it is detected that an interface of a PFC priority enters the deadlock state frequently, the PFC function of the priority can be disabled to avoid impact on the normal forwarding of packets.

Related Configuration

Configuring the PFC Deadlock Detection Time and Packet Processing Time for a Specified Priority

The PFC deadlock detection function is disabled for all priorities by default.

Run the **priority-flow-control deadlock cos-value** *cos-value-list* **detect** *detect-time* **recover** *recover-time* command to configure the PFC deadlock detection time and packet processing time for a specified priority.

Enabling the PFC Deadlock Detection Function for a Priority of an Interface

The PFC deadlock detection function is disabled for all priorities of all interfaces by default.

When the **priority-flow-control deadlock cos-value detect recover** command is executed to configure the global PFC deadlock detection function for a priority, you can run the **priority-flow-control deadlock cos-value** *cos-value-list* **enable** command to enable the PFC deadlock detection function for the specified priority of an interface.
After the PFC deadlock detection function is enabled for a specified priority of an interface, the interface will detect PFC deadlock events at this priority of the interface. If a PFC deadlock event occurs, the interface processes packets according to the configured policy. The **priority-flow-control deadlock drop** command is used to discard packets; if this command is not configured, packets are allowed to pass.

Configuring the PFC Deadlock Threshold Function

The PFC deadlock threshold function is disabled for all priorities by default.

After the **priority-flow-control deadlock cos-value detect recover** command is executed to configure the global PFC deadlock detection function for a priority, you can run the **priority-flow-control deadlock limit cos-value** *cos-value-list* **enable** command to enable the PFC deadlock threshold function for the specified priority.
After the PFC deadlock threshold function is enabled for a priority, if the number of deadlock times occurring for the priority within a detection period (1 minute by default) reaches the threshold (10 times by default), the PFC function is disabled for the priority, and the event and the disabling of the PFC function are logged (you can run the **show priority-flow-control deadlock status** command to display the logs).

2.1.3. Configuring DCBX

Working Principle

DCBX version

Currently, market products support the following DCBX versions:

- CIN-DCBX: indicates the first-generation of DCBX designed by Cisco, Intel, and Nuova that is compatible with the converged network adapter (CNA).
- CEE-DCBX: indicates the second-generation of DCBX that is compatible with the CAN.
- IEEE-DCBX: indicates a standard DCBX that is defined by IEEE 802.1Qaz.

Negotiating DCB capability

DCBX encapsulates each DCB capability into an LLDP type/length/value (TLV). Based on the LLDP TLV, peer devices discover the DCB capabilities of neighboring devices and then negotiate with the neighboring devices according to the discovered DCB capabilities. By the negotiation, DCB configuration can be synchronized and error detection can be implemented at both ends.

DCB capabilities that need to be negotiated include PFC, and APP. The Willing field is contained in corresponding TLV of the capabilities. When the Willing field is 1, peer DCB configuration is acceptable; when the Willing field is 0, peer DCB configuration is unacceptable. DCB capabilities are negotiated based on the Willing field. The following table lists the values of the Willing field at both ends and negotiation results.

Local Willing Field	Peer Willing Field	Negotiation Result
1	1	DCB configuration of the device with smaller media access control (MAC) address is selected as a common DCB configuration.
1	0	Peer DCB configuration is selected.
0	1	Local DCB configuration is selected.
0	0	If DCB configuration at both ends is the same, negotiation succeeds; otherwise, negotiation fails.

Related Configuration

The DCBX function is enabled by default, so DCB capability configuration can be performed without additional configuration.

1.4. Configuration

Configuration	Description and Command	
Configuring Basic Functions of the PFC	(Mandatory) It is used to enable the PFC capability.	
	priority-flow-control nodrop *cos-value-list* { **off** | **on**}	Configures the PFC function for specified priorities.
Configuring Basic Functions of the DCBX	(Mandatory) It is used to discover, detect, and negotiate peer DCB capabilities.	
	dcbx enable	Enables the global and interface DCBX function.
	dcbx mode { **auto** | **cee** | **ieee** }	Configures the DCBX mode of interfaces.
	dcbx { **pfc** | **app-proto** } **advertise**	Configures the DCBX TLV type that can be released on ports.
	dcbx { **pfc** | **app-proto** } **willing enable**	Enables DCB capability synchronization for ports.

- Remark: Configuration commands vary with specific products. For details about the commands, refer to DCB Commands.

1.1.4. Configuring Basic Functions of the PFC

Configuration Effect

- Implement PFC for port data traffic based on 802.1p values.

Notes

- PFC is applicable to point-to-point (P2P) full-duplex links only.

- The PFC function conflicts with ordinary flow control function.
- If the priorities with PFC enabled and the priorities with PFC disabled are mapped to a queue, packets with the latter priorities will be suppressed. Generally, it is not recommended to map such priorities to a same queue.
- The PFC function is not applicable to packets sent by a CPU.
- The PFC function can take effect after a port trusts the CoS values.
- Priorities with the PFC function enabled cannot be mapped to a TC or queue that contains any other priorities (including priorities with the PFC function enabled). In another word, priorities mapped to a TC or queue cannot have the PFC function enabled.

Configuration Steps

Enabling the PFC function

- The PFC function is disabled for packets with any priority by default. To implement PFC on specified priority, configure PFC.

Verification

- Check the display result.

Related Commands

Enabling the PFC function

Command	priority-flow-control nodrop ***cos-value-list*** {off **|** on}
Parameter Description	*cos-value-list*: Indicates a priority list (range: 0~7). **on**: Enables the PFC function of a priority. **off**: Disables the PFC function of a priority.
Command Mode	Global configuration mode
Usage Guide	N/A

Command	priority-flow-control nodrop ***cos-value-list*** {off **|** on}
Parameter Description	*cos-value-list*: Indicates a priority list (range: 0~7). **on**: Enables the PFC function of a priority. **off**: Disables the PFC function of a priority.
Command Mode	Interface configuration mode
Usage Guide	The PFC function can be automatically negotiated and configured based on DCBX. After the PFC function is configured on the CLI, configuration results fall into the following scenarios: 1) The PFC function is automatically negotiated and configured based on DCBX. The automatic configuration takes effect. 2) The PFC function is not automatically negotiated and configured based on DCBX. The configuration on the CLI takes effect.

Configuring Flow Rate Warning for a Priority of an Interface

Command	**priority-flow-control early-warning** *cos-value-list* **{ output pps** *rate-num*] **| input pps** *rate-num***}**
Parameter Description	*cos-value-list*: Indicates the configured priority list. The priority ranges from **0** to **7**. *rate-num*: Indicates the flow rate warning value. The value ranges from **1** to **148809523**.
Command Mode	Interface configuration mode
Usage Guide	N/A

Configuring the Global PFC Statistics Collection Interval

Command	**priority-flow-control counting-interval** *number*
Parameter Description	*number*: Indicates the statistics collection interval in seconds. The value ranges from **1** to **20**.
Command Mode	Global configuration mode
Usage Guide	N/A

Configuration Example

Enabling PFC for 802.1p value 3

Scenario Figure 1-6	
Configuration Steps	● Specify PFC IEEE 802.1p value 3 on a server. ● Specify PFC IEEE 802.1p value 3 on a switch.
A	QTECH(config)# priority-flow-control nodrop 3 on
Verification	Check the display result.
A	QTECH# show priority-flow-control status PFC Global PriorityEnablePFC: 3 QTECH# show priority-flow-control status interface FastEthernet 0/1 Interface PriorityEnabledPFC PriorityEnalbedPFCByUser ---------------- ----------------- -------------------------- FastEthernet 0/1 3 QTECH# show priority-flow-control statistics interface FastEthernet 0/1 interface Priority PauseSend PauseSend-Rate PauseReceived PauseReceived-Rate (packets/sec) (packets/sec) ---------- -------- ----------- ---------------- ------------- ------------------ Fa 0/1 0 0 0 0 0 Fa 0/1 1 0 0 0 0

	Fa 0/1 2 0 0 0 0 Fa 0/1 3 0 0 0 0 Fa 0/1 4 0 0 0 0 Fa 0/1 5 0 0 0 0 Fa 0/1 6 0 0 0 0 Fa 0/1 7 0 0 0 0 S6510-4C#show priority-flow-control statistics history-top int hu 1/1 Interface : HundredGigabitEthernet 1/1 Prioroty : 0 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 1 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 2 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 3 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 4 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec

	PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 5 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 6 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec Prioroty : 7 PauseSended : 0 PauseSended-Rate : 0 packets/sec PauseSended-Rate-1st : 0 packets/sec PauseSended-Rate-2nd : 0 packets/sec PauseSended-Rate-3rd : 0 packets/sec PauseReceived : 0 PauseReceived-Rate : 0 packets/sec PauseReceived-Rate-1st : 0 packets/sec PauseReceived-Rate-2nd : 0 packets/sec PauseReceived-Rate-3rd : 0 packets/sec

Common Errors

- The common flow control function is not disabled. As a result, PFC fails to be enabled.
- A port is not set to trust the CoS values.
- Priorities with the PFC function enabled are mapped to a TC that contains any other priorities (including priorities with the PFC function enabled).

1.4.1.1. Configuring the PFC Deadlock Detection Function

Configuration Effect

- If the PFC deadlock detection is enabled for a PFC priority on a port and deadlock is detected, allow packets to pass or discard packets within a period of time as per the configured policy.

- If the PFC deadlock threshold function is enabled for a PFC priority and it is detected that the number of deadlock times reaches the threshold, disable the PFC function of the PFC priority.

Notes

- The prerequisite for the PFC deadlock detection function to take effect for a priority is that the PFC function must be enabled for the priority.
- The prerequisite for the PFC deadlock threshold function to take effect for a priority is that the PFC deadlock detection function is configured for the priority.
- Modifying PFC deadlock time precision will affect the detection time and packet processing time of the configured PFC deadlock detection function.
- Disabling the PFC deadlock detection function for a priority will delete the configuration of the PFC deadlock threshold function of the priority.
- Disabling the global PFC deadlock detection function or PFC function for a priority will delete the configuration of the PFC deadlock detection function for the priority on all interfaces.

Configuration Steps

Enabling the PFC Deadlock Detection Function

- The PFC deadlock detection function is disabled for all priorities on all interfaces by default. Configure the global PFC deadlock detection function for a priority and set attribute parameters (use default parameter values if no parameter is configured), and then enable the PFC deadlock detection function for the relevant priority on a specified interface.

Enabling the PFC Deadlock Threshold Function

- The PFC deadlock threshold function is disabled for all priorities by default. The PFC deadlock threshold function can be enabled for a priority, for which the PFC deadlock detection function has been enabled. The threshold can be globally configured.

Verification

Checking the Configuration of the PFC Deadlock Detection Function and Threshold Function

- Run the **show priority-flow-control deadlock status** command to display the configuration of the PFC deadlock detection function and threshold function, and check whether the functions are enabled, configured parameter values, and the interfaces and priorities for which the PFC deadlock detection function is enabled.

Command	**show priority-flow-control deadlock status**
Parameter Description	N/A
Command Mode	Privileged EXEC mode, global configuration mode, and interface configuration mode
Usage Guide	Check the configuration and enabling status of the PFC deadlock function. QTECH#show priority-flow-control deadlock status Global PFC Deadlock: 1-2 Time Precision: Low<100ms> Packet Dispose: Forward Limit Cos-value: off Limit Period: 5 Limit Frequency: 10 Priority Attribute: Cos-value Detect-time Recover-time ------------- --------------- --------------- 1 1 1

	2 15 10000 Interface Enable: Interface Cos-value ------------ ---------- Hu0/1 1-2 Hu0/2 2 Limit Log(Close PFC Log): Interface Cos-valueFrequency Time ------------ ------------- -------------- ------------------------------ Hu0/1 1 20 2017-12-15 11:00:00 Hu0/12 20 2017-12-15 11:00:00 Hu0/11 20 2017-12-15 12:00:00 Hu0/12 20 2017-12-15 12:00:00 Hu0/11 20 2017-12-15 13:00:00 Hu0/1 2 20 2017-12-15 13:00:00 Hu0/11 20 2017-12-15 14:00:00 Hu0/1 2 20 2017-12-15 14:00:00 Hu0/11 20 2017-12-15 15:00:00 Hu0/1 2 20 2017-12-15 15:00:00 Field description:

Field	Description
Global PFC Deadlock	Priority value configured for the PFC deadlock function
Time Precision	PFC deadlock time precision. High precision (10 ms) is configured by defa
Packet Dispose	Packet processing mode in the PFC deadlock state. The value is **Forwa** preceding example.
Limit Cos-value	Priority value configured for the PFC deadlock threshold
Limit Period	Detection period set for the PFC deadlock threshold
Limit Frequency	PFC deadlock threshold
Priority Attribute	PFC deadlock priority attributes (cos-value, detection time, and recovery
Interface Enable	Interface on which the PFC deadlock function is enabled and priority inf (interface name and enabling priority)
Limit Log	PFC deadlock threshold log (recording the historical disabling of the PFC resulted from that the number of PFC deadlock events exceeds the thres
Interface	Interface on which the number of PFC deadlock events exceeds the thres
Cos-value	Priority for which the number of PFC deadlock events exceeds the thresh
Frequency	Number of PFC deadlock events that occur in a period
Time	Time at which the number of PFC deadlock events exceeds the threshold

Related Commands

↘ Configuring Attributes of the PFC Deadlock Detection Function for a PFC Priority

Command	**priority-flow-control deadlock cos-value** *cos-value-list* **detect** *detect-time* **recover** *recover-time*
Parameter Description	*cos-value-list*: Indicates the priority list. The priority ranges from **0** to **7**. *detect-time*: Indicates the deadlock detection time in time slices. The value ranges from **1** to **15**. The default value is **15**. *recover-time*: Indicates the deadlock recovery time in milliseconds. The value ranges from **10** to **3000**. The default value is **150**.
Command Mode	Global configuration mode

Usage Guide	1. The PFC deadlock detection function detects whether an interface at a PFC priority enters the PFC deadlock state. When it is detected that the interface enters the PFC deadlock state, the interface forwards or discards received packets within a period of time. You can run the **priority-flow-control deadlock drop** command to configure the packet processing mode. 2. You can use this command to configure attributes of the PFC deadlock detection function for PFC priorities. 3. *detect-time*: Indicates the time for determining whether an interface at a PFC priority enters the PFC deadlock state. When the interface stays in the PFC congested state in consecutive detect-time periods, it indicates that the interface enters the PFC deadlock state. 4. *recover-time*: Indicates the allowed packet processing time after an interface at a PFC priority enters the PFC deadlock state. 5. *detect-time* is in unit of time slice. You can run the **priority-flow-control deadlock precision low** command to configure the length of one time slice. *recover-time* is in unit of milliseconds.

Enabling the PFC Deadlock Detection Function for a Priority of an Interface

Command	**priority-flow-control deadlock cos-value** *cos-value-list* **enable**
Parameter Description	*cos-value-list*: Indicates the configured priority list. The priority ranges from **0** to **7**.
Command Mode	Interface configuration mode
Usage Guide	1. When the **priority-flow-control deadlock cos-value detect recover** command is not executed to globally configure the PFC deadlock detection function for a priority, the PFC deadlock detection function cannot be enabled on interfaces for the priority. 2. Likewise, if the **priority-flow-control deadlock cos-value detect recover** command, the PFC deadlock detection function will be disabled on all interfaces for the priority.

Configuring the Function of Discarding Packets in the PFC Deadlock State

Command	**priority-flow-control deadlock drop**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	If this command is not configured, an interface at a PFC priority forwards packets after entering the PFC deadlock state. If this command is configured, an interface at a PFC priority discards packets after entering the PFC deadlock state.

Setting the PFC Deadlock Time Precision to Low

Command	**priority-flow-control deadlock precision low**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	1. This command is used to configure the length of a single deadlock time slice. It affects the detect time unit in the **priority-flow-control deadlock cos-value** *cos-value-list* **detect** *detect-time* **recover** *recover-time* command.

	2. When this command is not configured, the PFC deadlock time precision is high (one time slice equals 10 ms). After this command is configured, the PFC deadlock time precision is low (one time slice equals 100 ms). 3. After the PFC deadlock function is configured for a priority, configuring this command will modify the time precision of the PFC deadlock function for the priority.

Enabling the PFC Deadlock Threshold Function for a Priority

Command	**priority-flow-control deadlock limit cos-value** *cos-value-list* **enable**
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	1. This command is not configured by default. Configuring this command will enable the PFC deadlock threshold function for the PFC priority corresponding to *cos-value-list*. If the number of deadlock times occurring for a priority within one detection period (1 minute by default) reaches the threshold (10 times by default), the PFC function is disabled for the priority via the **priority-flow-control nodrop** *cos-value-list* **off** command. 2. This command relies on the configuration of the **priority-flow-control deadlock cos-value** *cos-value-list* **detect** *detect-time* **recover** *recover-time* command. This command can be configured only after the command for configuring attributes of the PFC deadlock detection function is configured for a priority. Likewise, if the command for configuring attributes of the PFC deadlock detection function is deleted for a priority, the PFC deadlock threshold command of the priority is also deleted. 3. You can run the **priority-flow-control deadlock limit frequency** *number* **period** *time* command to configure the detection period and deadlock threshold.

Configuring the PFC Deadlock Threshold and Detection Period

Command	**priority-flow-control deadlock limit frequency** *number* **period** *time*
Parameter Description	N/A
Command Mode	Global configuration mode
Usage Guide	This command modifies the PFC deadlock threshold and detection period. When the **priority-flow-control deadlock limit cos-value** *cos-value-list* **enable** command is executed to enable the PFC deadlock threshold function for a PFC priority, the PFC deadlock threshold and detection period configured using this command are applied.

Configuration Example

Enabling the PFC Deadlock Detection Function and Deadlock Threshold Function for Priority 3

Scenario Figure 1-7	

Configuration Steps	● Set PFC IEEE 802.1p priority to 3 on the server. ● Enable the PFC function for Priority 3 on the switch.
A	QTECH(config)# priority-flow-control nodrop 3 on ● Set attributes of the PFC deadlock detection function for Priority 3 on the switch: Set the detection time to 5 time slices, recovery time to 200 ms, and time precision to low. QTECH(config)# priority-flow-control deadlock cos-value 3 detect 5 recover 200 QTECH(config)# priority-flow-control deadlock precision low ● Enable the PFC deadlock detection function for Priority 3 on Interface Hu0/1 of the switch. QTECH(config)#interface hundredGigabitEthernet 0/1 QTECH(config-if-HundredGigabitEthernet 0/1)#priority-flow-control deadlock cos-value 3 enable ● Enable the PFC deadlock threshold function for Priority 3, set the detection time to 5 minutes, and threshold to 100 times. QTECH(config-if-HundredGigabitEthernet 0/1)#exit QTECH(config)# priority-flow-control deadlock limit frequency 100 period 5 QTECH(config)# priority-flow-control deadlock limit cos-value 3 enable
Verification	Check the displayed configuration.
A	QTECH# show priority-flow-control deadlock status Global PFC Deadlock: 3 Time precision: Low<100ms> Packet dispose: Forward Limit cos-value: 3 Limit period: 5 Limit frequency: 100 Priority attribute: cos-value detect-time recover-time ------------- --------------- ---------------- 3 5 200 Interface Enable: interface cos-value ------------ ---------- Hu0/1 3 Limit Log(Close PFC Log): Interface Cos-valueFrequency Time ------------ ------------- -------------- ------------------------------

Common Errors

- The PFC deadlock detection function is enabled on an interface without globally configuring the PFC deadlock detection function for a priority or globally enabling the PFC function for a priority in advance.
- The PFC deadlock threshold function is configured without globally configuring attributes of the PFC deadlock detection function for a priority.

2.1.4. Configuring Basic Functions of the DCBX

Configuration Effect

- Discover peer DCB capabilities.
- Check whether DCB capabilities configured for both ends are the same.
- Automatically negotiate and configure DCB.

Notes

- The DCB capability negotiation function can take effect on interfaces only after the following requirements are met: the LLDP function is enabled in global configuration mode and interface configuration mode; the LLDP working mode of the interfaces is set to txrx; the TLV of corresponding DCB capability can be released; and corresponding DCB capability is supported. If the requirements are not met, the DCBX function will not take effect even if it is enabled on the interfaces.
- When configuring AP ports, make sure that the LLDP of different AP member ports is set to the same value, the AP port topology is correctly set, and the DCBX TLV released for different AP member ports is the same. Otherwise, abnormalities occur when DCB capabilities are negotiated on the AP ports at both ends.

Configuration Steps

Enabling DCBX

- The global and interface DCBX function is enabled by default.
- Enable DCBX in global or interface configuration mode to launch the DCB capability negotiation of the interface. Disabling the global DCBX function will disable the DCB capability negotiation of all interfaces. The DCBX function runs on physical ports or AP ports, and it is not applicable to stacking ports or VSL ports.

Configuring a DCBX mode

- (Mandatory) Interfaces work in the automatic switch mode by default.
- (Interface configuration mode) The DCBX function runs on physical ports or AP ports, and it is not applicable to stacking ports or VSL ports.

Configuring DCBX TLV types that can be released

- (Mandatory) All types of DCBX TLV can be released on interfaces by default.
- (Interface configuration mode) The DCBX function runs on physical ports or AP ports, and it is not applicable to stacking ports or VSL ports.

Configuring synchronization willingness

- Local end is unwilling to receive peer configuration by default. That is, the Willing parameter is not set.
- During negotiation of DCB capability synchronization, specify the Willing parameter for local ports to receive peer configuration. If local end is willing to receive peer configuration, configure the Willing parameter.

Verification

- Check the display result.

Related Commands

Enabling DCBX

Command	dcbx enable
Parameter Description	N/A

Command mode	Global configuration mode/Interface configuration mode
Usage Guide	Disabling the global DCBX function will disable the DCB capability negotiation of all interfaces. AP ports negotiate DCB capabilities based on AP member ports.

Configuring a DCBX mode

Command	dcbx mode **{** auto | cee | ieee **}**
Parameter Description	Interfaces work in the automatic switch mode by default. **auto**: Allows an interface to automatically switch a DCBX mode based on the received DCBX TLV type if an interface works in the automatic switch mode and the default mode is CEE-DCBX. **cee**: Disallows an interface to automatically switch a DCBX mode based on the received DCBX TLV type if an interface works in CEE-DCBX mode. **ieee**: Disallows an interface to automatically switch a DCBX mode based on the received DCBX TLV type if an interface works in IEEE-DCBX mode.
Command mode	Interface configuration mode
Usage Guide	AP ports negotiate DCB capabilities based on AP member ports.

Configuring DCBX TLV types that can be released

Command	dcbx **{** pfc | app-proto **}** advertise
Parameter Description	All types of DCBX TLV can be released on interfaces by default. **pfc**: Enables the DCBX TLV release function corresponding to the PFC capability. **app-proto**: Enables the DCBX TLV release function corresponding to the APP capability.
Command mode	Interface configuration mode
Usage Guide	AP ports negotiate DCB capabilities based on AP member ports. Types of DCBX TLV that can be released on AP ports will be applied to different AP member ports.

Configuring synchronization willingness

Command	dcbx **{** pfc | app-proto **}** willing enable
Parameter Description	**pfc**: Enables PFC capability synchronization willingness. **app-proto**: Enables APP capability synchronization willingness.
Command mode	Interface configuration mode
Usage Guide	AP ports negotiate DCB capabilities based on AP member ports. If the Willing parameter is set to the same value at both ends and the local and peer ports are willing to receive peer configuration, the configuration on the port with a smaller MAC address prevails. If the local and peer ports are unwilling to receive peer configuration, the same value of the Willing parameter indicates negotiation success; otherwise, it indicates negotiation failure. If DCB capabilities support only global configuration mode, they cannot be negotiated over DCBX. In this case, local configuration prevails. For example, if the PFC function supports only global configuration mode, even though the local PFC Willing bit is set to 1, the local port will not receive peer configuration. That is, the PFC capability cannot be negotiated, and local PFC configuration prevails.

Configuration Example

Automatically configuring DCB

Scenario Figure 1-7	
Configuration Steps	● Enable LLDP in global configuration mode and interface configuration mode (default configuration). ● Set the LLDP mode of an interface to txrx (default configuration). ● Enable DCBX of an interface (default configuration). ● Allow all types of DCBX TLV to be released on an interface (default configuration).
A	By default, switch A can instruct the server to automatically configure DCB, free users of manually configuring DCB.
Verification	Check the display result.
A	QTECH#show dcbx status interface GigabitEthernet 0/1 -- Port [GigabitEthernet 0/1] -- Local managing configuration information and port link state: Port status of DCBX : Enable Port mode of DCBX : AUTO mode Port active of DCBX : CEE active PFC is allowed to advertise : Yes APP is allowed to advertise : Yes PFC is willing to accept peer configuration : No APP is willing to accept peer configuration : No Port link state : UP Negotiated configuration information: PFC negotiated result : using local configuraiton APP negotiated result : using local configuration

Common Errors

- The LLDP function is not enabled in global configuration mode or interface configuration mode.
- The LLDP mode of an interface is not set to txrx.

1.5. Monitoring

Clearing

Running the **clear** commands may lose vital information and thus interrupt services.

Description	Command

Clears statistical information of PFC.	**clear priority-flow-control statistics** [**interface** *interface-name*]
Clears the PFC deadlock statistics of an interface	**clear priority-flow-control deadlock statistics** [**interface** *interface-name*]

Displaying

Description	Command
Displays global PFC configuration information.	**show priority-flow-control status**
Displays PFC configuration and status information of a port.	**show priority-flow-control status**[**interface** *interface-name*]
Displays statistical information of PFC.	**show priority-flow-control statistics [interface** *interface-name*]
Displays the PFC deadlock statistics of all priorities of an interface.	**show priority-flow-control deadlock statistics** [**interface** *interface-name*]
Displays PFC deadlock configuration and enabling status.	**show priority-flow-control status**
Displays DCBX configuration information and negotiation status of an interface.	**show dcbx status** [**interface** *interface-name*]
Displays DCB configuration information and capability negotiation status at both ends.	**show dcbx information** { **pfc** | **app-proto** } [**interface** *interface-name*]

Debugging

System resources are occupied when debugging information is output. Therefore, disable the debugging switch immediately after use.

Description	Command
Debugs the PFC deadlock function.	**debug pfc deadlock**
Debugs the PFC details.	**debug pfc detail**
Debugs the PFC error.	**debug pfc error**
Debugs the PFC event.	**debug pfc event**
Debugs the PFC statistics.	**debug pfc statistics**
Debugs ETS events.	**debug ets event**
Debugs DCBX events.	**debug dcbx event**
Debugs DCBX and DCB capability submodule interaction.	**debug dcbx detail**
Debugs error information during DCBX running.	**debug dcbx error**

Debugs packet transmission and receiving of DCBX.	**debug dcbx pkt**
Debugs change of the DCBX status machine.	**debug dcbx stm**
Debugs hot standby information of DCBX.	**debug dcbx ha**
Debugs the DCB library.	**debug dcb-lib**