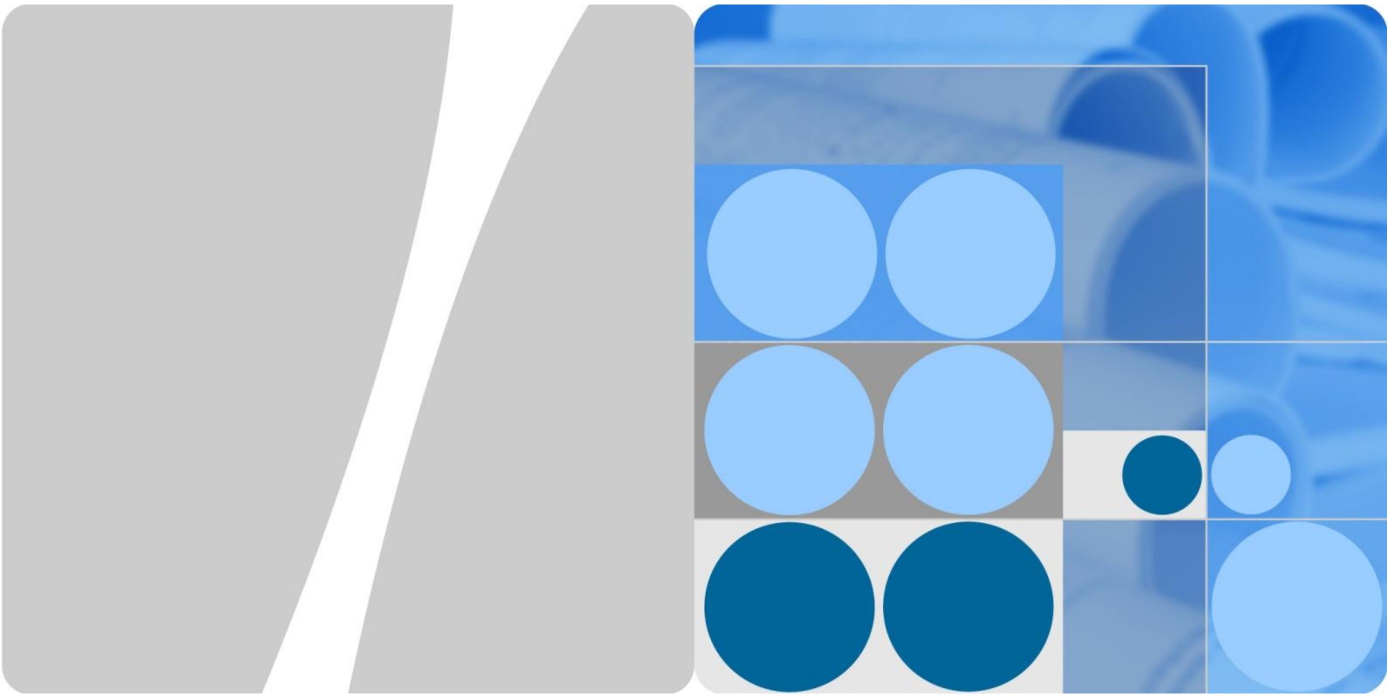


Technical White Paper-FCoE/DCB

Issue	01
Date	2013-04-09

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://enterprise.huawei.com

Email: ChinaEnterprise_TAC@huawei.com

Contents

Figures

Tables

1 FCoE

1.1 Introduction

Definition

Fiber Channel over Ethernet (FCoE) is a network convergence technology and is an I/O consolidation solution based on the Fiber Channel (FC) protocol. FCoE enables LANs and SANs to share network resources.

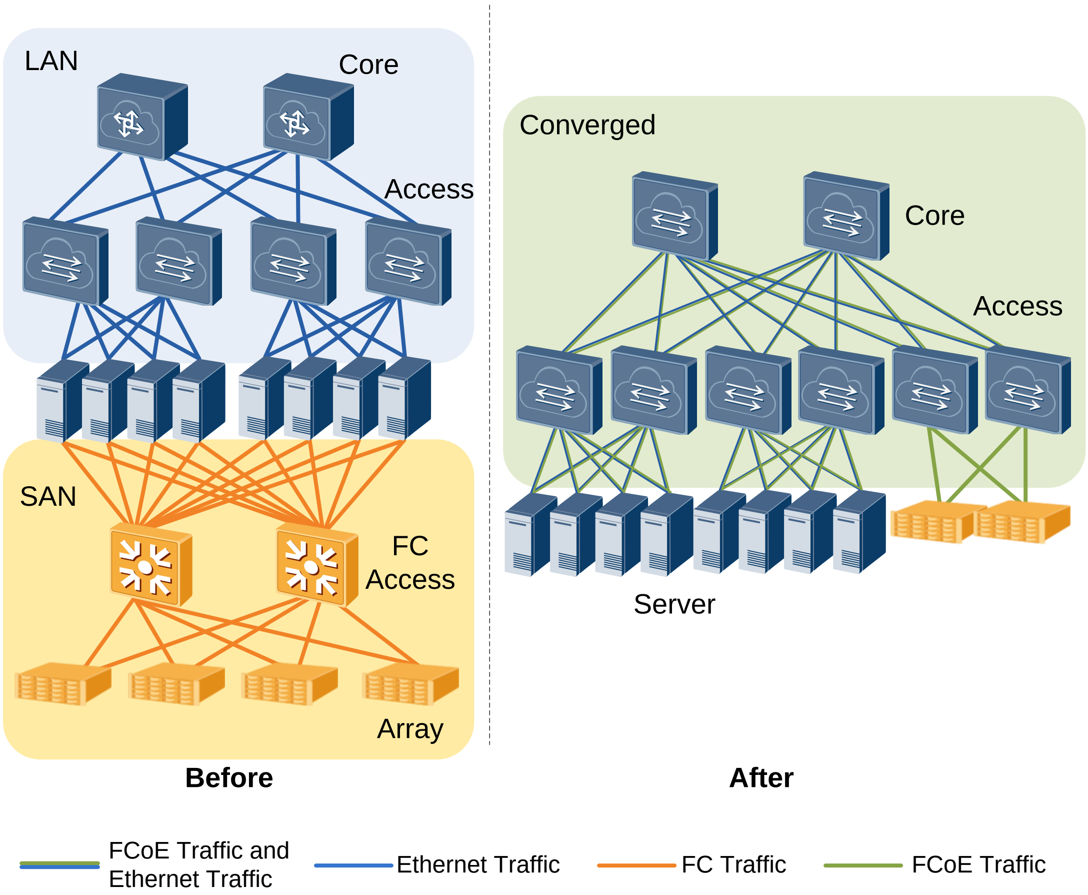
Purpose

As shown in Figure 1-1, the local area network (LAN) and storage area network (SAN) of a traditional data center are deployed and maintained independently. The LAN transmits services between servers and between servers and clients, and the SAN transmits services between servers and storage devices.

As data centers develop rapidly and increasing servers are deployed, independent deployment of LANs and SANs results in the following problems:

- Complex network: Service deployment is inflexible, network expansion is difficult, and network maintenance and management costs are high.
- Low energy efficiency: Each server is configured with at least 4 to 6 network adapters, including network interface cards (NICs) connected to LANs and host bus adapters (HBAs) connected to SANs. Such settings increase power consumption and cooling costs.

Figure 1-1 Before and after data center network convergence



After the LAN and SAN are converged, the SAN and Ethernet LAN share the same integrated network infrastructure, simplifying network infrastructure.

After SAN and LAN networks are converged, the following issues occur:

- FC traffic cannot be forwarded over Ethernet.
- Ethernet cannot ensure lossless forwarding as the FC network.

FCoE and Data Center Bridging (DCB) address these issues:

- FCoE encapsulates FC frames into Ethernet frames and controls FCoE traffic forwarding so that LANs and SANs share network resources. With FCoE technology, LAN and SAN networks can be converged.
- DCB builds a lossless Ethernet network on a data center network. This technology enables traditional Ethernet to implement congestion control as on the FC SAN and provides QoS guarantee for FCoE services.

Benefits

FCoE brings in the following benefits:

- Reduces Total Cost of Ownership (TCO): FCoE allows LANs and SANs to share network resources. It integrates and fully uses distributed resources, reduces investments on SAN network infrastructure, simplifies network topology, and reduces network management and maintenance costs. Servers use converged network adapters (CNAs), which reduce electricity and cooling costs in data centers.

- Saves investment: FCoE seamlessly integrates existing Ethernet and FC infrastructure on data center networks to maximize the return on investment on FC SAN infrastructure, including various FC SAN tools, and established FC SAN facilities and management architecture.
- Enhances service flexibility: FCoE provides capabilities for all servers to access storage devices and allows virtual machine (VM) migrations. This implementation improves system flexibility and availability.

1.2 References

The following table lists the references for this document.

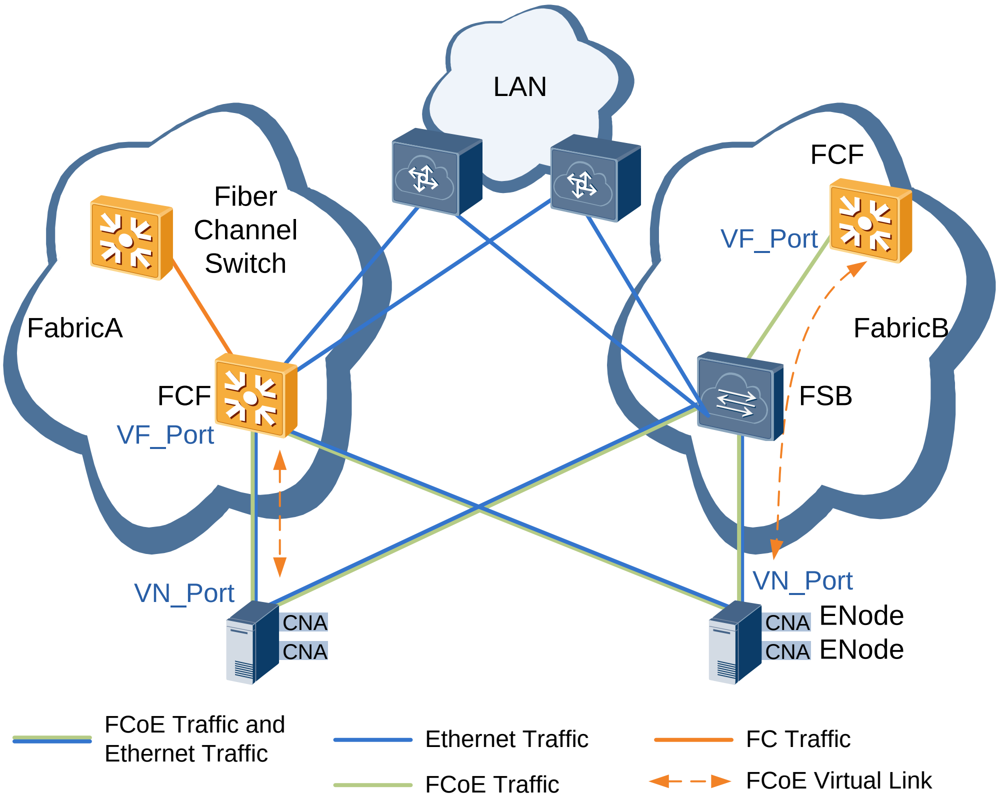
Document	Description	Remarks
FC-BB-5	Fibre Channel Backbone - 5 Rev 2.00	-

1.3 Principles

1.3.1 Basic Concepts of FCoE

As shown in Figure 1-2, FCoE involves the following entities: ENode, FCF, FSB, Fabric, FCoE Virtual Link, FIP, Interface Role, and FCoE VLAN.

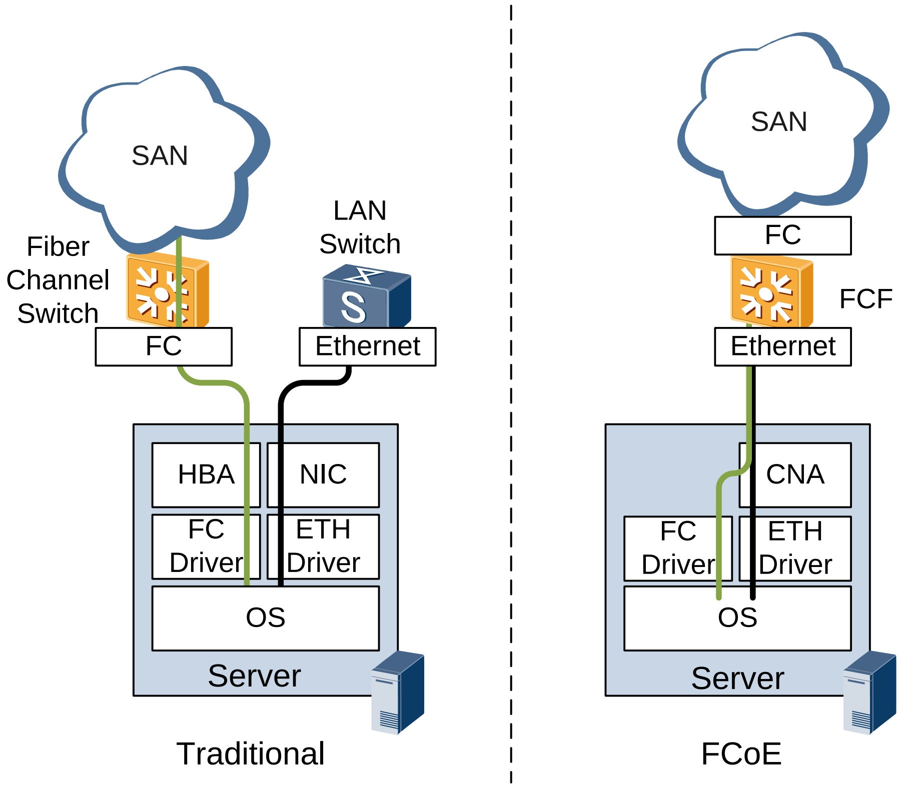
Figure 1-2 FCoE networking



- ENode

  An ENode is a converged network adapter (CNA) that supports FCoE and FC. As shown in Figure 1-3 a traditional server has two network adapters installed: network interface card (NIC) connected to a LAN and a host bus adapter (HBA) connected to a SAN. The CNA provides both NIC and HBA functions. It can forward Ethernet data, process FCoE frames, and encapsulate or decapsulate FCoE frames.

Figure 1-3 Difference between a traditional server and FCoE server



- FCF

  An FCoE forwarder (FCF) is a switch supporting both FCoE and FC, and is used to connect the SAN and LAN. An FCF can forward FCoE packets and encapsulate or decapsulate FCoE frames.

- FSB

  An FCoE Initialization Protocol Snooping Bridge (FSB) is a switch running FIP snooping. The FSB itself does not support FC. FIP snooping enables the FSB to obtain FCoE virtual link information by listening on FIP packets. This function is used to control FCoE virtual link setup and prevent malicious attacks.

- Fabric

  A fabric is the network topology where network nodes are connected through one or more switches.

- FCoE virtual link

  An FCoE virtual link is a point-to-point logical link between FCoE devices, for example, between an ENode and FCF. The connection between an ENode and FCF is not point-to-point when the ENode and FCF are connected through a lossless Ethernet network. The FCoE virtual link is used to solve this problem.

- FIP

  The FCoE Initialization Protocol (FIP) is a Layer 2 protocol that discovers FC terminals on an FCoE network, implements fabric login, and establishes FCoE virtual links. An ENode can log in to the fabric using FIP to communicate with the target FC device. FIP can also maintain FCoE virtual links.

- Interface role

On the traditional FC network, FC devices are connected through FC interfaces. FC interfaces are classified into node ports (N_Ports) and fabric ports (F_Ports):

  - N_Port: FC device interface that connects to an FC switch. An FC device can be a server or storage device.
  - F_Port: FC switch interface that connects to an FC device and provides fabric access services for the FC device.

  FCoE inherits the interface roles of FC. On an FCoE virtual link between an ENode and an FCF, the ENode interface is a VN_Port and the FCF interface is a VF_Port.

- FCoE VLAN

  FCoE frames are forwarded in specified VLANs. In the FC protocol stack, FC devices support multiple virtual storage area networks (VSANs), which are similar to Ethernet VLANs.

  An FCoE virtual link corresponds to one FCoE VLAN. An FCoE VLAN carries only FCoE traffic and does not carry any Ethernet traffic such as IP traffic.

1.3.2 FCoE Encapsulation

FCoE encapsulates FC frames into Ethernet frames so that FC traffic can be transmitted on an Ethernet network. From the FC perspective, FCoE is a different way of transmitting FC traffic. From the Ethernet perspective, FCoE is just another upper layer protocol to carry Ethernet frames.

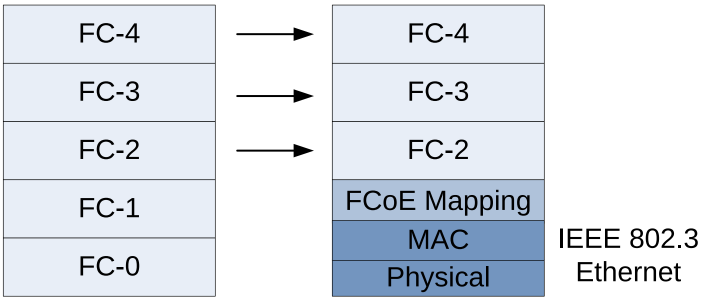
FCoE Protocol Stack

As shown in Figure 1-4, the FC protocol stack is divided into five layers:

- FC-0: defines the media type.
- FC-1: defines the frame encoding mode.
- FC-2: defines the frame format and flow control functions.
- FC-3: defines universal services.
- FC-4: defines mapping from the upper-layer protocol to the FC protocol.

FC-0 and FC-1 in the FCoE protocol stack map Physical and MAC layers in IEEE 802.3 Ethernet respectively. The FCoE protocol stack adds an adaptation layer between the upper-layer FC protocol stack and lower-layer Ethernet protocol stack.

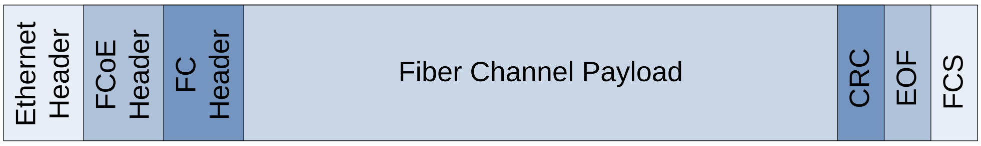
Figure 1-4 Mapping from FC to FCoE



FCoE Frame Encapsulation

FCoE encapsulates an FC frame into an Ethernet frame. Figure 1-5 shows FCoE frame encapsulation.

Figure 1-5 FCoE frame encapsulation



- The Ethernet Header defines the source and destination MAC addresses, Ethernet frame type, and FCoE VLAN.
- The FCoE Header specifies the FCoE frame version number and control information.
- Similar to an FC frame, the FC Header in an FCoE frame carries the source and destination addresses.

1.3.3 FIP Protocol

Principles

The FCoE Initialization Protocol (FIP) establishes and maintains FCoE virtual links between FCoE devices, for example, between ENodes and FC forwarders (FCFs).

An FCoE virtual link is established as follows:

- FIP discovers an FCoE VLAN and the FC virtual interface of the remote device.
- FIP completes initialization tasks such as fabric login (FLOGI) and fabric discovery (FDISC) for the FCoE virtual link.

After an FCoE virtual link is set up, FIP maintains the FCoE virtual link in the following way:

- Periodically detects whether FC virtual interfaces at both ends of the FCoE virtual link are reachable.
- Tears down the FCoE virtual link through Fabric logout (FLOGO).

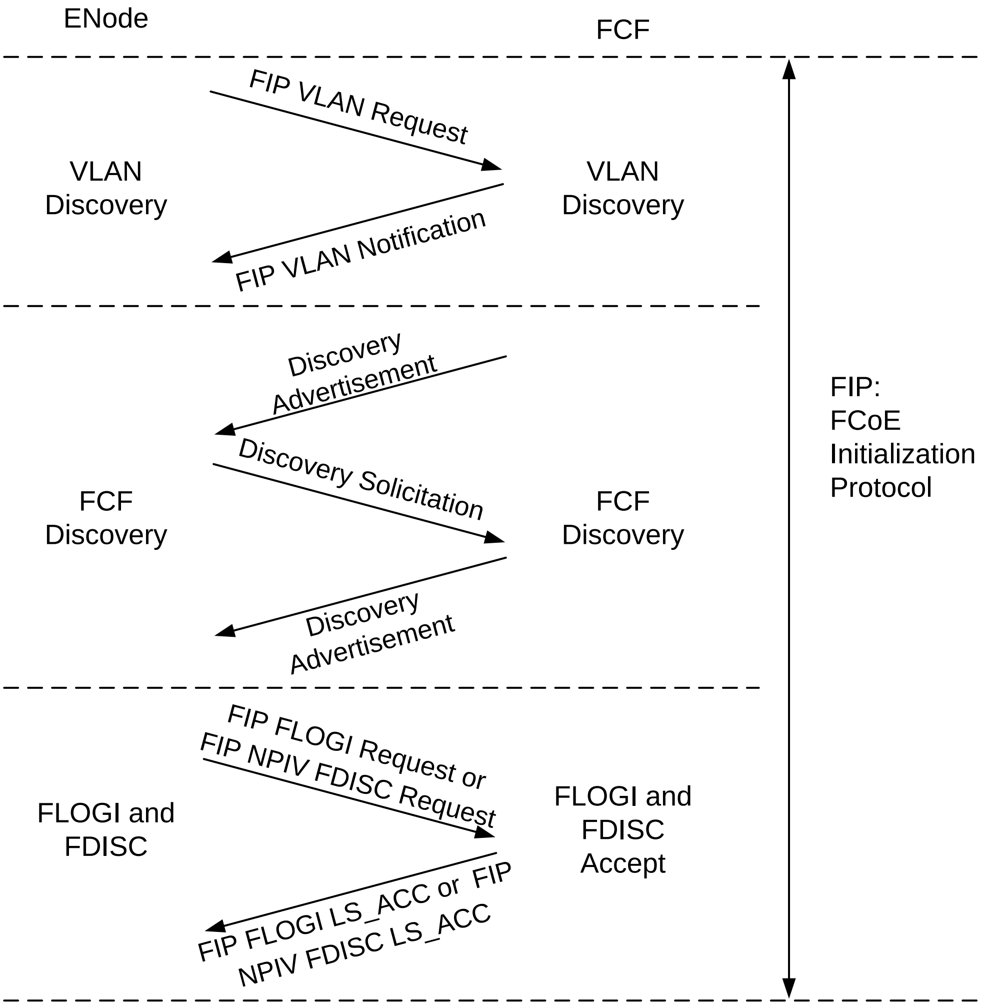
FCoE Virtual Link Setup

Figure 1-6 shows the process of setting up an FCoE virtual link between an ENode and an FCF. The ENode and FCF exchange FIP frames to establish the FCoE virtual link. After the FCoE virtual link is set up, FCoE frames are transmitted on the link.

NOTE

In FIP implementation, an ENode initiates all protocol packets. An FCF also initiates unsolicited FIP Advertisement packets, as described in **FIP FCF discovery**.

Figure 1-6 FCoE virtual link setup



An FCoE virtual link is set up through three phases: FIP VLAN discovery, FIP FCF discovery, and FIP FLOGI and FDISC. The FIP FLOGI and FDISC processes are similar to FLOGI and FDISC processes defined in traditional FC protocol.

1. FIP VLAN Discovery

   FIP VLAN discovery discovers the FCoE VLANs that will transmit FCoE frames. In this phase, an ENode can discover all the potential FCoE VLANs but does not select an FCF.

   The FIP VLAN discovery process is as follows:

   - An ENode sends an FIP VLAN discovery request to a multicast MAC address called All-FCF-MAC (01-10-18-01-00-02). All FCFs listen on packets destined for this MAC address.
   - All FCFs that are reachable in a common VLAN of the ENode report one or more FCoE VLANs to the ENode. The FCoE VLANs are available for the ENode's VN_Port login.

   FIP VLAN discovery is an optional phase as defined in FC-BB-5, the T11 standard covering FCoE. An FCoE VLAN can be manually configured by an administrator, or dynamically discovered using FIP VLAN discovery.

2. FIP FCF Discovery

   ENodes use FIP FCF discovery to locate FCFs that allow logins.

   The FIP FCF discovery process is as follows:

   - Each FCF periodically sends Discovery Advertisement messages in each configured FCoE VLAN. The Advertisement messages are destined for the multicast MAC address All-ENode-MAC (01-10-18-01-00-01) on which all ENodes listen. The FIP FCF discovery Advertisement message contains the FCF MAC address and FCoE virtual link parameters such as the FCF priority and timeout interval of FIP packets.
   - The ENode obtains FCF information from the received Discovery Advertisement messages, selects an FCF with the highest priority, and sends a unicast Discovery Solicitation message to the selected FCF.
   - After receiving the Discovery Solicitation message, the FCF sends a unicast Discovery Advertisement message, allowing the ENode to log in.

   FCFs send Discovery Advertisement messages periodically, but new ENodes joining a network do not want to wait for Discovery Advertisement messages from all FCFs. Therefore, FC-BB-5 allows ENodes to send Discovery Solicitation messages to the multicast MAC address All-FCF-MAC. FCFs that receive the solicitation message send a unicast Discovery Advertisement message to the requesting ENode. Based on the received Discovery Advertisement messages, the ENode selects an FCF with the highest priority to set up a virtual link.

3. FIP FLOGI and FDISC

   After discovering all FCFs and selecting one for login, an ENode sends FIP FLOGI or FIP FDISC packets for establishing an FCoE virtual link with the VF_Port on the selected FCF. Then FCoE frames can be exchanged on the established FCoE virtual link. FIP FLOGI and FIP FDISC packets are unicast packets and correspond to FLOGI and FDISC packets in FC respectively. FIP FLOGI and FIP FDISC packets are used for allocating MAC addresses to ENodes so that the ENodes can log in to the fabric.

   FIP FLOGI is similar to FIP FDISC. The difference is as follows: FIP FLOGI refers to FCoE virtual ink setup when an ENode first logs in to the fabric. FIP FDISC refers to FCoE virtual link setup for each VM when multiple VMs exist on an ENode. FIP FLOGI is used an example.

   The FIP FLOGI process is as follows:

   - An ENode sends an FIP FLOGI Request to the FCF.
   - The FCF responds to the FIP FLOGI Request of the ENode and allocates a locally unique MAC address: Fabric Provided MAC Address (FPMA) to the ENode. Alternatively, the FCF responds to the FLOGI Request of the ENode, agreeing that the ENode uses its locally unique MAC address: Server Provided MAC Address (SPMA).

FCoE Virtual Link Maintenance

On the traditional FC network, FC can immediately detect faults on a physical link. In FCoE, FC cannot immediately detect faults on a physical link because of Ethernet encapsulation. FIP provides a Keepalive mechanism to solve the problem.

FCoE monitors an FCoE virtual link as follows:

- An ENode periodically sends FIP Keepalive packets to an FCF. If the FCF does not receive FIP Keepalive packets within 2.5 times the Keepalive interval, the FCF considers the FCoE virtual link faulty and terminates the FCoE virtual link.

- An FCF periodically sends multicast Discovery Advertisement messages with the destination MAC address as ALL-ENode-MAC to all ENodes. If an ENode does not receive multicast Discovery Advertisement messages within 2.5 times the Keepalive interval, the ENode considers the FCoE virtual link faulty and terminates the FCoE virtual link.

If an FCF does not receive FIP Keepalive packets from an ENode, the FCF sends an FIP Clear Virtual Link message, requesting FCoE virtual link teardown. If the ENode logs out, the ENode can send a Fabric Logout request to the FCF, requesting the FCF to delete the virtual link.

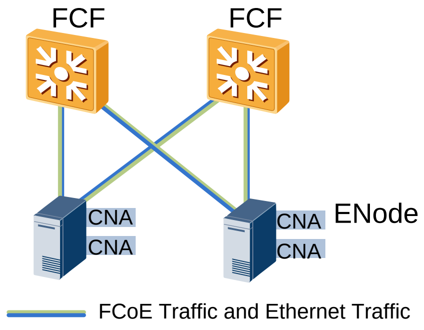
1.3.4 FIP Snooping

An ENode and an FCF can establish a direct connection or remote connection. FIP snooping solves security problems in remote connection mode.

Direct Connection

As shown in Figure 1-7, when an ENode is directly connected to an FCF, the FCoE virtual link and its mapping physical link are point-to-point. Although packets forwarded on the physical link are encapsulated with FCoE, FCoE frame forwarding process is similar to FC frame forwarding because both ends of the physical link support FC.

Figure 1-7 Direct connection



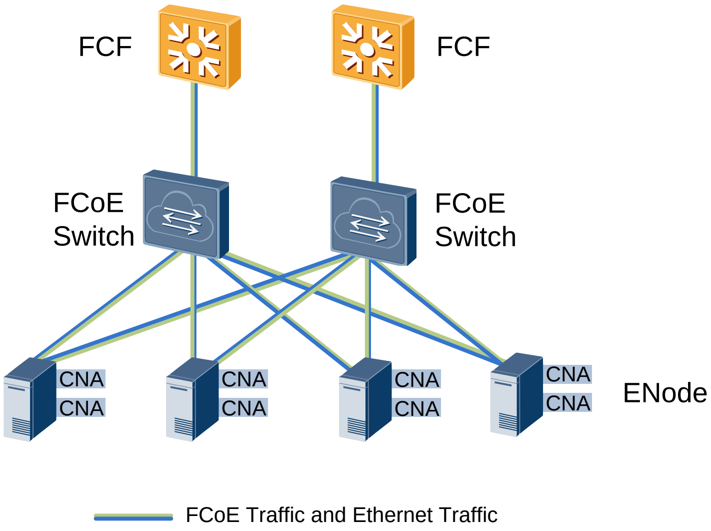
In direct connection mode, FCoE frame processing complies with FC except for data encapsulation at the data link layer. In this mode, FCoE has the same security as FC.

The direct connection mode allows SAN administrators to use original software to manage the SAN when FCoE is used.

Remote Connection

Because the FCF cost is high and a large number of servers are deployed in a data center, establishing direct connections between all servers and FCFs is impractical. As shown in Figure 1-8, access switches are deployed between FCFs and ENodes in remote connection mode. Access switches function as FCoE switches and cannot provide some FCF functions, such as FIP snooping bridge (FSB).

Figure 1-8 Remote connection



NOTE

In remote connection mode, one or more FCoE switches are deployed between ENodes and FCFs.

FIP Snooping

On an FC network, an FC switch is considered a trusted device. Other FC devices such as ENodes must log in to the FC switch before they can connect to the FC network. The FC switch then assigns addresses to the FC devices. FC links are point-to-point, and an FC switch can completely control traffic received and sent by FC devices. Therefore, an FC switch ensures that devices use the assigned addresses to exchange packets and protect FC devices against malicious attacks.

When an FCoE switch is deployed between an ENode and an FCF, FCoE frames are forwarded on the FCoE switch based on the Ethernet protocol because the FCoE switch does not support the FC protocol. In this case, FCoE frames may not be destined for the FCF, and the point-to-point connection between the ENode and FCF is terminated.

To achieve equivalent robustness as an FC network, the FCoE switch must forward FCoE traffic from all ENodes to the FCF. FIP snooping enables the FSB to obtain FCoE virtual link information by listening on FIP packets. This function is used to control FCoE virtual link setup and prevent malicious attacks.

The FCoE switch running FIP snooping is called an FIP snooping bridge (FSB). The CE6800 supports FIP snooping.

Step 1 Configure an FC instance.

```
[~CE6800] fcoe FSB
[~CE6800-fcoe-FSB] vlan 2094
[~CE6800-fcoe-FSB] commit
[~CE6800-fcoe-FSB] quit
```

Step 2 Configure a role for an interface.

```
[~CE6800-10GE1/0/1] fcoe role vnp
[~CE6800-10GE1/0/1] commit
```

```
[~CE6800-10GE1/0/1] quit
```

----End

2 DCB

2.1 Introduction

Definition

Data Center Bridging (DCB) is a set of extensions to Ethernet for use in a data center environment, which is defined by the IEEE 802.1 working group. DCB is used to build lossless Ethernet, meeting QoS requirements on a converged data center network.

Purpose

A converged data center network has LAN traffic, SAN traffic, and IPC traffic. A traditional data center deploys a network for each type of service. As the data center scale increases, the following problems occur:

- Each server requires multiple dedicated adapters (network cards) and a different cabling system.
- The equipment room must be large enough to hold more devices and meet requirements for power consumption and refrigeration of these devices.
- Multiple networks cannot be managed in a unified manner, requiring different maintenance personnel.
- Network deployment, configuration, management, and maintenance become more difficult.

Multi-network integration can solve the preceding problems. However, different types of traffic have different QoS requirements. SAN traffic is sensitive to packet loss and relies on in-order delivery, which means that packets are delivered in the same order in which they were sent. LAN traffic allows packet loss and is delivered on a best-effort (BE) basis. IPC traffic is exchanged between servers and requires low latency. DCB was developed to meet different QoS requirements of these types of traffic (especially the SAN traffic) on the Ethernet.

2.2 References

The following table lists the references for this document.

Document	Description	Remarks
IEEE 802.1 Qbb	Priority-based Flow control	-
IEEE 802.1 Qaz	• Enhanced transmission selection • Data Center Bridging Exchange (DCBX) Protocol	

2.3 Principles

Among DCB configuration tasks, configuring PFC and configuring ETS are mandatory and can be performed in any sequence. Configuring DCBX is optional. When PFC is configured to work in **auto** mode, configure DCBX.

Table 2-1 lists DCB features.

Table 2-1 List of DCB features

Feature	Purpose
Priority-based Flow control (PFC)	Provides priority-based traffic control on a shared link.
Enhanced transmission selection (ETS)	Improves bandwidth efficiency on a shared link.
Data Center Bridging Exchange (DCBX) Protocol	Negotiates Ethernet parameters on both ends of a link automatically to reduce the management cost.

2.3.1 PFC

Background

SAN traffic is sensitive to packet loss on a converged network.

The Ethernet Pause mechanism ensures the lossless transmission service. When a downstream device detects that its receive capability is lower than the transmit capability of its upstream device, it sends Pause frames to the upstream device, requesting the upstream device to stop sending traffic for a period of time. The Ethernet Pause mechanism stops all traffic on a link, whereas FCoE requires link sharing:

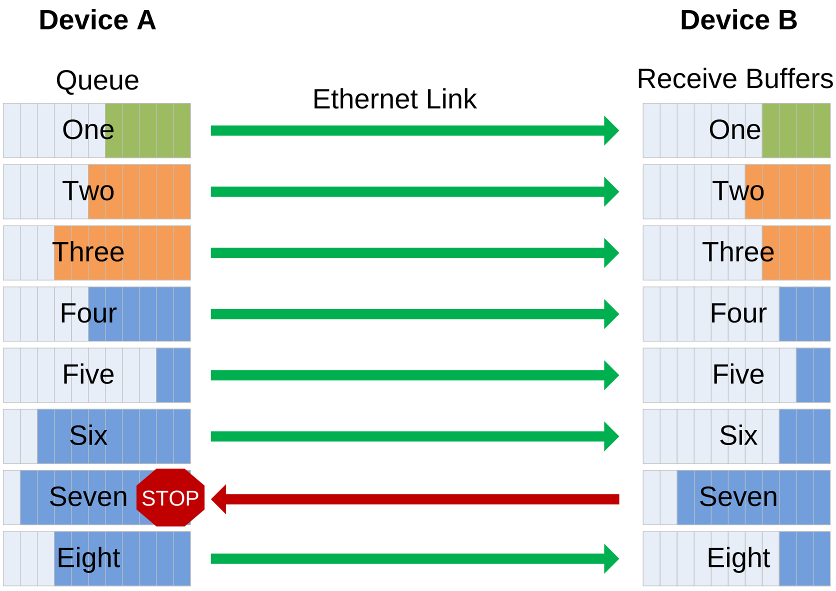
- Burst traffic of one type cannot affect forwarding of traffic of other types.
- A large amount of traffic of one type in a queue cannot occupy buffer resources of traffic of other types.

Priority-based Flow Control (PFC) addresses the contradiction between the Ethernet Pause mechanism and link sharing.

Principles

PFC is also called Per Priority Pause or Class Based Flow Control (CBFC). It enhances the Ethernet Pause mechanism. As shown in Figure 2-1, eight priority queues on the transmit interface of DeviceA correspond to eight buffers on the receive interface of DeviceB. When a receive buffer on DeviceB is to be congested, DeviceB sends a backpressure signal to DeviceA, requesting DeviceA to stop sending packets in the corresponding priority queue.

Figure 2-1 PFC working mechanism



A backpressure signal is an Ethernet frame. Figure 2-2 shows the PFC frame format.

Figure 2-2 PFC frame format

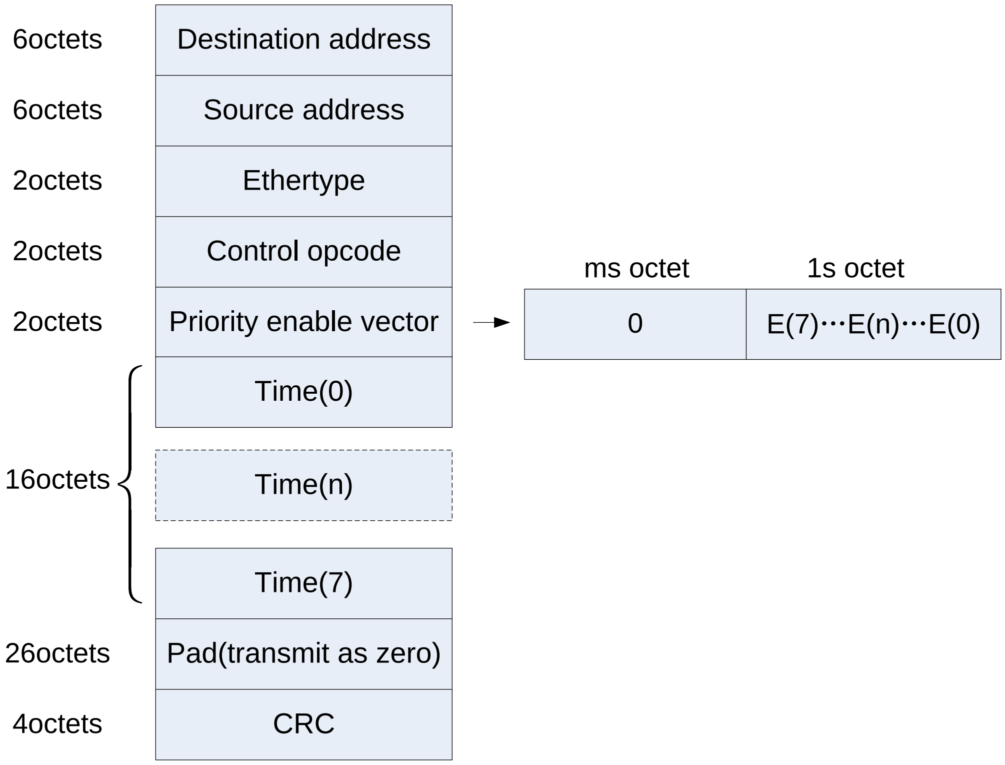


Table 2-2 Fields in a PFC frame

Item	Description
Destination address	Destination MAC address, which has a fixed value of 01-80-c2-00-00-01.
Source address	Source MAC address.
Ethertype	Ethernet frame type. The value is 88-08.
Control opcode	Control code. The value is 01-01.
Priority enable vector	Priority-enable vector. E(*n*) corresponds to queue *n* and determines whether backpressure is enabled for queue *n*. When E(*n*) is 1, backpressure is enabled for queue *n*. When E(*n*) is 0, backpressure is disabled for queue *n*.
Time(0)-Time(7)	Backpressure timer. If Time(n) is 0, backpressure is canceled.
Pad(transmit as zero)	Reserved. The value is 0 during PFC frame transmission.
CRC	Cyclic Redundancy Check (CRC).

When receiving backpressure signals, a device only stops traffic in one or several priority queues, but does not stop traffic on the entire interface. PFC can pause or restart any queue, without interrupting traffic in other queues. This feature enables traffic of various types to share one FCoE link. The system does not apply the backpressure mechanism to the priority queues with PFC disabled and directly discards packets in these queues when congestion occurs.

In an FCoE environment, an administrator can apply PFC to queues of FCoE traffic to ensure lossless transmission of FCoE service.

Configuring PFC

The CE6800 supports PFC.

```
[~CE6800] interface 10ge 1/0/1
[~CE6800-10GE1/0/1] dcb pfc enable mode auto
[~CE6800-10GE1/0/1] quit
[~CE6800] commit
```

2.3.2 ETS

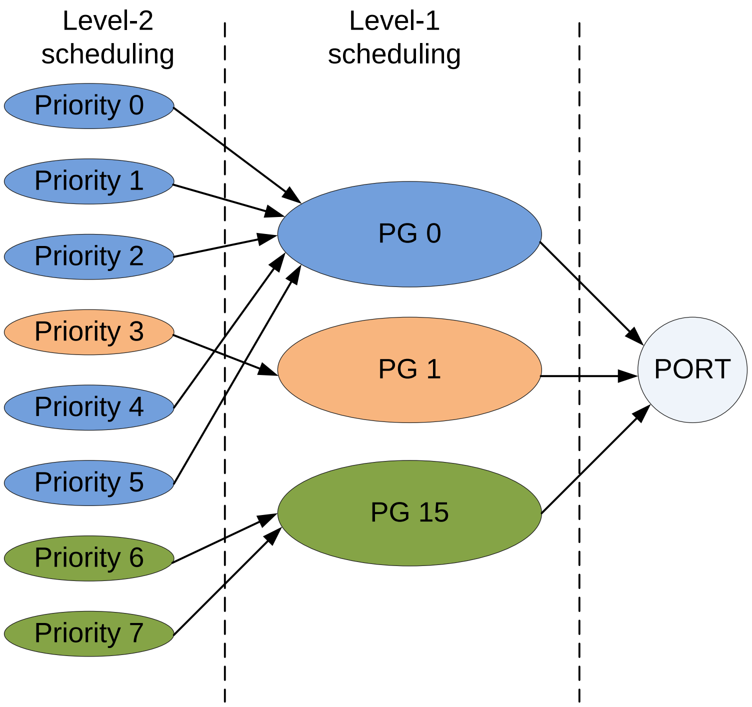
Background

A converged data center network has LAN traffic, SAN traffic, and IPC traffic. The converged network has high QoS requirements. Traditional QoS cannot meet requirements of the converged network, whereas ETS uses a hierarchical flow control mechanism to implement QoS on the lossless Ethernet.

Principles

ETS provides two-level scheduling: scheduling based on priority groups and scheduling based on queues, as shown in Figure 2-3. An interface first schedules priority groups, and then schedules priority queues.

Figure 2-3 ETS Process



Compared with common QoS, ETS provides scheduling based on priority groups. ETS adds traffic of the same type to a priority group so that traffic of the same type obtains the same CoS.

Scheduling Based on Priority Groups

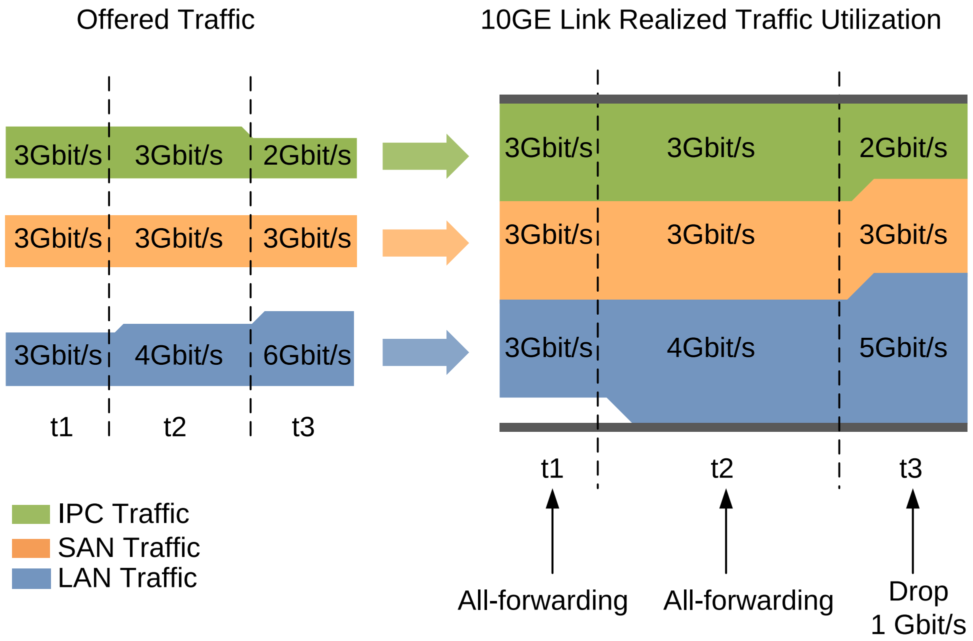
A priority group is a group of priority queues using the same scheduling mode. You can add queues with different priorities to a priority group. Scheduling based on the priority group is called level-1 scheduling.

ETS defines three priority groups: PG0, PG1, and PG15. PG0, PG1, and PG15 process LAN traffic, SAN traffic, and IPC traffic respectively.

As defined by ETS, PG0, PG1, and PG15 use PQ+DRR. PG15 uses priority queuing (PQ) to schedule delay-sensitive IPC traffic. PG0 and PG1 use Deficit Round Robin (DRR). Bandwidth can be allocated to priority groups based on actual networking.

As shown in Figure 2-4, the queue with priority 3 carries FCoE traffic, so this queue is added to the SAN group (PG1). Queues with priorities 0, 1, 2, 4, and 5 carry LAN traffic, so these queues are added to the LAN group (PG0). The queue with priority 7 carries IPC traffic, so this queue is added to the IPC group (PG15). The total bandwidth of the interface is 10 Gbit/s. PG1 and PG0 each obtain 50% of the total bandwidth, 5 Gbit/s.

Figure 2-4 Congestion management based on priority groups



At t1 and t2, all traffic can be forwarded because the total traffic on the interface is within the interface bandwidth. At t3, the total traffic exceeds the interface bandwidth and LAN traffic exceeds given bandwidth. At this time, LAN traffic is scheduled based on ETS parameters and 1 Gbit/s LAN traffic is discarded.

ETS also provides traffic shaping based on priority groups. This traffic shaping mechanism limits traffic bursts in a priority group to ensure that traffic in this group is sent out at an even rate.

Priority-based Scheduling

ETS also provides priority-based scheduling, level-2 scheduling.

In addition, ETS provides priority-based queue congestion management, queue shaping, and queue congestion avoidance.

Configuring ETS

The CE6800 supports ETS.

Step 1 Configure an ETS profile.

```
[~CE6800] dcb ets-profile ets1
```

Step 2 Apply an ETS profile.

```
[~CE6800] interface 10ge 1/0/1
[~CE6800-10GE1/0/1] dcb ets enable ets1
[~CE6800-10GE1/0/1] quit
[~CE6800] commit
```

----End

2.3.3 DCBX

Background

To implement lossless Ethernet on a converged data center network, both ends of an FCoE link must have the same PFC and ETS parameter settings. If PFC and ETS parameters are manually configured, the administrator's workload is heavy and configuration errors may occur. DCBX, as a link discovery protocol, enables DCB devices at both ends of a link to discover and exchange DCB configurations, reducing workloads of administrators.

Principles

DCBX provides the following functions:

- Discovers the DCB configuration of the remote device.
- Detects the DCB configuration errors of the remote device.
- Configures the remote device if permitted.

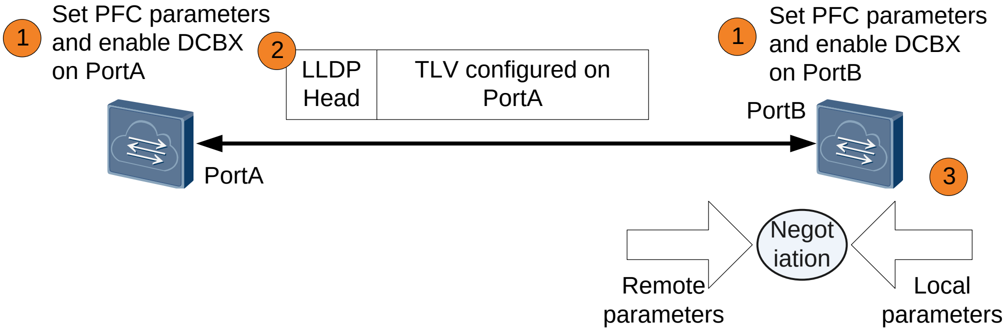
DCBX enables DCB devices at both ends to exchange the following DCB configurations:

- ETS priority group
- PFC

DCBX encapsulates DCB configurations into Link Layer Discovery Protocol (LLDP) TLVs so that devices at both ends of an FCoE virtual link can exchange DCB configurations.

PFC is used as an example to describe DCBX implementation through LLDP.

Figure 2-5 DCBX implementation through LLDP



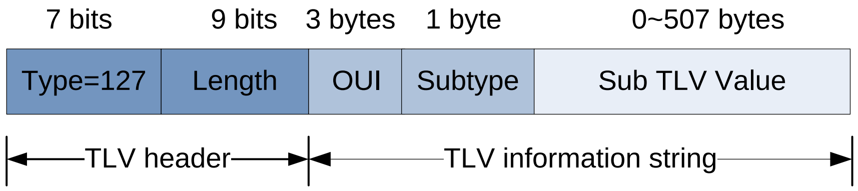
As shown in Figure 2-5, LLDP is enabled on PortA and PortB, and PortA is configured to send DCBX TLVs. The implementation is as follows:

1. Set PFC parameters on PortA and PortB, and enable DCBX. The DCBX module instructs PortA and PortB to encapsulate their PFC parameters into LLDPDUs and send the LLDPDUs to each other.
2. The LLDP module of PortA sends LLDPDUs with DCBX TLVs to PortB at intervals.
3. PortB parses the DCBX TLVs in the received LLDPDUs and sends PFC parameters of PortA to the DCBX module. The DCBX module compares PFC parameters of PortA with its PFC parameters. Through negotiation, PFC parameters on the two ends are consistent, and a configuration file is then generated.

DCBX TLV

As shown in Figure 2-6, the DCB configuration is encapsulated into specified TLVs. The Type field has a fixed value of 127, and the OUI field has a fixed value of 0x0080c2.

Figure 2-6 DCBX TLV format



DCBX TLVs include the ETS Configuration TLV, ETS Recommendation TLV, and PFC Configuration TLV. Table 2-3 describes the DCBX TLVs.

Table 2-3 DCBX TLVs

TLV	Subtype	Length	Description
ETS Configuration TLV	09	25	Local ETS configuration: Priority group configuration: priority group ID and bandwidth usage of a priority group Priority queue configuration: priority queue ID and its priority group ID
ETS Recommendation TLV	0A	25	Recommended ETS configuration, used for ETS configuration negotiation between both ends of an FCoE virtual link: Priority group configuration: priority group ID and bandwidth usage of a priority group Priority queue configuration: priority queue ID and its priority group ID
PFC Configuration TLV	0B	6	Local PFC configuration: Priority queue ID Whether PFC is applied to a queue

Configuring DCBX

The CE6800 supports DCBX.

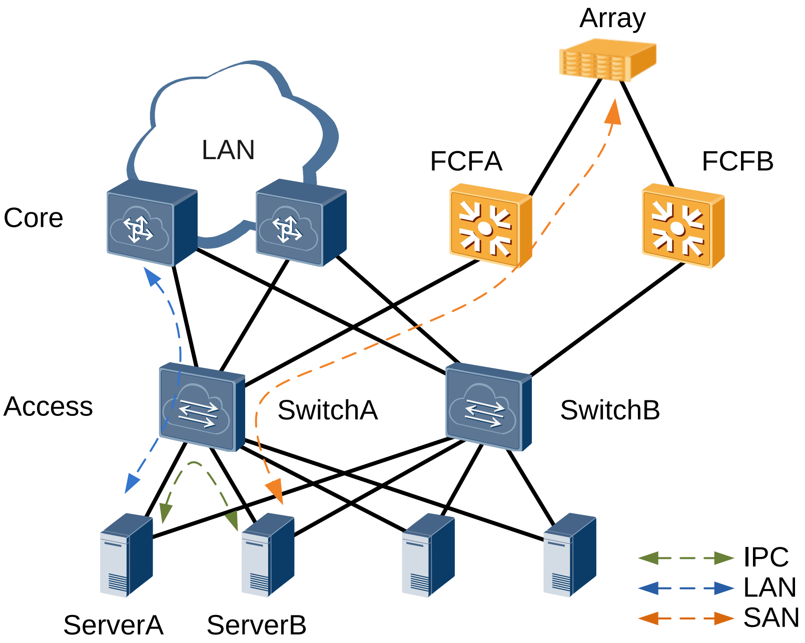
```
[~CE6800] lldp enable
[~CE6800] interface 10ge 1/0/1
[~CE6800-10GE1/0/1] lldp tlv-enable dcbx
[~CE6800-10GE1/0/1] quit
[~CE6800] commit
```

3 Applications

FCoE&DCB Networking

FCoE converges LANs and SANs of data centers. Access switches are often deployed between servers and FCFs to reduce investments. As shown in Figure 3-1, SwitchA functions as an access switch and needs to forward LAN, SAN, and IPC traffic. The two links ServerA->Switch A->FCF A->Array and ServerA->Switch B->FCF B->Array ensure link reliability between ServerA and the array.

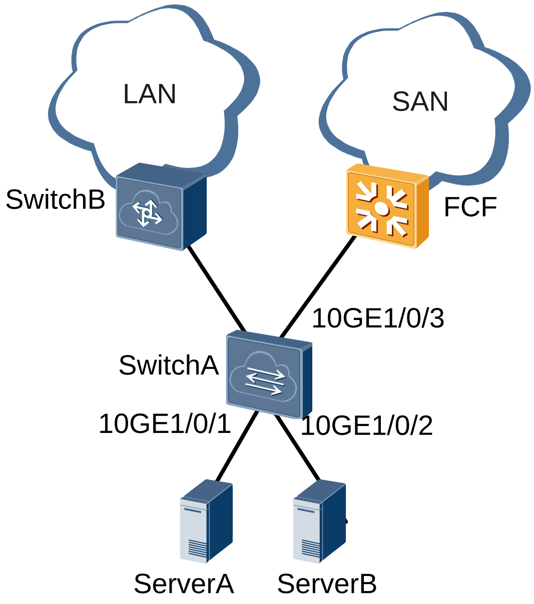
Figure 3-1 FCoE&DCB networking



On SwitchA, FIP snooping is configured to ensure correct SAN traffic forwarding, and DCB is configured to guarantee QoS of LAN, SAN, and IPC traffic.

Figure 3-2 shows the networking diagram of configuring FCoE/DCB.

Figure 3-2 Networking diagram of configuring FCoE/DCB



# Configuration file of SwitchA

```
#
sysname SwitchA
#
dcb pfc
#
dcb ets-profile ets1
 priority-group 0 queue 0 to 2 4 to 6
 priority-group 15 queue 7
 priority-group 0 drr weight 60
 priority-group 1 drr weight 40
#
fcoe FSB
 vlan 2094
#
lldp enable
#
diffserv domain ds1
 8021p-inbound 3 phb af1 green
 8021p-outbound af1 green map 3
#
interface 10GE1/0/1
 port link-type trunk
 port trunk allow-pass vlan 2094
 lldp tlv-enable dcbx
 trust upstream ds1
 dcb pfc enable mode auto
 dcb ets enable ets1
#
interface 10GE1/0/2
 port link-type trunk
 port trunk allow-pass vlan 2094
 lldp tlv-enable dcbx
 trust upstream ds1
 dcb pfc enable mode auto
 dcb ets enable ets1
#
interface 10GE1/0/3
 port link-type trunk
 port trunk allow-pass vlan 2094
 lldp tlv-enable dcbx
 fcoe role vnp
 dcb pfc enable mode auto
#
Return
```

A Acronyms and Abbreviations

Acronym

Acronym	Full Name
FCoE	Fibre Channel over Ethernet
FC	Fibre Channel
SAN	Storage Area Network
NIC	Network Interface Card
HBA	Host Bus Adapter
DCB	Data Center Bridging
CNA	Converged Network Adapter
FCF	FCoE Forwarder
FSB	FCoE Initialization Protocol Snooping Bridge
FIP	FCoE Initialization Protocol
PQ	Priority Queue
DRR	Deficit Round Robin
DCBX	Data Center Bridging Exchange Protocol
LLDP	Link Layer Discovery Protocol