Global Knowledge®

Expert Reference Series of White Papers

Fibre Channel Essentials

Fibre Channel Essentials

Raymond B. Dooley, Global Knowledge Instructor, CCSI, CCNP, CCDP

Introduction

As a major component of any major enterprise network, the corporate data center has evolved into an integral and highly complex environment. The data center is now important enough to be a separate strategic focus for Cisco and many other networking companies. The demand is high for skilled personnel in this area.

But the objective of this white paper is not to describe the design or operations of the data center. To establish a preamble to our objective, consider all the options that fill the role of a server in the data center:

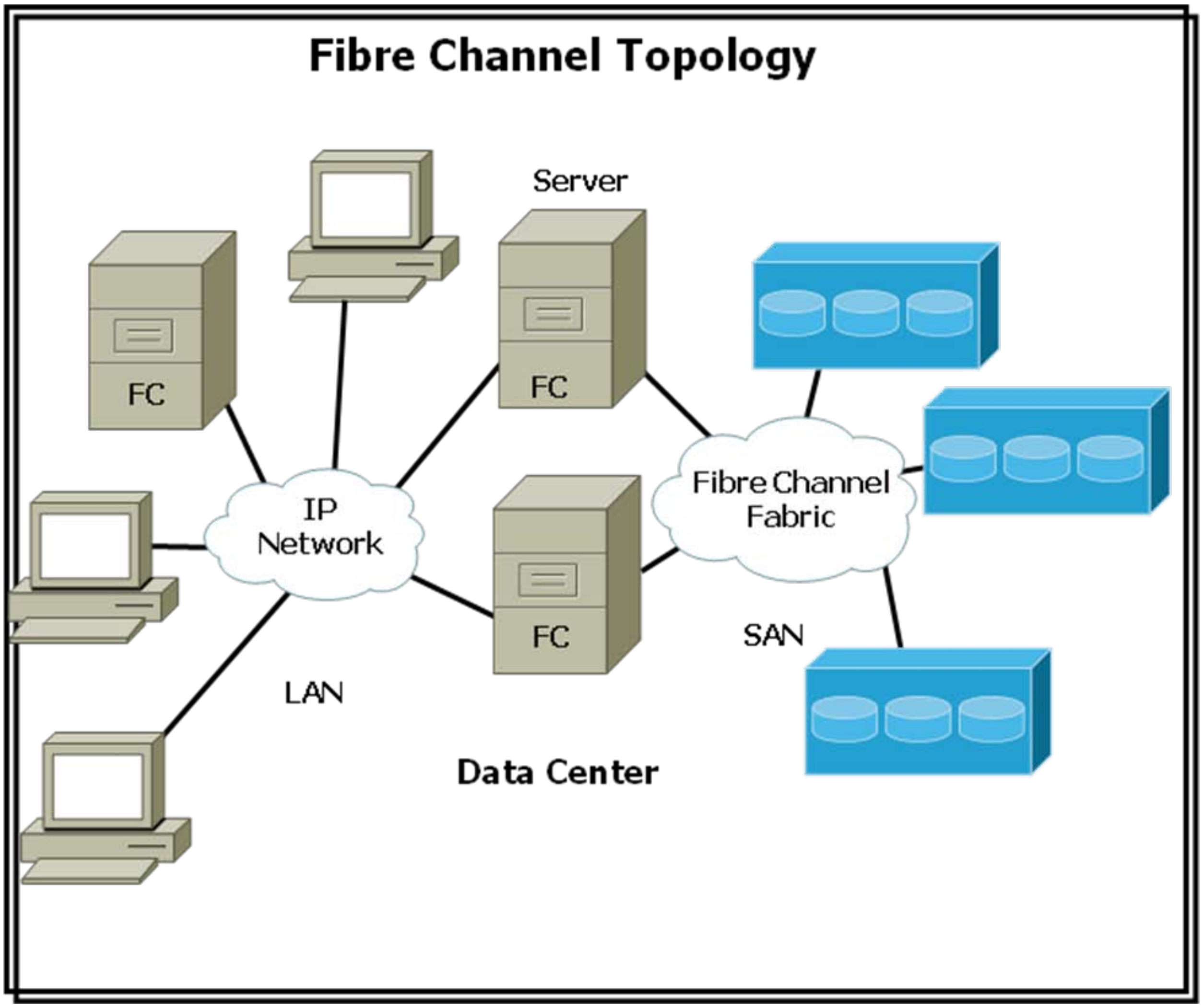
- Dell
- IBM
- HP
- Sun
- Linux
- Red Hat
- IBM Mainframe
- Blade Servers
- NIC Teams with VMware

This is not an exhaustive list. All of the servers listed have one thing in common: the need to retrieve data from storage media such as disks or optic memory arrays. The protocol utilized to transport data between servers and storage for High Performance Computing (HPC) and to define a Storage Area Network (SAN) is Fibre Channel (FC). Fibre Channel is an “invisible” protocol, because it is only known to the technicians that implement it and keep it working. Just as a PC user is unaware of the underlying technology with which the PC retrieves data from the hard drive, the data center client is likewise in the dark.

To provide an explanation of the essentials of this technology, an overview of the following will be provided:

- Fibre Channel/SAN topology
- FC ports
- FC vs. OSI
- FC port addressing and naming
- Classes of service
- Flow control
- Error control
- FC fabric components
- VSANs, zones, services

Fibre Channel Essentials



In the graphic above, the server could be any of the ones listed. The Fibre Channel fabric is composed of Fibre Channel switches such as the Cisco MDS 9000 series, the Nexus 5000 series, as well as switches from Brocade, McData and others. Modern topologies include:

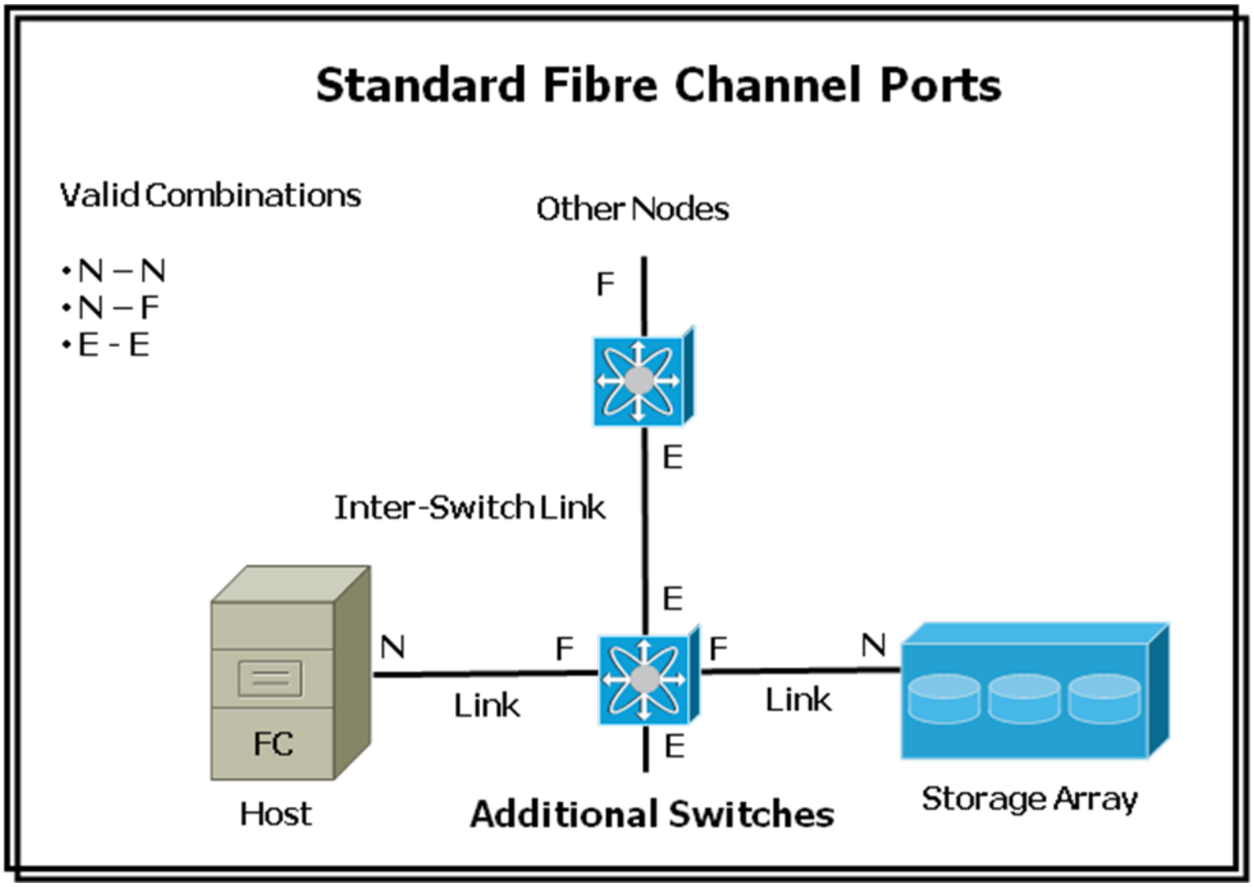
- Point-to-point
- Switched fabric

Point-to-point is any two FC ports connected together. When a switch is inserted, multiple point-to-point connections are made to the switch, creating a switched fabric. Switched fabric is the most common technology in SANs and allows thousands of devices and up to 239 meshed switches. The devices are:

- Server (host)
- Storage array or just a bunch of disks (JBOD)
- Tape storage
- Fibre switch

Each device manufacturer provides Fibre Channel port technology for the device. An FC port is not just a physical interface; it is an intelligent component of the fabric. The servers use a Host Bus Adapter (HBA). The disk, tape storage devices, and the switch have a port/s as well.

FC is documented in American National Standards Institute (ANSI) specifications. In older implementations an arbitrated loop topology along with an appropriate hub device was supported. That technology is not included here.



- A node port (N Port) is a port on a node that connects to a fabric.
- A fabric port (F port) is a port on a switch that connects to an N port.
- An extension port (E port) is a port on a switch that connects to an E Port on another switch. Another name of this connection is an Inter-Switch Link (ISL).

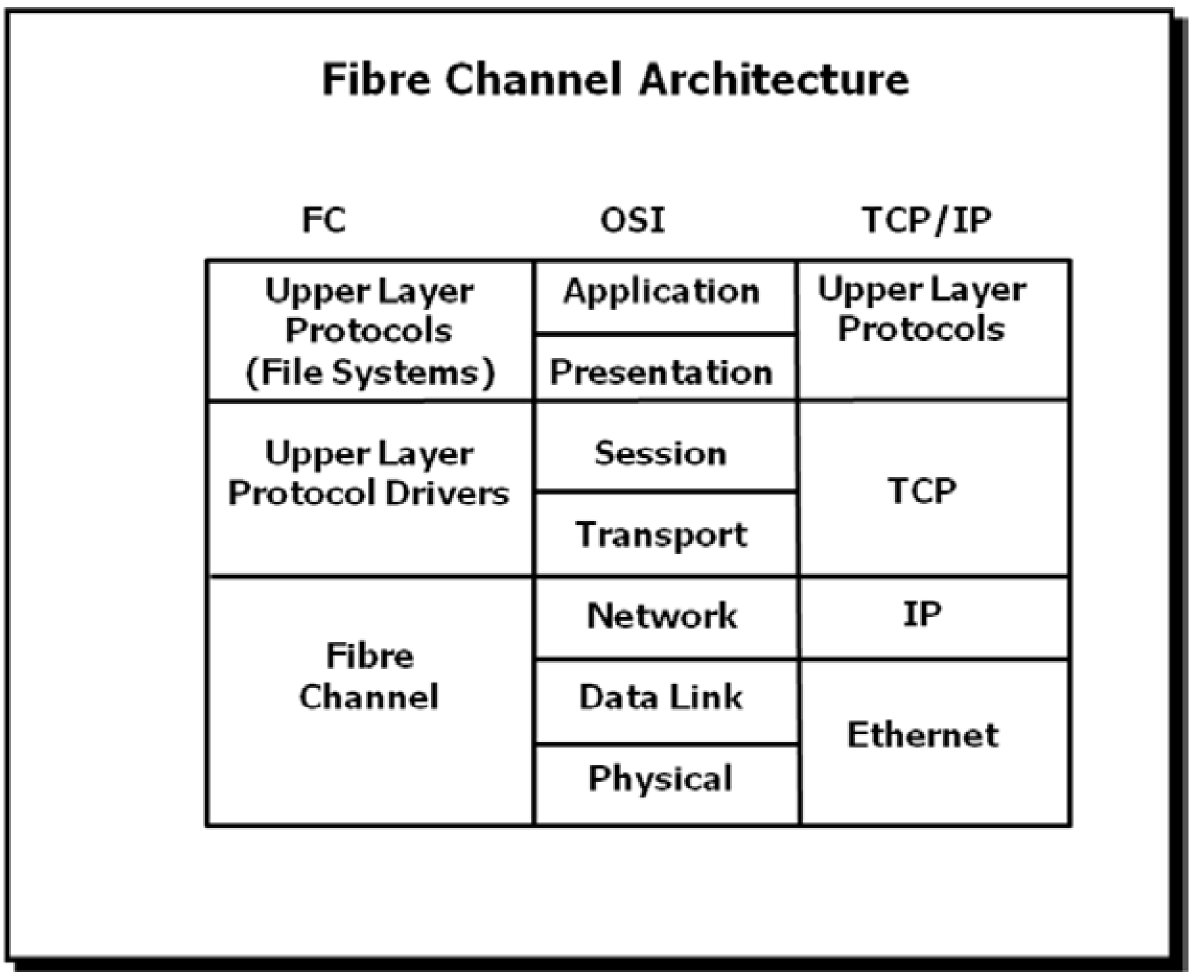
There are other ports for connectivity between the switch and the older arbitrated loop hubs not defined here.

World Wide Names (WWNs) are unique identifiers that are assigned to manufacturers by the International Electrical and Electronic Engineers (IEEE) and hard-coded into FC devices. The names are usually 128 bits long (it could be 64 if older). Since one server may have dual ports, the WWN has two formats:

- nWWN (node WWN)
- pWWN (port WWN)

Each port has a Fibre Channel identifier (FCID), which is 24 bits in length and consists of three parts (domain, area, and port) so that many devices can be identified within a domain. FCIDs typically are assigned dynamically by the switch during start-up.

Obviously, the port is much more than a physical interface. With all the parts named and addressed, a name service is necessary (provided by switch) to resolve WWNs (names) to FCIDs (port addresses). This is somewhat analogous to IP and Domain Name Service in a TCP/IP network.

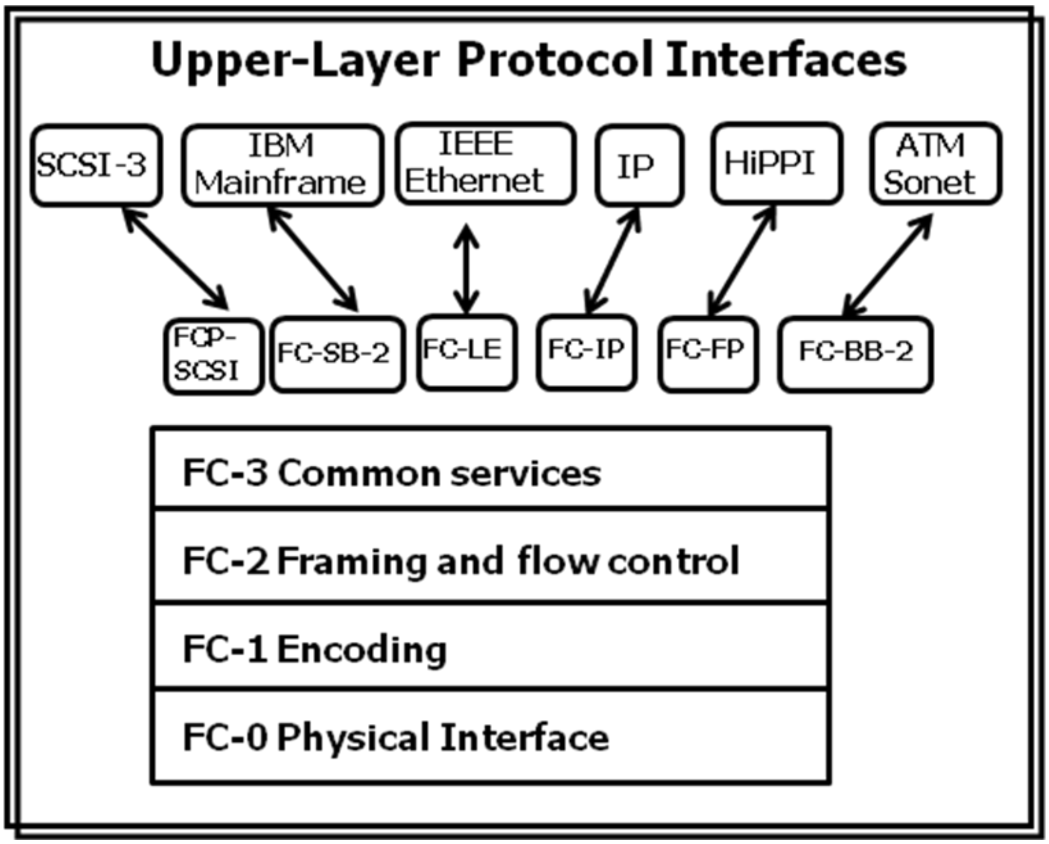


Fibre Channel defines the lower three layers of the OSI Reference Model: physical; data link; and network.

The upper layer protocols (ULPs) are file systems such as Microsoft's NT File System (NTFS), Common Internet File System (CIFS), and so on.

The ULP's drivers provide the transport and session functions. The most prevalent is Small Computer Systems Interface (SCSI). Others include Virtual Interface (VI) and High Performance Parallel Interface (HiPPI). Ninety-five percent of the payload for FC is SCSI commands, data, and messages. With the introduction of newer and greater server technologies, High Performance Computing (HPC) using something other than SCSI (perhaps HiPPI) is becoming more common.

FC functions are broken down even further than the three layers in the OSI model.

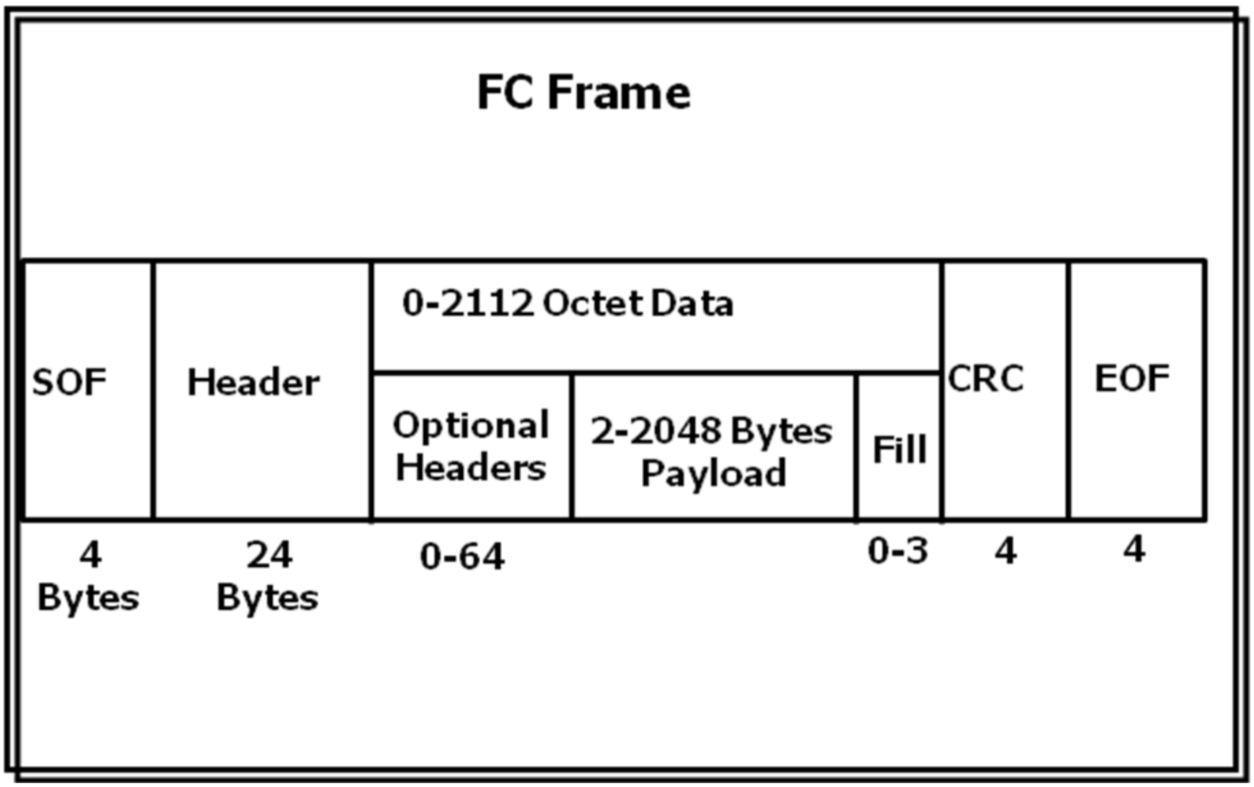


The further definition of FC functions in the graphic above is simple with some additional explanation. Contrary to the name of the technology (Fibre Channel), the physical media can be either copper or optic fiber.

The busiest function in FC is FC-2, framing and flow control. The robustness, reliability, and scalability of FC depend on these functions, which will be covered in more detail later.

FC-3 provides support services for Transport level protocol interfaces (FC-4, at the top of the graphic). These services include name server, secure key server, management server, and time server.

SCSI is the most common interface to FC-3. IBM mainframes also have an interface. Ethernet frames with any payload can run on FC. FC also supports TCP/IP packets, ATM/SONET frames, and the previously described High Performance Parallel Interface. This is a bit misleading, because it makes FC look like a legitimate layer 3 tunneling method for all of the protocols listed. It is not true due to the distance limitations of FC. When these types of packets are transmitted on FC, it is usually within a data center environment.



A simplified representation of an FC frame includes a Start of Frame and End of Frame word (4 octets), a header, and a Cyclic Redundancy Check field for error detection, optional headers, and payload.

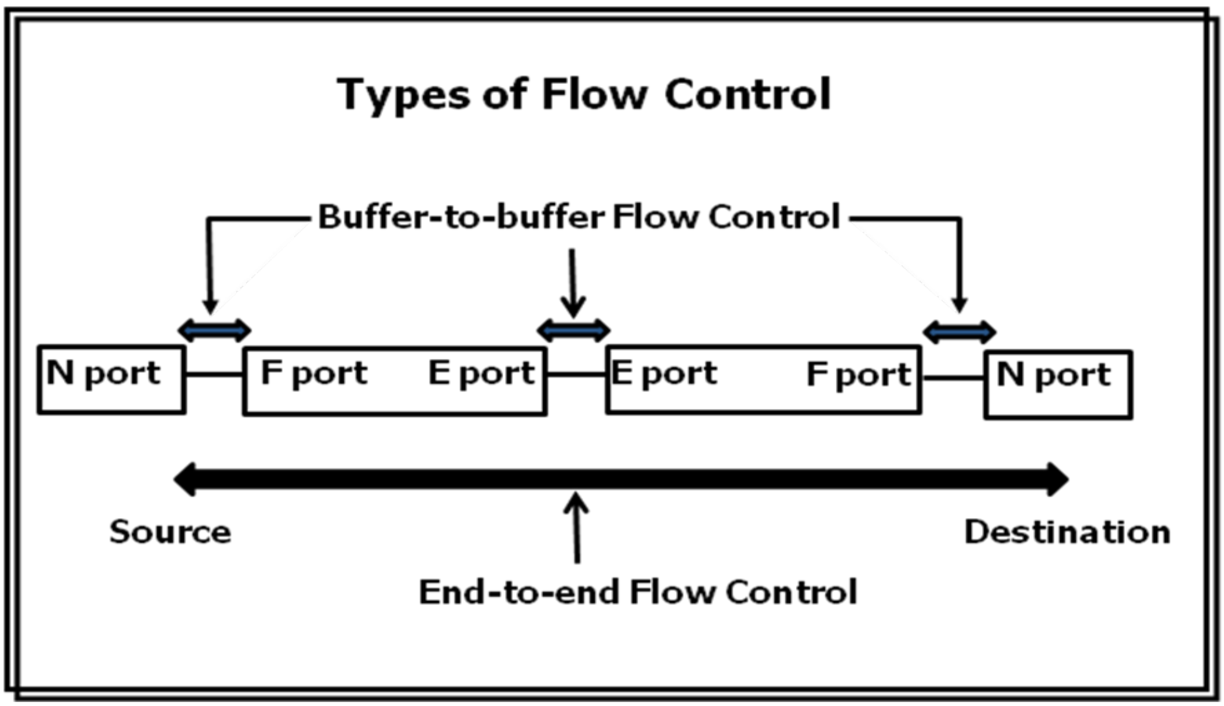
The header includes source and destination port addresses (FCIDs) and control fields. The data payload can range from 0 to 2,112 octets. An SCSI payload is typically 2,048 octets.

The smallest unit of data is a word (4 octets). Words are packaged into frames. An FC frame is similar to an IP packet. A sequence is a series of frames sent from one port to another. An exchange is a series of sequences between two ports to identify a discrete transaction (save, retrieve, on-line transaction process, etc.).

Frame errors can occur at any of the FC layers previously defined.

- Frame errors
- Resource errors
- Delimiter errors
- Delivery errors

SCSI error recovery causes all frames in an errored sequence to be retransmitted, which is a high-overhead solution. Enhanced error recovery provides a more efficient recovery method.



The purpose of flow control is to avoid lost frames of data due to congestion on the links' connecting devices. The problem with most flow-control schemes, such as Ethernet flow control, is that by the time the flow-control process is activated, data already has been lost. FC overcomes this problem with a proactive method where control is in the hands of the receiving port.

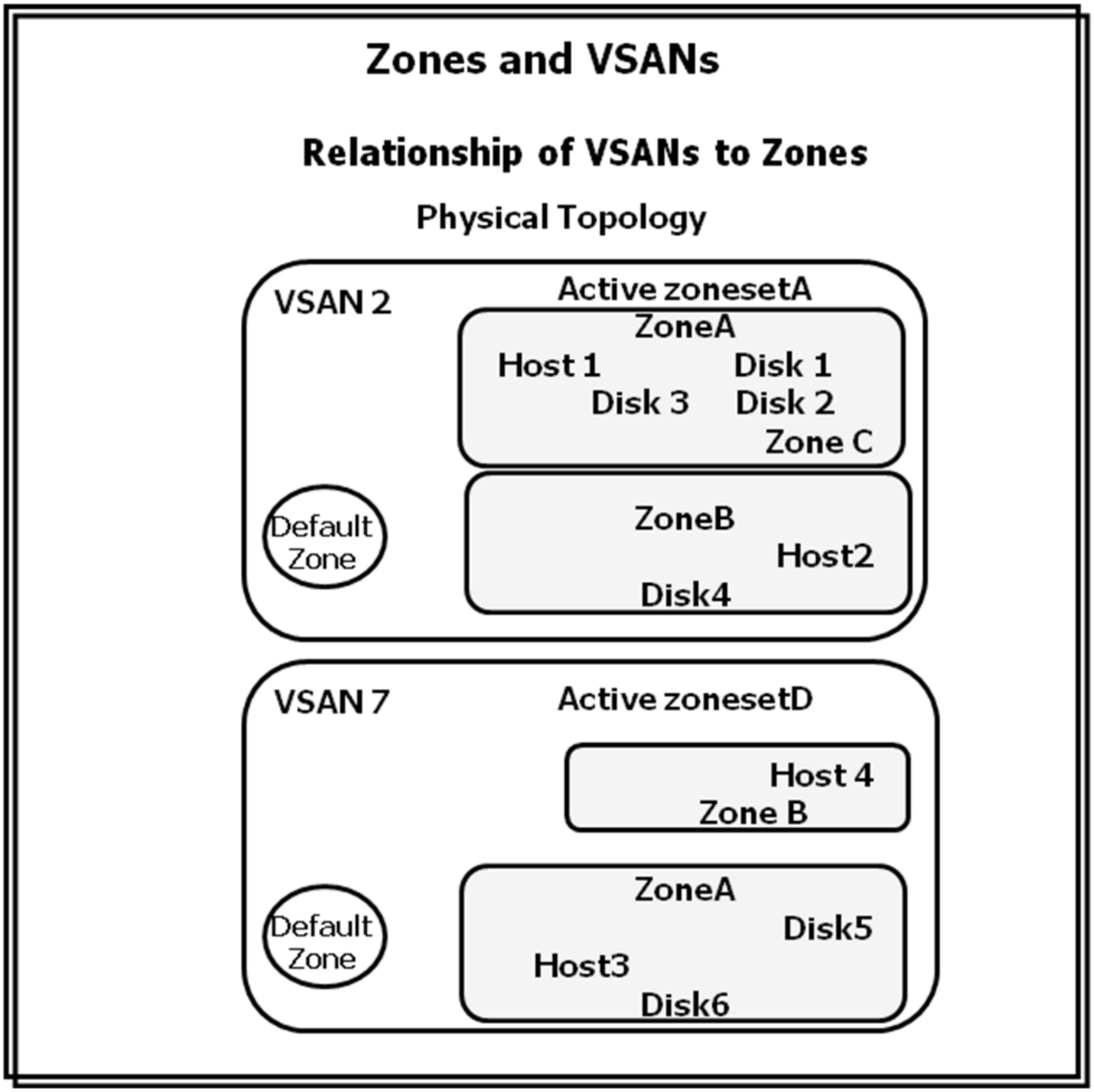
An FC frame cannot be transmitted by the sending port until a buffer credit has been granted by the receiving port. The math to calculate buffer credits is based on the distance between the ports, and the time required to process the frame and get it on the media. Ten credits are required to maintain full bandwidth on a 2 Gbps, 10 km link with 2,000 octet frames. A good rule is one credit per kilometer for 2 Gbps and a 2Kbyte frame.

Buffer-to-buffer flow control is between the ports on one link (hop). End-to-end flow control is between the two ultimate end node ports. Since there is more than one type of data and many devices in a storage network, each class of data has unique requirements. The following table explains FC classes of service:

Classes of Service

Class	Description	Typical Use
1	A dedicated connection between two N ports. End-to-end flow control.	Continuous, time-critical data such as voice or video.
2	Connectionless service with notification of delivery or non-delivery of frames. No dedicated connection is set up. Bandwidth of links is shared. Frames are delivered in order. Uses both buffer-to-buffer and end-to-end flow control.	LAN traffic such as IP, FTP, etc.
3	Similar to Class 2 with no notification of non-delivery of frames. Sometimes called datagram service. Error recovery is responsibility of the Upper Layer Protocol.	SCSI
4	Establishes a unidirectional Virtual Circuit (VC) in each direction between two nodes. Can support Class 1, 2, and 3 traffic. Each VC is regulated separately and Quality of Service may be implemented.	Used in newer SAN devices to support scalability and additional features.
5	Isochronous, just-in-time service – not defined at this time.	Not defined.
6	Similar to class 1 but frames are sent from one N node to numerous N nodes to support multicast.	Multicast

There is a price to pay for the reliability and predictability features of FC. On a 2 Gbps FC link, the maximum data transfer rate is 200 Mbps.



In the previous topic, FC classes of service were defined to allow flow control to meet the requirements of a multitude of storage data types. A similar method is used to define Quality of Service solutions. This idea is also implemented in FC switching devices with Virtual Storage Area Networks (VSANs) and Zones. Zones are further divided into Sub-zones.

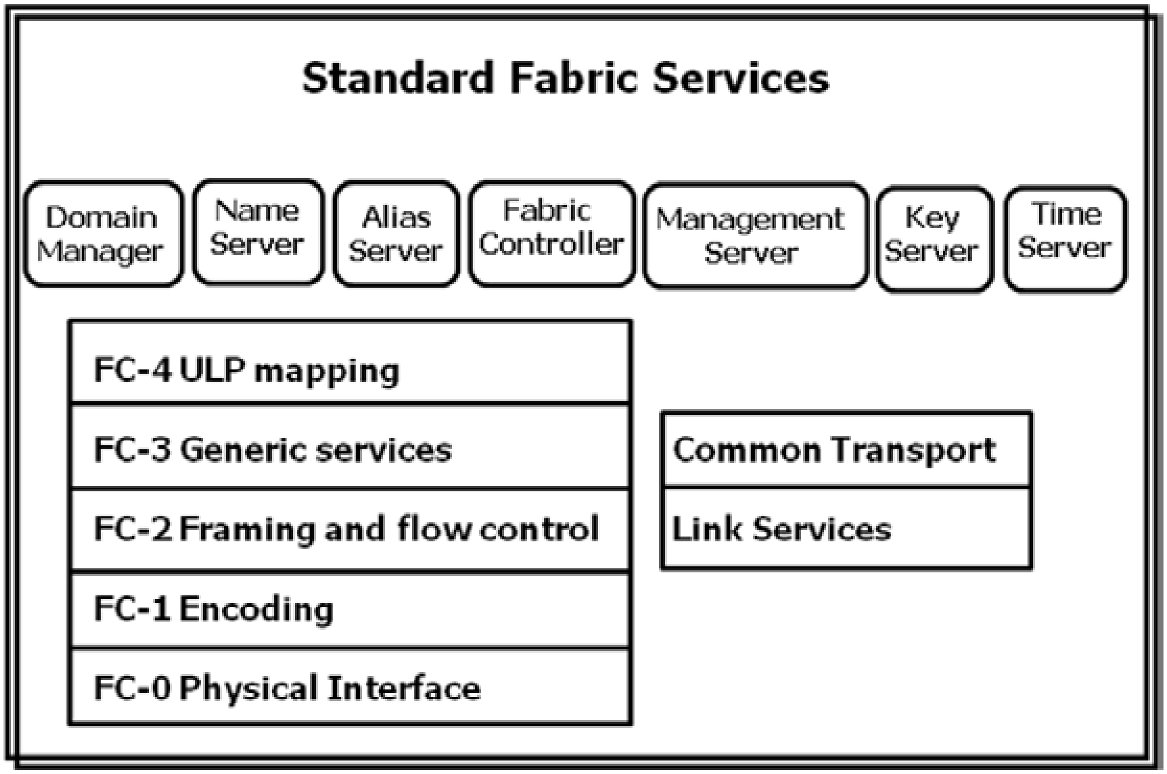
Specific switch ports are defined to specific VSANs based on the device using the port and the type of data being transferred. This is similar to a VLAN (Virtual LAN) in Ethernet. For additional security and sharing of device ports, zones are defined within each VSAN. VSANs are numbered from 1 to 4,095. VSANs and zones are complimentary and could be compared loosely to the functions of a VLAN and Private VLAN in Ethernet switching.

In the previous graphic there are two VSANs: 2 and 7.

- VSAN2
  - ZonesetA
    - ZoneA
      - Host1
      - Disk1
      - Disk1
      - Disk3
    - ZoneB
      - Host2
      - Disk4
    - ZoneC
      - Host2
      - Disk1
      - Disk2

- VSAN7
  - ZonesetD
    - ZoneA
      - Host3
      - Disk5
      - Disk6
    - ZoneD
      - Host4
      - Disk6

Once VSANs are implemented, the frames traversing the ISL links between switches will be "tagged" with a VSAN number similar to VLAN trunking in Ethernet. The ISL link then becomes an Enhanced Inter-Switch Link (EISL).



The previous graphic shows the fundamental FC fabric services provided by fabric switches. Vendors add additional proprietary services as well. The major functions are as follows:

- Domain Manager
    - Allocates domain IDs
    - Allocates port addresses (FCIDs)
    - Participates in switch selection
    - Performs fabric builds and reconfiguration
- Name Server
    - FCIDs
    - pWWNs and nWWNs
    - FC operating parameters
    - Classes of service
    - Maintains database and dynamically updates when changes occur
- Management Server
    - Fabric configuration service
    - Zone service
    - Unzoned name service
    - Read-only

The FC specification does not identify how frames are routed from port to port by the switches in the fabric. Fibre Shortest Path First (FSPF) uses a routing algorithm similar to Open Shortest Path First (OSPF) in IP. It is only one of the methods for efficiently routing FC frames and the details of this function could easily be another paper.

To support additional connectivity requirements in the data center and to extend Fibre Channel over longer distances and diverse networks, FC can be "tunneled" in two ways. First FC frames can be encapsulated in IP, which is Fibre Channel P (FCIP). Secondly, FC can be encapsulated in an Ethernet frame, which is Fibre Channel over Ethernet (FCOE). Just as Inter-VLAN routing is possible with IP and Ethernet. Inter-VSAN routing is possible with Fibre Channel. The name of this feature is Inter-VSAN Routing (IVR).

Conclusion

In summary, Fibre Channel is an ANSI standard-based protocol for the movement of data between servers and storage devices. FC is defined by port addresses and names and each device port connected to a switch port creating a FC Fabric. FC represents the first three layers of the OSI reference model and has a defined frame format and interface with upper layer protocols and drivers. The most common payload for FC frame is SCSI.

Fibre Channel defines a strict flow control process using buffer credits. VSANs, Zones, and Sub-zones are also defined to create data architecture and secure environment. The entire process is controlled by sophisticated fabric switches from Cisco, Brocade, McData and others. These services include domain services, name services, management services, and routing.

Learn More

Learn more about how you can improve productivity, enhance efficiency, and sharpen your competitive edge Check out the following Global Knowledge courses:

Data Center Network infrastructure Design (DCNID)

Designing Cisco Storage Networking Solutions (DCSNS)

For more information or to register, visit www.globalknowledge.com or call 1-800-COURSES to speak with a sales representative. Our courses and enhanced, hands-on labs offer practical skills and tips that you can immediately put to use. Our expert instructors draw upon their experiences to help you understand key concepts and how to apply them to your specific work situation. Choose from our more than 1,200 courses, delivered through Classrooms, e-Learning, and On-site sessions, to meet your IT and business training needs.

About the Author

Ray is a co-founder and CEO of ICM (a Global Knowledge partner) in Redmond, WA. He has been a Global Knowledge Instructor since 1994 and teaches the entire CCNP curriculum, plus advanced classes in MPLS, BGP, ATM, and Cisco Troubleshooting. He has recently become involved in data center and storage network design.